January 2015

# A Variable-Structure Variable-Order Simulation Paradigm for Power Electronic Circuits 

Anandakumar Subbiah<br>Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_dissertations

## Recommended Citation

Subbiah, Anandakumar, "A Variable-Structure Variable-Order Simulation Paradigm for Power Electronic Circuits" (2015). Open Access Dissertations. 1318.
https://docs.lib.purdue.edu/open_access_dissertations/1318

# PURDUE UNIVERSITY <br> GRADUATE SCHOOL <br> Thesis/Dissertation Acceptance 

This is to certify that the thesis/dissertation prepared
By Anandakumar Subbiah
Entitled A Variable-Structure Variable-Order Simulation Paradigm for Power Electronic Circuits

For the degree of Doctor of Philosophy

Is approved by the final examining committee:
OLEG WASYNCZUK

MARYAM SAEEDIFARD

MICHAEL A. CAPANO

SCOTT D. SUDHOFF

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

OLEG WASYNCZUK
Approved by Major Professor(s): $\qquad$


# A VARIABLE-STRUCTURE VARIABLE-ORDER SIMULATION PARADIGM 

 FOR POWER ELECTRONIC CIRCUITSA Dissertation<br>Submitted to the Faculty<br>of<br>Purdue University<br>by<br>Anandakumar Subbiah<br>In Partial Fulfillment of the<br>Requirements for the Degree<br>of<br>Doctor of Philosophy

December 2015

Purdue University
West Lafayette, Indiana

To my family

## ACKNOWLEDGMENTS

I sincerely thank Dr. Oleg Wasynczuk for his patient advising, introducing me to computational science and its use as a modeling tool in the power and energy area. His comments and constant encouragement helped me shape my research career. I thank Drs. Scott Sudhoff, Michael Capano and Maryam Saeedifard for serving in my advisory committee. I thank Dr. Sudhoff for suggesting that I take Solid State Devices (ECE 606), which helped me gain more insight about semiconductor fundamentals. I thank the Power and Energy Devices and Systems faculty for providing a pleasant environment throughout my stay here at Purdue. Last but not the least, I thank my family for their love, care and support.

## TABLE OF CONTENTS

Page
LIST OF TABLES ..... vii
LIST OF FIGURES ..... viii
ABSTRACT ..... x
1 INTRODUCTION ..... 1
1.1 Literature Survey ..... 2
1.2 Motivation ..... 5
1.3 Outline of Document ..... 6
2 ELECTRONIC TRANSPORT MODEL ..... 7
2.1 One Dimensional Diode ..... 7
2.2 Ionization Model ..... 8
2.3 Poisson's Equation ..... 9
2.4 Recombination Model ..... 10
2.5 Continuity Equation ..... 12
2.6 Coupled Carrier Transport Equations ..... 13
2.7 Chapter Summary ..... 16
3 DISCRETIZATION AND MODEL ENCAPSULATION ..... 17
3.1 Abstract Discretization ..... 17
3.2 Poisson's Equation ..... 19
3.3 Continuity Equations ..... 22
3.3.1 Central difference method ..... 23
3.3.2 Scharfetter-Gummel method ..... 26
3.3.3 Boundary charge densities ..... 31
3.4 Temporal Discretization ..... 31
3.5 Encapsulation ..... 33
Page
3.6 Chapter Summary ..... 36
4 COUPLED DEVICE AND ELECTRICAL NETWORK SIMULATION ..... 37
4.1 Diode Resistor Inductor Network ..... 37
4.1.1 Models of basic circuit elements ..... 44
4.2 Coupled Circuit and Device Simulation ..... 45
4.2.1 Interpretation of interface matrices ..... 49
4.3 Comparison with Other Formulations ..... 53
4.4 Chapter Summary ..... 54
5 TEMPORAL INTEGRATION AND PARAMETER EXTRACTION ..... 56
5.1 Backward Euler Temporal Integration ..... 56
5.2 Parameter Extraction ..... 60
5.3 Diode Resistor Inductor Circuit Example ..... 65
5.4 Chapter Summary ..... 69
6 EXPERIMENTAL VALIDATION ..... 70
6.1 Full System Simulation ..... 70
6.1.1 Computational performance ..... 73
6.1.2 Step size and sample rate ..... 74
6.1.3 ac bus waveforms ..... 76
6.1.4 dc bus waveforms ..... 77
6.1.5 Diode 1 variables ..... 78
6.1.6 Power and energy calculations ..... 79
6.2 Chapter Summary ..... 82
7 VARIABLE-STRUCTURE VARIABLE-ORDER SIMULATION PARADIGM ..... 83
7.1 Variable-Structure Variable-Order Paradigm ..... 83
7.1.1 Fixed graded mesh ..... 85
7.1.2 Linear solver ..... 86
7.1.3 Structurally varying tableau ..... 94
7.2 Single-Phase Bridge Rectifier Simulation ..... 98
Page
7.3 VSVO Simulation Results ..... 101
7.4 Speed-up calculations ..... 105
7.5 Chapter Summary ..... 108
8 SUMMARY, CONCLUSIONS, AND FUTURE RESEARCH ..... 110
8.1 Future Work ..... 112
8.1.1 Multi-dimensional/multi-terminal devices ..... 112
8.1.2 Coupled Electro-thermal modeling ..... 113
8.1.3 Numerical aspects ..... 113
REFERENCES ..... 115
A DERIVATIONS IN ONE DIMENSION ..... 123
A. 1 Poisson's Equation ..... 123
A. 2 Derivation of Jacobian Elements ..... 126
A.2.1 Poisson's equation ..... 127
A.2.2 Central difference method ..... 128
A.2.3 Scharfetter-Gummel method ..... 130
A.2.4 Newton-Raphson iterator ..... 133
A.2.5 $\quad \mathbf{f}_{\mathbf{i}, \mathrm{X}}$ matrices ..... 134
B MANUAL TABLEAU DERIVATION ..... 138
B. 1 Appendix Summary ..... 142
C Adaptive Spatial Mesh ..... 144
C. 1 Appendix Summary ..... 147
D COMPUTER SPECIFICATIONS ..... 148
VITA ..... 150

## LIST OF TABLES

Table Page
2.1 Ionization parameters at $T=300 \mathrm{~K}$. ..... 10
2.2 Recombination parameters for Si . ..... 11
2.3 Drift-diffusion parameters. ..... 13
3.1 Parameters of an example $p n$-diode with step junction. ..... 25
3.2 Bernoulli function values. ..... 29
4.1 Resistor-inductor-diode netlist. ..... 38
4.2 Incidence matrix partitions. ..... 43
5.1 Estimated Circuit parameters. ..... 63
5.2 Doping density versus forward voltage drop $V_{F}$. ..... 64
5.3 Extracted physical parameters of S1A PIN diode. ..... 65
6.1 Single-phase diode bridge rectifier circuit parameters. ..... 71
6.2 Single-phase diode bridge rectifier netlist. ..... 71
6.3 Profiler output listing computational bottlenecks. ..... 74
6.4 Simulated and measured energy comparisons. ..... 82
7.1 Average time for solving $224 \times 224$ linear system (smaller the better). ..... 88
7.2 Performance comparison of 10000 solves of a linear system of 882 equa- tions. ..... 92
7.3 Simulated and measured energy comparisons. ..... 105
7.4 VSVO profile results. ..... 106
7.5 Estimated speed-up. ..... 108

## LIST OF FIGURES

Figure Page
2.1 One-dimensional PIN diode structure. ..... 8
3.1 Finite volume discretization (top) and a prototypical volume (bottom). ..... 23
3.2 Charge densities and Péclet number near $p n$-junction using central differ- ence method. ..... 26
3.3 Charge densities and Péclet number near $p n$-junction using Scharfetter- Gummel method. ..... 30
4.1 A simple $R L$ and diode circuit with no ground node. ..... 37
4.2 $R L$ and diode circuit with parasitic capacitors and ground nodes added. ..... 39
4.3 Circuit-diode coupling. ..... 40
4.4 Directed graph of the circuit in Figure 4.2. ..... 40
4.5 Single-phase diode bridge rectifier with resistor-inductor load. ..... 52
5.1 Backward Euler algorithm flowchart. ..... 58
5.2 Simulation algorithm flowchart. ..... 61
5.3 Circuit used for diode characterization. ..... 62
5.4 Simulated voltage $v_{d}$ across (top) and current $i_{d}$ (bottom) through the sample diode. ..... 66
5.5 comparison of I-V characteristics of the diode. ..... 67
5.6 PIN diode switching current (top) and voltage (bottom) dynamics. ..... 68
6.1 Single-phase diode bridge rectifier. ..... 70
6.2 Incidence graph for the circuit with parasitic capacitors and ground node added. ..... 72
6.3 Measured and simulated step-size comparison. ..... 75
6.4 Step-size comparison zoomed near $5 \times 10^{-4}$ s ..... 76
6.5 Voltage (top) and current (bottom) waveforms on ac-side. ..... 77
6.6 Voltage (top) and current (bottom) waveforms on dc-side. ..... 78
Figure Page
6.7 Diode voltage (top) and current (bottom) waveforms. ..... 79
6.8 Diode voltage (top) and current (bottom) waveforms near $4 \times 10^{-4}$ and $5 \times 10^{-4} \mathrm{~s}$, respectively ..... 80
6.9 Diode instantaneous power loss waveform ..... 80
6.10 Instantaneous power waveforms on ac (top) and dc (bottom) sides. ..... 81
7.1 Four Newton iterator matrices are sampled at the marked instances on current and voltage response of a diode in RLD network. ..... 87
7.2 Eigen spectrum of respective Newton iterator matrices in clockwise start- ing at top left for operating points marked left to right as in Figure 7.1. ..... 88
7.3 Qualitative figure to explain BGE solver's performance. ..... 93
7.4 Diode activity detector output for the example $R L D$ network. ..... 97
7.5 Diode switch on-off states versus time. ..... 99
7.6 Diode 1 current and its on-off state versus time. ..... 100
7.7 Instantaneous power waveforms of Diode 1 using VSVO. ..... 102
7.8 Voltage (top) and current (bottom) waveforms of Diode 1 comparison. ..... 103
7.9 Zoomed voltage and current waveforms of Diode 1 using VSVO. ..... 103
7.10 Voltage (top) and current (bottom) waveforms on ac-side. ..... 104
7.11 Voltage (top) and current (bottom) waveforms on dc-side. ..... 104
7.12 Thread utilization with 3 threads for 2 and 4 -diodes on cases. ..... 108
A. 1 One-dimensional discretization. ..... 123
A. 2 Canonical element. ..... 124
B. 1 A simple $R L$ and diode circuit. ..... 138
C. 1 Charge densities computed on adaptive mesh. ..... 144
C. 2 Number of nodes (\# Nodes) versus time-step index for adaptive mesh. ..... 146


#### Abstract

Subbiah, Anandakumar Ph.D., Purdue University, December 2015. A VariableStructure Variable-Order Simulation Paradigm for Power Electronic Circuits. Major Professor: Oleg Wasynczuk.

Solid-state power converters are used in a rapidly growing number of applications including variable-speed motor drives for hybrid electric vehicles and industrial applications, battery energy storage systems, and for interfacing renewable energy sources and controlling power flow in electric power systems. The desire for higher power densities and improved efficiencies necessitates the accurate prediction of switching transients and losses that, historically, have been categorized as conduction and switching losses. In the vast majority of analyses, the power semiconductors (diodes, transistors) are represented using simplified or empirical models. Conduction losses are calculated as the product of circuit-dependent currents and on-state voltage drops. Switching losses are estimated using approximate voltage-current waveforms with empirically derived turn-on and turn-off times. With recent increases in switching speeds, these approximations are no longer valid in many applications. Although it is possible to simulate power converters using physics-based models of power semiconductors based upon coupled drift, diffusion, continuity (CDDC) equations, such simulations are generally prohibitively slow. In this thesis, a variable-structure variable-order simulation paradigm is set forth in which the detailed CCDC-based models are used to calculate the switching transients and corresponding losses. As devices (e.g., diodes) become active or inactive, the structure and order of the simulation is dynamically changed without sacrificing accuracy. A time-step control algorithm is devised such that the overall simulation captures only the relevant transients. Finally, paralleliza-


tion strategies are identified that can produce a $483 \%$ improvement in simulation speed compared with a conventional solution of the CDDC equations.

## 1. INTRODUCTION

An important consideration in the design of power electronic converters are the losses that are produced over their entire operating range. The losses not only affect the efficiency of the converter, but also its physical size, weight, and ultimately cost, due to the necessity to channel the losses away from the devices and limit the temperatures to safe operating values. In the vast majority of simulations and analyses, the power semiconductors (diodes, transistors) are represented using highly simplified or empirical models. Typically, the losses are categorized as either switching or conduction. Conduction losses are calculated as the product of circuit-dependent currents times on-state voltage drops for Insulated-Gate Bipolar Transistors (IGBTs) and as $i^{2} R_{D S, o n}$ in Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs). Switching losses are estimated using approximate voltage-current waveforms with empirically-derived turn-on and turn-off times established using "test" circuits.

Current trends to improve efficiency and reduce the size and weight of power electronic converters have lead to higher switching frequencies and correspondingly faster devices. The assumptions previously made to estimate switching and conduction losses are no longer always valid. Attempts have been made to improve the accuracy of simulations by developing more detailed circuit models similar to those used in SPICE [1]. However, numerous assumptions are also made in their derivation that may limit the scope or applicability of the resulting model.

A completely different class of simulators such as Medici [2], which uses physical properties (e.g., geometry, doping levels, lifetimes), is typically used when designing the power semiconductor devices that are used in converter circuits. Therein, semiconductor devices are modeled at a much higher level of detail based upon their physical properties and processes that take place within the device such as ionization, recombination, diffusion, and drift due to electric fields. These provide a much more
detailed representation of the devices; however, when numerous devices are interconnected to form a complete circuit, the resulting simulations are generally very slow. The primary objective of this thesis is to establish a simulation paradigm in which the converter losses (switching and conduction) can be accurately and more rapidly established at the circuit level, based upon the physical (non-empirical) parameters of the semiconductor devices utilized within the circuit.

### 1.1 Literature Survey

Simulation of the physical phenomena in devices and power electronic circuits is an active area of research. Developments in computing can be used for advantage in this type of simulation. The following literature survey includes an overview of past and recent research in power semiconductor device and power electronic circuit simulation.

The currents in semiconductors are mainly due to two types of charge carriers namely electrons and holes. Based upon statistics for semiconductors, many-particle and short-range-particle interactions, the Boltzmann Equation can be derived [3]. Solving Boltzmann equations analytically and numerically is prohibitively expensive. A simpler model involving three spatial coordinates and time can be derived. The zeroth order moment of the Boltzmann equation and the adjoining Poisson's equation yields the so-called drift diffusion (DD) equations [4,5]. Van Roosbroeck first derived the DD equations in [6]. This system of DD, Poisson's, and current density equations together describe averaged physical quantities of the carrier densities in a semiconductor device. Electron density, hole density, and the electric potential are the variables in this model, which forms a strongly-coupled partial differential equation (PDE) system.

In this section, it is sufficient to analyze the type of PDEs in this model. The model employs two parabolic PDEs to describe the rate-of-change of charge carriers (continuity equations), and an elliptic PDE that describes the electrostatic potential (Pois-
son's equation). Appropriate boundary conditions and two equations describing the electron and hole current densities in the device completely model the device physics. A proof of existence and uniqueness for these PDEs is desirable before attempting their numerical solution. Zláal and Miloš [7] provided this proof for the solution of the DD equations. A parabolic PDE equation/system with a unique solution can be solved numerically using different spatial-temporal discretization techniques [8]. A space-time-dependent PDE discretized both spatially and temporally is often referred to as a fully discretized system. The discretizing techniques and solution of the resulting equations are discussed in [9] and summarized here for reference.

Spatial discretization of a parabolic PDE system will result in a system of ordinary differential equations (ODEs). The resulting ODEs can be solved by discretizing them in time using one of the many available integration algorithms and solving for the variables at each time step. These algorithms are available in standard computer applications such as MATLAB [10] or DASSL [11], to name a few. This technique is called the method of lines (MOL) and a rich mathematical literature is available on this method due to its popularity. When the order of spatial and temporal discretization is reversed, i.e. temporal followed by spatial, the result is a boundary-value problem that is solved at each time step. This method is called Rothe's method, which facilitates implementation of adaptive grids wherein the PDE solver uses different spatial grids at each time step. Rothe's method is relatively new and only a limited literature is available on this method [12]. When the same fixed spatial grid is used in the MOL and Rothe's methods, then the methods and solutions are equivalent. The space-time finite element method, another way of spatial temporal discretization, is equivalent to Rothe's method with mathematically sound spatial-temporal step-size control [8].

A numerical solution of the quantities of interest is sought using the fully discretized PDEs. Gummel [13] introduced an iterative procedure to decouple and solve the linear system resulting from discretization of a 1-D transistor. This approach can be treated as nonlinear Gauss-Seidel method and is used in [14-17]. This it-
erative technique can be applied when the computational resources are limited or the discretized system is very large, which is the case for higher (2 or 3 ) dimensions [15]. Scharfetter and Gummel in [18] introduced an exponentially-weighted finite difference scheme for the continuity equation, known as Scharfetter-Gummel discretization. This approach is widely used in the numerical simulation of semiconductor devices $[4,14,16,19]$. This exponential-weighting has a stabilizing effect on the discretized system. These two seminal works $[13,18]$ of Gummel have spawned a large number of papers on semiconductor simulation to date.

Finite difference methods can be used for spatial discretization of DD equations [19-26]. Finite element methods are an alternative as described in $[4,14,15$, $17,24,26-33]$. The use of finite difference and finite element methods is described here in a broad sense. In practical implementations, Poisson's equation alone can be discretized using a finite element approach $[14,17,22,28,31,32]$ while finite difference [14] or finite element methods [17, 22, 28, 31, 32] can be used for the continuity equations. Alternatively, a hybrid scheme is described in [34], which uses ScharfetterGummel discretization embedded in the finite element method. It is apparent that the majority of device simulation literature uses a MOL approach to solve the PDEs. The charge carrier concentration and electric potential variables are decoupled as in [13] and solved iteratively or the entire system is solved using Newton iterations [22].

The solution of the DD equations changes only in specific regions of the semiconductor. Even for stationary DD equations, a spatially adaptive grid will be useful [35] where the carrier densities or electric potential change drastically. This phenomenon is termed spatial stiffness. Additionally, the temporal variations of these variables in the unstationary DD equations span several orders of magnitude. This phenomenon is termed temporal stiffness. The unstationary or dynamic problem has both spatial and temporal stiffness. It is intuitive to use adaptation in both space and time [24] in this case. Most of the adaptive mesh implementations up to date were developed for the stationary DD equations [35, 36]. The work by Burgler et al [37] start with stationary DD equations, use FEM for discretizing Poisson's equation and a
divergence-free upwinding discretization scheme for the continuity equations. Then, a posteriori error [38], estimated based on electrical potential and terminal currents, is used to refine or coarsen the mesh.

Simpler models can be derived with the help of numerical solutions of the DD equations and the resulting data. A good survey on the models briefly discussed here is in $[39,40]$. Many of the models resort to simpler version of DD equations known as ambipolar diffusion equation (ADE) [41-45] or other simpler differential equations $[46,47]$ because of the computationally expensive physics-based models. A similar method involves breaking the device into regions of interest and solving chargecontrol equations [43] or model as lumped charges [48, 49]. Parameter extraction via experimental data $[49,50]$ or simulation of ADE [42], etc. and curve fitting the data obtained therefrom can be used to develop empirical models. Another way to model is to augment the small-signal model of the semiconductor devices with subcircuits $[41,43]$ to closely depict the switching transients. Even though the previous models can be used to estimate the switching losses, only a few [51] actually try to do this. Others, which do not rely on simulation, use a three-component power loss model with turn-on, conduction, and turn-off losses. Analytical expressions are used to approximate these components and hence the total loss [52-54]. Apart from estimating the losses in the device, simulating the transients during turn-on and turnoff is also of interest [49].

### 1.2 Motivation

There is significant interest in simulating the device accurately when used in a power circuit. Of the many reasons for simulation, a few are to gain confidence in the design, to estimate the losses in the device, to predict the transients during switching and mitigate them when they affect the performance. There are different levels of device models to achieve one or more of these objectives. Physics-based models meet most of these objectives but require large computational resources. Simpler empirical
models, when they do not meet some of these objectives, require modification with ad hoc networks. However, these models can be integrated with commercially available simulation environments and are fast. Other models that are somewhere in between these two extremes are based on certain assumptions that limit their applicability.

As noted previously, the primary objective of this thesis is to establish a simulation paradigm in which the converter losses (switching and conduction) can be accurately and more rapidly established at the circuit level based upon the physical (non-empirical) parameters of the semiconductor devices utilized within the circuit. The proposed approach is based upon developing a computationally efficient numerical solution of the coupled DD, continuity, and Poisson's equations of the devices, and the algebraic and differential equations corresponding to the external circuit elements and their interconnection.

### 1.3 Outline of Document

The outline of this document is as follows. The semiconductor charge transport model is reviewed in Chapter 2 including models of the relevant physical processes such as ionization, electrostatics, charge transport, charge recombination, and boundary equations. Discretization of the resulting PDEs (temporal and spatial) is described in Chapter 3 wherein an encapsulated device model is set forth. A procedure to couple these encapsulated device models with the circuit equations to result in a set of differential-algebraic equations (DAEs) is established in Chapter 4. The temporal integration of the DAEs and extraction of the physical parameters of a commercial PIN diode is described in Chapter 5. Simulation of a single-phase diode bridge rectifier circuit is validated by experimental measurements in Chapter 6. A new variable-structure variable-order (VSVO) simulation paradigm is set forth in Chapter 7 and simulation results are compared with validated complete-system-simulation results. The major conclusions, strategies to simulate similar problems, and potential future research topics are identified in Chapter 8.

## 2. ELECTRONIC TRANSPORT MODEL

A detailed model of the semiconductor diode is used in this research for power-loss estimation. To this end, the semiconductor charge transport model is reviewed. Models of the relevant physical processes including ionization, electrostatics, charge transport, and charge recombination are discussed in the subsequent sections. All of the corresponding equations are combined into one set whose consistent numerical solution can be used to describe the voltages and currents both within the device and at its terminals.

### 2.1 One Dimensional Diode

A PIN diode is considered in this research which is extensively used in highpower circuits. They are typically used in rectifiers (ac to dc converters) and as freewheeling diodes in other types of converters. The general physical structure of a semiconductor PIN diode is shown in Figure 2.1 (top). The anode and cathode ohmic contacts are at the coordinates $x=x_{a}=0$ and $x=x_{c}=X$ respectively. The ohmic contacts are assumed to be placed in the neutral regions of the semiconductor. These contacts form semiconductor-metal junctions that provide electrical connection with the external electrical circuit. A good ohmic contact exhibits a negligible voltage drop and resistance compared to the voltage drop and bulk resistance of the diode, respectively.

In this diode, an intrinsic semiconductor region $i$ is sandwiched between highly doped $p^{+}$and $n^{+}$regions and hence the name PIN diode. Two semiconductorsemiconductor junctions, $p^{+}-i$ and $i-n^{+}$, are formed as a result of this arrangement. The bottom of Figure 2.1 shows the doping profile of a typical PIN diode. The quantity $\log N$ is the logarithm of absolute sum of the acceptor $\left(N_{A}\right)$ and donor $\left(N_{D}\right)$


$$
\log N=\log \left(N_{A}+N_{D}\right)
$$



Fig. 2.1.: One-dimensional PIN diode structure.
doping densities. Over the distance $X_{j p}$ the acceptor density decreases smoothly from a large number of the order of $10^{18}$ to $10^{16}$. At $x=X_{j p}$ there is step junction formed by $p^{+}$and $i$ regions. The intrinsic region extends over a distance of $W_{d}$. The doping profile increases over the distance $X_{j n}$ starting from $x=X_{j p}+W_{d}$. The terminal behavior of the diode can usually be satisfactorily described by a one-dimensional model. The quantities of interest such as terminal currents can be found by appropriately scaling terminal current densities by the device cross-sectional area. Some of the relevant semiconductor processes are reviewed and modeled in the subsequent sections.

### 2.2 Ionization Model

A semiconductor specimen is called as an intrinsic semiconductor when the amount of impurity atoms in it is insignificant. The number of electrons and holes in such a specimen are equal. When impurity atoms are added to the intrinsic semiconductors,
this balance is shifted and results in an extrinsic semiconductor. Silicon belongs to Group IV in the periodic table. Elements from Group III or V is added to the intrinsic semiconductor to produce $p$ or $n$-type extrinsic semiconductors. The primary $n$-type dopant in Si is phosphorus from Group V while aluminum from Group III is the primary $p$-type dopant. The ionization energy of both dopants are approximately 45 meV . The ionization model may be approximated as [55],

$$
\begin{equation*}
n=\frac{N_{D}}{1+\frac{g_{D}}{N_{C}} e^{\frac{E_{D}}{k T}}}, \tag{2.1}
\end{equation*}
$$

wherein $E_{D}$ is the donor ionization energy and $g_{D}$ is the donor-site degeneracy factor. The effective density of states at the conduction band edge may be expressed as,

$$
\begin{equation*}
N_{C}=5.3886 \times 10^{15} \times T^{3 / 2} \mathrm{~cm}^{-3} . \tag{2.2}
\end{equation*}
$$

In (2.1), $n$ appears on both sides of the equation. It is possible to solve for $n$ in terms of $T$ even if (2.1) is implicit. A similar expression for the $p$-type ionization model [55] is,

$$
\begin{equation*}
p=\frac{N_{A}}{1+\frac{g_{A}}{N_{C}} e^{\frac{E_{A}}{k T}}}, \tag{2.3}
\end{equation*}
$$

where $E_{A}$ is the acceptor ionization energy and $g_{A}$ is the acceptor-site degeneracy factor. The effective density of states $\left(N_{V}\right)$ at the valence band edge may be expressed as,

$$
\begin{equation*}
N_{V}=2.0015 \times 10^{15} \times T^{3 / 2} \mathrm{~cm}^{-3} . \tag{2.4}
\end{equation*}
$$

The ionization parameters for the Si material at a temperature of 300 K are summarized in Table 2.1. At $T=300 \mathrm{~K}$ the effective density of states at the conduction and valence bands are as in [56]. Based on this number, a temperature dependency is included as in (2.2) and (2.4).

### 2.3 Poisson's Equation

The charge distribution within the semiconductor device obeys the Gauss's law. According to Gauss's law, the divergence of the electric field should equal the charge

Table 2.1.: Ionization parameters at $T=300 \mathrm{~K}$.

| $g_{D}$ | $n$-type Degeneracy factor | 2 |
| :--- | :--- | :--- |
| $g_{A}$ | $p$-type Degeneracy factor | 4 |
| $E_{D}$ | Ionization energy (Phosphorus) | 45 meV |
| $E_{A}$ | Ionization energy (Aluminum) | 44.39 meV |

enclosed. By observing that the electric field is the negative gradient of the electric potential, it is rewritten as Poisson's equation. The right-hand side $\rho$ describes the charge distribution within the body of the semiconductor. The general form of the Poisson's equation may be expressed as,

$$
\begin{equation*}
-\nabla \cdot(\varepsilon \nabla \psi)=\rho \tag{2.5}
\end{equation*}
$$

where $\psi$ is the intrinsic electric potential, $\rho$ is the net electric charge density, $\varepsilon$ is the permittivity. In one dimension,

$$
\begin{equation*}
\varepsilon_{r} \varepsilon_{0} \frac{d^{2} \psi}{d x^{2}}=-q_{e}\left(p(x)-n(x)+N_{D}^{+}(x)-N_{A}^{-}(x)\right) \tag{2.6}
\end{equation*}
$$

where $q_{e}$ is the electron charge, $\varepsilon_{r}$ is the relative permittivity of the semiconductor material, $\varepsilon_{0}$ is the permittivity of the free space, $p$ and $n$ are the mobile hole and electron densities, respectively, $N_{D}^{+}$and $N_{A}^{-}$are the density of immobile ionized donor and acceptor ions, respectively, which are established from the ionization equations (2.1) and (2.3). It is assumed here that the permittivity of the semiconductor is constant throughout the material and hence $\varepsilon=\varepsilon_{r} \varepsilon_{0}$ is factored out.

### 2.4 Recombination Model

When the semiconductor is perturbed from its equilibrium state, certain processes occur in order to restore that equilibrium. Recombination-Generation (R-G) is such
a process where charges are annihilated or created, respectively. Some of the other recombination processes are recombination via shallow level, recombination involving excitons, Auger recombination and band-to-band recombination [55]. Of these, only the dominant R-G or SRH (Shockley, Read, Hall) recombination generation process is modeled. The general form for the recombination rate $\mathcal{R}$ is,

$$
\begin{equation*}
\mathcal{R}=r_{n}=r_{p}=\frac{n(x, t) p(x, t)-n_{i e}^{2}}{\tau_{p}\left(n+n_{1}\right)+\tau_{n}\left(p+p_{1}\right)}, \tag{2.7}
\end{equation*}
$$

where,

$$
\begin{align*}
& r_{p}=-\left.\frac{\partial p(x, t)}{\partial t}\right|_{\mathrm{R}-\mathrm{G}},  \tag{2.8}\\
& r_{n}=-\left.\frac{\partial n(x, t)}{\partial t}\right|_{\mathrm{R}-\mathrm{G}}, \tag{2.9}
\end{align*}
$$

and $n_{i e}$ is the intrinsic electron concentration,

$$
\begin{equation*}
n_{i e}=\sqrt{N_{C} N_{V} e^{-E_{g} / 2 k T}} . \tag{2.10}
\end{equation*}
$$

Also, $\tau_{n}$ and $\tau_{p}$ are time constants which depend on trap density $\left(N_{T}\right)$ and capture coefficients $\left(c_{p}\right.$ or $\left.c_{n}\right)$. Specifically,

$$
\begin{align*}
\tau_{p} & =\frac{1}{c_{p} N_{T}},  \tag{2.11a}\\
\tau_{n} & =\frac{1}{c_{n} N_{T}}, \tag{2.11b}
\end{align*}
$$

Table 2.2.: Recombination parameters for Si .

| Symbol | Parameter | Value |
| :--- | :--- | :--- |
| $\tau_{p}$ | Time constant | $10^{-4} \mathrm{~s}$ |
| $\tau_{n}$ | Time constant | $10^{-4} \mathrm{~s}$ |
| $E_{g}$ | Gap energy | 1.11 eV |

and,

$$
\begin{align*}
& p_{1}=n_{i e} e^{\left(E_{i}-E_{T}\right) / k T},  \tag{2.12a}\\
& n_{1}=n_{i e} e^{\left(E_{T}-E_{i}\right) / k T}, \tag{2.12b}
\end{align*}
$$

where $E_{T}$ is the trap density and $E_{i}$ is the intrinsic level (near mid-gap). It is assumed that $E_{T} \approx E_{i}$ whereupon $p_{1}=n_{1}=n_{i e}$. The recombination parameters are summarized in Table 2.2.

### 2.5 Continuity Equation

The rate of change of charges within a small volume of the semiconductor, apart from recombination, depends on the drift and diffusion phenomena. The drift phenomenon is the response or the transport of charges ( $p$ or $n$ ) due to the electric field in the device. The diffusion phenomenon is the response or the transport of charges due to the distribution gradient of the respective charges i.e. transport of charges from the regions of higher concentration to the lower ones. The continuity equation can be expressed,

$$
\begin{align*}
\frac{\partial n}{\partial t} & =\left.\frac{\partial n}{\partial t}\right|_{\mathrm{drift}}+\left.\frac{\partial n}{\partial t}\right|_{\text {diffusion }}+\left.\frac{\partial n}{\partial t}\right|_{\mathrm{R}-\mathrm{G}}  \tag{2.13}\\
\frac{\partial n}{\partial t} & =\frac{1}{q_{e}} \nabla \cdot J_{n}+\left.\frac{\partial n}{\partial t}\right|_{\mathrm{R}-\mathrm{G}} \tag{2.14}
\end{align*}
$$

where $\frac{1}{q_{e}} \nabla \cdot J_{n}$ represents the net influx of electrons per unit time, $\left.\frac{\partial n}{\partial t}\right|_{\mathrm{R}-\mathrm{G}}$ is the generation rate due to thermal ionization less the rate of recombination. In one dimension,

$$
\begin{equation*}
\frac{\partial n(x, t)}{\partial t}=\frac{1}{q_{e}} \frac{d J_{n}(x, t)}{d x}+\left.\frac{\partial n(x, t)}{\partial x}\right|_{\mathrm{R}-\mathrm{G}} . \tag{2.15}
\end{equation*}
$$

A similar relationship applies to holes,

$$
\begin{equation*}
\frac{\partial p(x, t)}{\partial t}=-\frac{1}{q_{e}} \frac{d J_{p}(x, t)}{d x}+\left.\frac{\partial p(x, t)}{\partial x}\right|_{\mathrm{R}-\mathrm{G}} \tag{2.16}
\end{equation*}
$$

The current densities $J_{n}$ and $J_{p}$ are sum of the drift and diffusion currents. They are expressed as,

$$
\begin{align*}
J_{n}(x, t) & =q_{e} \mu_{n} n(x, t) \mathcal{E}(x, t)+q_{e} D_{p} \frac{\partial n(x, t)}{\partial x}  \tag{2.17a}\\
J_{p}(x, t) & =q_{e} \mu_{p} p(x, t) \mathcal{E}(x, t)-q_{e} D_{p} \frac{\partial p(x, t)}{\partial x} \tag{2.17b}
\end{align*}
$$

The first term on the right-hand side of each equation represents the drift current while the other term represents the diffusion current. The drift/diffusion parameters [57] for Si corresponding to a temperature of 300 K are summarized in Table 2.3

Table 2.3.: Drift-diffusion parameters.

| Symbol | Parameter | Value |
| :--- | :--- | :--- |
| $\mu_{n}$ | Electron mobility | $973.90 \mathrm{~cm}^{2} / \mathrm{V}-\mathrm{s}$ |
| $\mu_{p}$ | Hole mobility | $495.00 \mathrm{~cm}^{2} / \mathrm{V}-\mathrm{s}$ |
| $D_{n}=\mu_{n} k T / q_{e}$ | Diffusion constant | $35.224 \mathrm{~cm}^{2} / \mathrm{s}$ |
| $D_{p}=\mu_{p} k T / q_{e}$ | Diffusion constant | $12.820 \mathrm{~cm}^{2} / \mathrm{s}$ |

### 2.6 Coupled Carrier Transport Equations

$$
\begin{align*}
\frac{\partial n(x, t)}{\partial t} & =\frac{1}{q_{e}} \frac{d J_{n}(x, t)}{d x}-\frac{n(x, t) p(x, t)-n_{i e}^{2}}{\tau_{p}\left(n+n_{1}\right)+\tau_{n}\left(p+p_{1}\right)}  \tag{2.18a}\\
\frac{\partial p(x, t)}{\partial t} & =-\frac{1}{q_{e}} \frac{d J_{p}(x, t)}{d x}-\frac{n(x, t) p(x, t)-n_{i e}^{2}}{\tau_{p}\left(n+n_{1}\right)+\tau_{n}\left(p+p_{1}\right)}  \tag{2.18b}\\
J_{n}(x, t) & =q_{e} \mu_{n} n(x, t) \mathcal{E}(x, t)+q_{e} D_{n} \frac{\partial n(x, t)}{\partial x}  \tag{2.18c}\\
J_{p}(x, t) & =q_{e} \mu_{p} p(x, t) \mathcal{E}(x, t)-q_{e} D_{p} \frac{\partial p(x, t)}{\partial x}  \tag{2.18d}\\
\frac{d^{2} \psi(x, t)}{d x^{2}} & =-\frac{q_{e}}{\varepsilon_{r} \varepsilon_{0}}\left(p(x)-n(x)+N_{D}^{+}(x)-N_{A}^{-}(x)\right) \tag{2.18e}
\end{align*}
$$

$$
\begin{equation*}
\mathcal{E}(x, t)=-\frac{d \psi}{d x} \tag{2.18f}
\end{equation*}
$$

All of the above equations modeling the transport of charges and the supplementing equations are combined and called the Coupled Carrier Transport (CCT) equations for convenience. There are six equations with six unknown distributions ( $n, p, J_{n}$, $J_{p}, \psi$ and $\left.E\right)$. It is assumed that the immobile donor acceptor densities, $N_{D}^{+}$and $N_{A}^{-}$, and the diffusion, mobility, and recombination parameters are known. Of these, six unknown distributions, $p$ and $n$ distribution can alone be used to determine the other four. Usually, a numerical solution of CCT seeks $p, n$ and $\psi$ distributions often known as primary variables. These equations form the basis for programs such as Medici [2], which can be used to establish the steady-state and dynamic characteristics of semiconductor devices.

These model equations are then supplemented with appropriate boundary conditions as discussed in the following paragraphs. The ohmic contacts form Dirichlet boundaries at the semiconductor-metal contact interfaces. The electric potential and charge densities are assumed to be fixed at these interfaces. The quasi-Fermi potentials at the anode and cathode are fixed by the respective terminal potentials. Quasi-Fermi potential is a nonphysical quantity used to quantify the carrier concentration under non-equilibrium conditions. When the quasi-Fermi potentials are fixed, the carrier concentration at the boundaries are also fixed. The boundary electric potential $\psi$ is fixed such that there exists a zero-space charge for these carrier concentrations, $p+N_{D}^{+}=n+N_{A}^{-}$. Inserting $p n=n_{i e}^{2}$ into this relation gives the following quadratic.

$$
\begin{equation*}
p^{2}+p\left(N_{D}^{+}-N_{A}^{-}\right)-n_{i e}^{2}=0 \tag{2.19}
\end{equation*}
$$

The solution to this quadratic equation along with the inequalities $N_{A}^{-} \gg n_{i e}, N_{A}^{-} \gg$ $N_{D}^{+}$in $p$ region and $N_{D}^{+} \gg n_{i e}, N_{D}^{+} \gg N_{A}^{-}$in $n$ region leads to the following boundary charge densities.

$$
\begin{equation*}
p\left(x_{a}\right)=\frac{N_{A}^{-}}{2}+\sqrt{\left(\frac{N_{A}^{-}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{A}^{-} \tag{2.20a}
\end{equation*}
$$

$$
\begin{align*}
& n\left(x_{a}\right)=\frac{n_{i e}^{2}}{N_{A}^{-}}  \tag{2.20b}\\
& n\left(x_{c}\right)=\frac{N_{D}^{+}}{2}+\sqrt{\left(\frac{N_{D}^{+}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{D}^{+}  \tag{2.20c}\\
& p\left(x_{c}\right)=\frac{n_{i e}^{2}}{N_{D}^{+}} \tag{2.20d}
\end{align*}
$$

These boundary charge densities and quasi-Fermi potentials are used to compute the respective boundary electrical potentials given by (2.21).

$$
\begin{align*}
& \psi\left(x_{a}\right)=v_{a}-\frac{k T}{q_{e}} \ln \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right)  \tag{2.21a}\\
& \psi\left(x_{c}\right)=v_{c}+\frac{k T}{q_{e}} \ln \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right) \tag{2.21b}
\end{align*}
$$

The external-circuit voltage applied at the diode terminals anode and cathode are $v_{a}$ and $v_{c}$ respectively. Equations (2.18), (2.20), and (2.21) can be solved numerically to find the aforementioned primary variables.

The charge densities vary from $10^{15}$ to $10^{18}$ within the semiconductor according to the doping concentration. The variation in electric potential may range from a few volts to several hundred. This large difference between the variables can cause numerical difficulties. It is a common practice to scale the charge densities and the other relations accordingly in these situations. For numerical implementation, the scaling factor $N_{M}$, defined as the maximum of $N_{D}^{+}$and $N_{A}^{-}$concentration, is used. For simplicity, all derivations hereafter use the scaled quantities and is written without the use of tilde where there is no ambiguity.

$$
\begin{array}{cl}
\tilde{p}=\frac{p}{N_{M}} ; & \tilde{n}=\frac{n}{N_{M}} \\
\tilde{p}_{1}=\frac{p_{1}}{N_{M}} ; & \tilde{n}_{1}=\frac{n_{1}}{N_{M}} \\
\tilde{J}_{p}=\frac{J_{p}}{N_{M}} ; & \tilde{J}_{n}=\frac{J_{n}}{N_{M}}  \tag{2.22}\\
\tilde{n}_{i e}=\frac{n_{i e}}{N_{M}} ; & \tilde{N}_{A}^{-}=\frac{N_{A}^{-}}{N_{M}} \\
\tilde{N}_{D}^{+}=\frac{N_{D}^{+}}{N_{M}} ; & N_{M}=\max \left(N_{D}^{+}, N_{A}^{-}\right)
\end{array}
$$

### 2.7 Chapter Summary

The physical processes that are responsible for the charge transport within a semiconductor are modeled using appropriate equations from referenced sources [5557]. Several assumptions are made in deriving these equations. The temperature of the device is assumed to be lumped and is constant (e.g. $T=300 \mathrm{~K}$ ). Drift and diffusion parameters are assumed to be constant which can possibly vary according to the doping level and electric field. Carrier lifetimes that can vary with doping level are also assumed to be constant. Recombination processes that are non-dominant and impact ionization models are neglected. When the results of the semiconductorcircuit simulation is discussed in Chapter 6, it will be evident that these are reasonable assumptions. The important model equations, namely CCT, is derived which is used to establish a numerical solution.

## 3. DISCRETIZATION AND MODEL ENCAPSULATION

The CCT equations (2.18) must be discretized in both time and space to establish a solution numerically. A numerical solution of the fully discretized system gives the primary unknown distributions $(p, n, \psi)$ in the device. Post-processing this information gives the current density, turn-on, turn-off, and conduction losses of the device under consideration. In the first section, the discretization approach is described abstractly and the basic patterns are set forth. Spatial discretization is performed first to result in a system of ordinary differential equations, which can be solved using any one of a number of established temporal integration algorithms. In the second section, a finite element discretization of the Poisson's equation is explained followed by a discussion on finite volume discretization of the continuity equation based on central and exponentially-weighted difference (Scharfetter-Gummel) approaches. With some foresight, device terminal equations are augmented to the discretized CCT equations to result in model encapsulation as illustrated in the last section.

### 3.1 Abstract Discretization

It is useful to consider the continuity and Poisson equations from CCT equations. Poisson's equation is rewritten and the right-hand sides of the continuity equations are condensed into general functions to form the following set of equations,

$$
\begin{align*}
\frac{\partial p}{\partial t} & =f_{p}(p, n, \psi)  \tag{3.1a}\\
\frac{\partial n}{\partial t} & =f_{n}(p, n, \psi)  \tag{3.1b}\\
0 & =f_{\psi}\left(p, n, \psi, N_{A}^{-}, N_{D}^{+}\right) \tag{3.1c}
\end{align*}
$$

where it is assumed that $p=p(x, t), n=n(x, t)$, etc.

The solution is to be established for a time partition $t^{0}<t^{1}<\ldots<t^{m}<\ldots<t^{M}$ with intervals $\Delta t^{m}=t^{m}-t^{m-1}$. The next time step's ( $m^{t h}$ ) solution is computed using $(m-1)^{t h}$-step solution and the time-step size $\Delta t^{m}$. The backward Euler algorithm is used first to discretize (3.1) in time to get

$$
\begin{align*}
p^{m}-p^{m-1} & =\Delta t^{m} f_{p}\left(p^{m}, n^{m}, \psi^{m}\right)  \tag{3.2a}\\
n^{m}-n^{m-1} & =\Delta t^{m} f_{n}\left(p^{m}, n^{m}, \psi^{m}\right)  \tag{3.2b}\\
0 & =f_{\psi}\left(p^{m}, n^{m}, \psi^{m}, N_{A}^{-}, N_{D}^{+}\right) \tag{3.2c}
\end{align*}
$$

The set of equations in (3.2) form a nonlinear system requiring iterative solution. Newton-Raphson (NR) iteration is preferred over fixed-point methods due to its rapid (quadratic) convergence in the neighborhood of the exact solution. It should be kept in mind that this set of equations is discretized only in time and yet to be discretized in space. The NR iteration indices are denoted using $k, k+1, \ldots$ and $\chi$ as place holder for $p, n$, or $\psi$. The previous time step's convergent solution is denoted as $\chi^{m-1, \infty}$, and $\chi^{m, k}$ is the $k^{t h}$ iterate at the $m^{\text {th }}$ time step. It is convenient to define $\Delta \chi=\chi^{m, k+1}-\chi^{m, k}$. The NR iterator is derived using the truncated Taylor's series expansion of the nonlinear equation about the convergent solution. Expanding the functions in (3.2) using Taylor series up to first order and rearranging results in the equations below (arguments of $f_{\chi}$ are omitted for brevity).

$$
\begin{align*}
& 0=p^{m}-p^{m-1}-\Delta t^{m} f_{p}+\left(I-\frac{\partial f_{p}}{\partial p}\right) \Delta p-\frac{\partial f_{p}}{\partial n} \Delta n-\frac{\partial f_{p}}{\partial \psi} \Delta \psi  \tag{3.3a}\\
& 0=n^{m}-n^{m-1}-\Delta t^{m} f_{n}+\left(I-\frac{\partial f_{n}}{\partial n}\right) \Delta n-\frac{\partial f_{n}}{\partial p} \Delta p-\frac{\partial f_{n}}{\partial \psi} \Delta \psi  \tag{3.3b}\\
& 0=f_{\psi}+\frac{\partial f_{\psi}}{\partial p} \Delta p+\frac{\partial f_{\psi}}{\partial n} \Delta n+\frac{\partial f_{\psi}}{\partial \psi} \Delta \psi \tag{3.3c}
\end{align*}
$$

The gradient functions $f_{\chi}$ are evaluated using $\chi^{m, k}$ values of the associated arguments. The NR iterator is the matrix-vector representation of the previous set of equations. The Jacobian elements are given by $\partial f_{\chi} / \partial \chi$. When the Jacobian is computed in every iteration using $\chi^{m, k}$ values, the iterative method is called full Newton method. In some cases, it suffices to evaluate the Jacobian once per time step using $\chi^{m-1, \infty}$ values.

This is known as simplified-Newton or quasi-Newton method, a rigorous analysis of its convergence properties can be found in [58]. The computational performance of these variations of NR are examined in Chapter 6 to identify a suitable method for this thesis.

The linear system to be solved using NR iteration is shown in (3.4). Next, it is useful to consider a spatial partition $0=x_{1}<x_{2} \ldots<x_{i}<\ldots<x_{N}=X$ with interval $h_{i}=x_{i+1}-x_{i}$. Then each $f_{\chi}$ and $\partial f_{\chi} / \partial \chi$ can be discretized in space to get vectors and matrices with dimensions $N \times 1$ and $N \times N$, respectively. The pattern seen in (3.4) is the block-matrix pattern of a fully discretized system. In the subsequent sections, a spatial discretization is performed first followed by a temporal discretization. The order of the spatial discretization and temporal discretization does not matter when spatial and time partitions such as those described previously are used. However, this relation no longer holds [8] when a spatially adaptive mesh is used.

$$
\begin{align*}
& {\left[\begin{array}{ccc}
I / \Delta t^{m}-\partial f_{p} / \partial p & -\partial f_{p} / \partial n & -\partial f_{p} / \partial \psi \\
-\partial f_{n} / \partial p & I / \Delta t^{m}-\partial f_{n} / \partial n & -\partial f_{n} / \partial \psi \\
\partial f_{\psi} / \partial p & \partial f_{\psi} / \partial n & \partial f_{\psi} / \partial \psi
\end{array}\right]\left[\begin{array}{c}
\Delta p \\
\Delta n \\
\Delta \psi
\end{array}\right]}  \tag{3.4}\\
& =\left[\begin{array}{c}
-\left(p^{m, k}-p^{m-1, \infty}\right) / \Delta t^{m}+f_{p}\left(p^{m, k}, n^{m, k}, \psi^{m, k}\right) \\
-\left(n^{m, k}-n^{m-1, \infty}\right) / \Delta t^{m}+f_{n}\left(p^{m, k}, n^{m, k}, \psi^{m, k}\right) \\
-f_{\psi}\left(p^{m, k}, n^{m, k}, \psi^{m, k}, N_{A}^{-}, N_{D}^{+}\right)
\end{array}\right]
\end{align*}
$$

### 3.2 Poisson's Equation

Poisson's equation, which relates the carrier densities and electric field, produces an algebraic relationship between $p, n$, and $\psi$ (third equation in (3.1)) and is an elliptic PDE. Numerical methods such as the finite element method can be used to numerically solve an elliptic PDE. The problem can be formulated using variational or Galerkin approaches, both resulting in a similar discretized linear system of equations. The mathematical theory behind the existence and uniqueness of the solution to this
type of problem is well established [59]. In the one-dimensional case, the Poisson's equation may be expressed as,

$$
\begin{equation*}
\frac{d^{2} \psi}{d x^{2}}=-\frac{\rho(x)}{\varepsilon} \tag{3.5}
\end{equation*}
$$

In one dimension, $\psi(x)$ is approximated as piecewise linear or, equivalently, as a weighed sum of basis functions,

$$
\begin{equation*}
\psi(x)=\sum_{i=1}^{N} \psi_{i} \varphi_{i}(x) \tag{3.6}
\end{equation*}
$$

where $\psi_{i}=\psi\left(x_{i}\right)$ are the nodal values of the electrical potentials. The basis function $\varphi_{i}(x)$ with a support of $\left[x_{i-1}, x_{i+1}\right]$ is defined as,

$$
\varphi_{i}(x)=\left\{\begin{array}{lc}
\frac{x-x_{i-1}}{x_{i}-x_{i-1}} & : x \in\left[x_{i-1}, x_{i}\right]  \tag{3.7}\\
\frac{x_{i+1}-x}{x_{i+1}-x_{i}} & : x \in\left[x_{i}, x_{i+1}\right] \\
0 & \text { otherwise. }
\end{array}\right.
$$

Dirichlet boundary values of $\psi$ are specified at $x=0$ and $x=X$. Substituting (3.6) into (3.5) and using integration by parts.

$$
\begin{equation*}
\int_{-x_{a}}^{x_{c}} \varepsilon \psi^{\prime}(x) \sum_{i=1}^{N} \varphi_{i}^{\prime} d x=\int_{-x_{a}}^{x_{c}} \rho(x) \sum_{i=1}^{N} \varphi_{i} d x \tag{3.8}
\end{equation*}
$$

The support of the basis functions is such that for the $i^{\text {th }}$ equation, $\{i-1, i, i+1\}$ nodes are involved. A detailed derivation of the $i^{t h}$ equation is provided in Appendix A.1. The final result is given by,

$$
\begin{equation*}
-\frac{\varepsilon}{h_{i-1}} \psi_{i-1}+\left(\frac{\varepsilon}{h_{i-1}}+\frac{\varepsilon}{h_{i}}\right) \psi_{i}-\frac{\varepsilon}{h_{i}} \psi_{i+1}=\frac{h_{i-1}}{6} \rho_{i-1}+\left(\frac{h_{i-1}}{3}+\frac{h_{i}}{3}\right) \rho_{i}+\frac{h_{i}}{6} \rho_{i}, \tag{3.9}
\end{equation*}
$$

where,

$$
\begin{equation*}
\rho_{i}=q_{e} N_{M}\left(p_{i}-n_{i}+N_{D, i}^{+}-N_{A, i}^{-}\right), \tag{3.10}
\end{equation*}
$$

and,

$$
\begin{equation*}
h_{i}=x_{i+1}-x_{i} . \tag{3.11}
\end{equation*}
$$

The left-hand side of (3.9), when expressed for $i=1, \ldots N$, can be written as a matrixvector product. The matrix is known as stiffness matrix $(\check{\mathbf{S}})$. The $i^{t h}$ equation of the
stiffness matrix has nonzero entries in the columns $\{i-1, i, i+1\}$ resulting in a tridiagonal matrix. $\boldsymbol{\psi}$ is a vector of the nodal values of electrical potential $\psi$. The vector on the right-hand side $\check{\boldsymbol{\rho}}$ is the source vector.

$$
\begin{equation*}
\check{\mathrm{S}} \psi-\check{\rho}=0 \tag{3.12}
\end{equation*}
$$

The source vector $\check{\rho}$ is expressed using a matrix $\check{\mathbf{D}}$ times a charge vector product. The resulting equation is given in (3.12).

$$
\begin{equation*}
\check{\mathbf{S}} \boldsymbol{\psi}-\check{\mathbf{D}}\left(\mathbf{p}-\mathbf{n}+\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)=\mathbf{0} \tag{3.13}
\end{equation*}
$$

The matrices $\check{\mathbf{S}}$ and $\check{\mathbf{D}}$ require modification to take boundary conditions into account. A concise representation of (3.13) is (A.14). The boundary values of the electric potentials are known at $x=0$ and $x=X$. Thus, a linear solve of an $N-2 \times N-2$ dimensional system is required to establish the $N-2$ unknown potentials. Instead of excluding these potentials from (3.13), they are still included, however, with modification to the stiffness matrix by zeroing the first and last rows and setting the $(1,1)$ and $(N, N)$ entries to unity. A similar modification of the first and last rows of the $\check{\mathbf{D}}$ matrix is performed. This is done such that boundary $\psi$ is equal to the values specified by a boundary vector. The boundary vector is defined using the knowledge of the diode terminal anode voltage $\left(v_{a}\right)$, cathode voltage $\left(v_{c}\right)$ and charge densities such that the first and last entries correspond to the nodal values $\psi_{1}$ and $\psi_{N}$. The boundary potentials given by (2.21) are repeated here for the sake of completeness in setting up a boundary value problem.

$$
\begin{align*}
\psi_{1} & =v_{a}-\frac{k T}{q_{e}} \ln \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right) \\
\psi_{N} & =v_{c}+\frac{k T}{q_{e}} \ln \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right) \tag{3.14}
\end{align*}
$$

The complete system of equations that are to be solved to establish the electric potential given the charge densities are summarized as follows,

$$
\begin{equation*}
0=\mathbf{S} \boldsymbol{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)-\mathbf{b}\left(v_{a}, v_{c}\right) \tag{3.15a}
\end{equation*}
$$

$$
\begin{align*}
& \mathbf{S}=\left[\begin{array}{cccccc}
1 & 0 & 0 & \cdots & 0 & 0 \\
-\frac{\varepsilon}{h_{1}} & \frac{\varepsilon}{h_{1}}+\frac{\varepsilon}{h_{2}} & -\frac{\varepsilon}{h_{2}} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{\varepsilon}{h_{N-2}}+\frac{\varepsilon}{h_{N-1}} & -\frac{\varepsilon}{h_{N-1}} \\
0 & 0 & 0 & \cdots & 0 & 1
\end{array}\right]  \tag{3.15b}\\
& \mathbf{D}=q_{e} N_{M}\left[\begin{array}{cccccc}
0 & 0 & 0 & \cdots & 0 & 0 \\
\frac{h_{1}}{6} & \frac{h_{1}+h_{2}}{3} & \frac{h_{2}}{6} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{h_{N-2}+h_{N-1}}{3} & \frac{h_{N-1}}{6} \\
0 & 0 & 0 & \cdots & 0 & 0
\end{array}\right]  \tag{3.15c}\\
& \mathbf{b}\left(v_{a}, v_{c}\right)=\left[\begin{array}{c}
0 \\
v_{a}-\frac{k T}{q_{e}} \ln \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right) \\
0 \\
\vdots \\
0 \\
v_{c}+\frac{k T}{q_{e}} \ln \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right)
\end{array}\right] \tag{3.15d}
\end{align*}
$$

### 3.3 Continuity Equations

The first two equations in (3.1) represent continuity equations. These equations are discretized using the finite volume approach. A one-dimensional spatial partition, similar to that in Section 3.1, and a prototypical finite volume is depicted in Figure 3.1. The nodal values of the charge densities, namely $p_{i}=p\left(x_{i}\right)$ and $n_{i}=n\left(x_{i}\right)$, are the average value over the volume $\left(h_{i-1}+h_{i}\right) / 2$. The electric field $\mathcal{E}_{i+1 / 2}$ or flux in a given interval $\left[x_{i}, x_{i+1}\right]$ of the partition is found using a finite difference formula (negative gradient of the electric potential). The gradient of electrical potential, since $\psi$ is approximated by piecewise linear functions, is piecewise constant. For the computation of current density $J$ entering the volume face at $x_{i+1 / 2}=\left(x_{i}+x_{i+1}\right) / 2$, values of the charge densities at the face are required. The current or flux is assumed
to enter the volume face at $x_{i-1 / 2}$ and leave at $x_{i+1 / 2}$. The current density values can be approximated using central difference or Scharfetter-Gummel methods. These methods are discussed in the following subsections.


Fig. 3.1.: Finite volume discretization (top) and a prototypical volume (bottom).

### 3.3.1 Central difference method

In the finite volume method, the Divergence theorem is applied to the differential equations to be discretized. Then, the derivative terms are approximated by appropriate finite difference formulae using the nodal values of the spatial partition under consideration. When applied to the continuity equation, the rate of change of charge densities averaged over the finite volume is the difference in the flux (current densities in semiconductors) entering and leaving the $i^{\text {th }}$ control volume.

$$
\begin{align*}
\frac{\partial p_{i}}{\partial t} & =-\frac{1}{q_{e}} \frac{J_{p, i+1 / 2}-J_{p, i-1 / 2}}{\left(h_{i-1}+h_{i}\right) / 2}-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)}  \tag{3.16a}\\
\frac{\partial n_{i}}{\partial t} & =\frac{1}{q_{e}} \frac{J_{n, i+1 / 2}-J_{n, i-1 / 2}}{\left(h_{i-1}+h_{i}\right) / 2}-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} \tag{3.16b}
\end{align*}
$$

The current densities due to holes and electrons are computed numerically by discretizing the respective equation from the CCT equation , producing

$$
\begin{align*}
J_{p, i+1 / 2} & =q_{e}\left(\mu_{p} \frac{\left(p_{i}+p_{i+1}\right)}{2} \mathcal{E}_{i+1 / 2}-D_{p} \frac{p_{i+1}-p_{i}}{h_{i}}\right),  \tag{3.17a}\\
J_{n, i+1 / 2} & =q_{e}\left(\mu_{n} \frac{\left(n_{i}+n_{i+1}\right)}{2} \mathcal{E}_{i+1 / 2}+D_{n} \frac{n_{i+1}-n_{i}}{h_{i}}\right),  \tag{3.17b}\\
\mathcal{E}_{i+1 / 2} & =-\frac{\psi_{i+1}-\psi_{i}}{h_{i}}=-\frac{\Delta \psi_{i+1 / 2}}{h_{i}} . \tag{3.17c}
\end{align*}
$$

The main difference between the central difference and Scharfetter-Gummel methods is in the approximation of the current densities. A linear interpolation of the charge densities at the volume faces is used here. The volume face is located midway between two nodes. In the central difference approach, the charge densities at the volume face is the average of the values at the neighboring nodes. Substituting (3.17) into (3.16) results in,

$$
\begin{align*}
\frac{\partial p_{i}}{\partial t}= & -\frac{1}{\left(h_{i-1}+h_{i}\right) / 2}\left[\mu_{p}\left(\frac{p_{i}+p_{i+1}}{2} \frac{\psi_{i}-\psi_{i+1}}{h_{i}}-\frac{p_{i-1}+p_{i}}{2} \frac{\psi_{i-1}-\psi_{i}}{h_{i-1}}\right)\right. \\
& \left.-D_{p}\left(\frac{p_{i+1}-p_{i}}{h_{i}}-\frac{p_{i}-p_{i-1}}{h_{i-1}}\right)\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)},  \tag{3.18a}\\
\frac{\partial n_{i}}{\partial t}= & \frac{1}{\left(h_{i-1}+h_{i}\right) / 2}\left[\mu_{n}\left(\frac{n_{i}+n_{i+1}}{2} \frac{\psi_{i}-\psi_{i+1}}{h_{i}}-\frac{n_{i-1}+n_{i}}{2} \frac{\psi_{i-1}-\psi_{i}}{h_{i-1}}\right)\right. \\
& \left.+D_{n}\left(\frac{n_{i+1}-n_{i}}{h_{i}}-\frac{n_{i}-n_{i-1}}{h_{i-1}}\right)\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} . \tag{3.18b}
\end{align*}
$$

The finite difference formulae should possess certain properties including conservation, boundedness, and transportiveness, which collectively ensure reliable (convergent) results [60]. The numerical approach preserves conservation when the flux leaving a volume face is the same as that entering the adjacent volume face. When the discretization method results in a diagonally dominant system matrix, the boundedness property is satisfied. This ensures that the solution increases or decreases monotonically. In this context, flow of species is nothing but the flow of current densities. When the flow is drift-dominated the species density at a node varies more due to the upwind node (next node against the flow) than due to a downwind (next node along the flow). A transportive discretization method takes this flow's direction
and influence of neighboring node species densities at a given node accordingly into account. The central difference method is conservative but during drift-dominated operation, it loses its transportive and boundedness properties.

Table 3.1.: Parameters of an example $p n$-diode with step junction.

| Variable | Symbol | Value |
| :---: | :---: | :---: |
| Diode length | $X$ | 0.005 m |
| $p n$-junction area | $a_{c}$ | $0.01 \mathrm{~cm}^{2}$ |
| Acceptor concentration | $N_{A}$ | $10^{19} \mathrm{~cm}^{-3}$ |
| Donor concentration | $N_{D}$ | $5 \times 10^{16} \mathrm{~cm}^{-3}$ |

The ratio of drift and diffusion components of current are quantified using the so-called Péclet number $\left(P_{e}\right)$. In particular, the Péclet number $P_{e}$ is the ratio of drift to diffusion component coefficients within each volume. For the central difference method to satisfy boundedness, the Péclet number must satisfy $P_{e} \leq 2$ [60]. The $P_{e}$ for the semiconductor CCT is derived using (3.17). This number is the same for both the electron and hole currents.

$$
\begin{align*}
\left|P_{e}\right| & =\left|\frac{q_{e} \mu_{n}\left(\psi_{i}-\psi_{i+1}\right)}{h_{i} q_{e} \frac{D_{n}}{h_{i}}}\right| \\
& =\left|\frac{\mu_{n}\left(\psi_{i}-\psi_{i+1}\right)}{D_{n}}\right|  \tag{3.19}\\
& =\left|\frac{\psi_{i}-\psi_{i+1}}{V_{T}}\right|
\end{align*}
$$

The Péclet number is large in the semiconductor diode when it operates under reverse bias. To illustrate this, the diode model is discretized using central difference method on a logarithmically spaced 100-node grid for an example $p n$ diode with step junction. The doping density is constant through out the region and drops to zero at the junction to form a step junction. The species densities undergo a rapid change with respect to position near the vicinity of depletion region formed around this step
junction. The diode is simulated with a bias voltage of -50 V is shown in Figure 3.2. The parameters of the diode is given in Table 3.1.


Fig. 3.2.: Charge densities and Péclet number near $p n$-junction using central difference method.

As shown, the scaled charge densities oscillate near the edges of the depletion region where $\left|P_{e}\right|$ is greater than two and the boundedness property is not satisfied. The central difference method does not take the flow direction into account while computing the flux at volume faces. Consequently, the central difference method does not satisfy the transportiveness property, especially at high $P_{e}$. These two difficulties are circumvented in the following discretization approach.

### 3.3.2 Scharfetter-Gummel method

The drift part of the current density in (3.17) uses equal weight for the $p$ or $n$ at the neighboring nodes. In the Scharfetter-Gummel method [13, 18], an exponential
weight, which is a function of electrical potential $\psi$, is used. The discretization of the continuity equation starts with the current density equations,

$$
\begin{align*}
J_{p, i+1 / 2} & =q_{e} \mu_{p} p \mathcal{E}_{i+1 / 2}-q_{e} D_{p} \frac{\partial p}{\partial x}  \tag{3.20a}\\
J_{n, i+1 / 2} & =q_{e} \mu_{n} n \mathcal{E}_{i+1 / 2}+q_{e} D_{n} \frac{\partial n}{\partial x} \tag{3.20b}
\end{align*}
$$

The electric field $\mathcal{E}_{i+1 / 2}$ is the the discrete approximation of $-\nabla \psi$ in one-dimensional interval $\left[x_{i}, x_{i+1}\right]$. A piecewise linear approximation of $\psi$ implies that the electric field is piecewise constant. When the current density in interval $\left[x_{i}, x_{i+1}\right]$ is constant the current density equations can be treated as a simple boundary value problem in each interval. The Einstein's relation is,

$$
\begin{equation*}
\frac{D_{p}}{\mu_{p}}=\frac{D_{n}}{\mu_{n}}=\frac{k T}{q_{e}}=V_{T} \tag{3.21}
\end{equation*}
$$

where $V_{T}$ is thermal voltage. The hole current density equation (3.20a) along with Einstein's relation is written as,

$$
\begin{align*}
J_{p, i+1 / 2} & =q_{e} \mu_{p} \mathcal{E}_{i+1 / 2} p-q_{e} D_{p} \frac{d p}{d x}  \tag{3.22a}\\
\frac{J_{p, i+1 / 2}}{-q_{e} D_{p}} & =\frac{d p}{d x}-\frac{\mathcal{E}_{i+1 / 2}}{V_{T}} p,  \tag{3.22b}\\
j & =\frac{d p}{d x}-\nu p, \tag{3.22c}
\end{align*}
$$

where,

$$
\begin{equation*}
\nu=\frac{\mathcal{E}_{i+1 / 2}}{V_{T}}, \quad j=\frac{J_{p, i+1 / 2}}{-q_{e} D_{p}} . \tag{3.22~d}
\end{equation*}
$$

The definitions (3.22d) are made to make the derivation more clear and readable. The only variable that changes with position $x$ in (3.22c) is hole density $p$. If the current density in the interval considered is not assumed to be constant then this equation cannot be solved. The first-order derivative corresponds to the diffusion current component while the charge density times $\nu$ (equivalent drift velocity) is the drift component. Multiplying both sides of (3.22c) by an exponential factor $\exp (-\nu x)$
and integrating both sides with respect to $x$ yields the sequence of equations in (3.23). The boundary values of the hole densities at $x_{i}$ and $x_{i+1}$ are $p_{i}$ and $p_{i+1}$ respectively.

$$
\begin{align*}
j e^{-\nu x} & =\frac{d}{d x} p e^{-\nu x}-\nu p e^{-\nu x}  \tag{3.23a}\\
j e^{-\nu x} & =\frac{d}{d x} p e^{-\nu x}  \tag{3.23b}\\
\int_{x_{i}}^{x_{i+1}} j e^{-\nu x} d x & =\int_{x_{i}}^{x_{i+1}} \frac{d}{d x} p e^{-\nu x} d x  \tag{3.23c}\\
\left.j \frac{e^{-\nu x}}{-\nu}\right|_{x_{i}} ^{x_{i+1}} & =\left.p e^{-\nu x}\right|_{x_{i}} ^{x_{i+1}}  \tag{3.23d}\\
-\frac{j}{\nu}\left(e^{-\nu x_{i+1}}-e^{-\nu x_{i}}\right) & =p_{i+1} e^{-\nu x_{i+1}}-p_{i} e^{-\nu x_{i}}  \tag{3.23e}\\
-\frac{j}{\nu}\left(e^{-\nu\left(x_{i+1}-x_{i}\right)}-1\right) & =p_{i+1} e^{-\nu\left(x_{i+1}-x_{i}\right)}-p_{i}  \tag{3.23f}\\
j & =-\nu \frac{p_{i+1} e^{-\nu h_{i}}-p_{i}}{e^{-\nu h_{i}}-1}  \tag{3.23g}\\
J_{p, i+1 / 2} & =q_{e} D_{p} \frac{\mathcal{E}_{i+1 / 2}}{V_{T}} \frac{p_{i+1} \exp \left(-\mathcal{E}_{i+1 / 2} h_{i} / V_{T}\right)-p_{i}}{\exp \left(-\mathcal{E}_{i+1 / 2} h_{i} / V_{T}\right)-1} \tag{3.23h}
\end{align*}
$$

The definitions (3.22d) are used to recover the actual hole current density equation (3.23h). It should be noted that $\mathcal{E}_{i+1 / 2} h_{i}=-\Delta \psi_{i+1 / 2}$. Equation (3.23h) is further rewritten to get the current density at the control volume face.

$$
\begin{equation*}
J_{p, i+1 / 2}=q_{e} \mu_{p} \mathcal{E}_{i+1 / 2} \frac{\exp \left(\Delta \psi_{i+1 / 2} / V_{T}\right) p_{i+1}-p_{i}}{\exp \left(\Delta \psi_{i+1 / 2} / V_{T}\right)-1} \tag{3.24}
\end{equation*}
$$

where $\mu_{p}=D_{p} / V_{T}$. A similar expression for electron current density at the control faces can be derived.

$$
\begin{equation*}
J_{n, i+1 / 2}=q_{e} \mu_{n} \mathcal{E}_{i+1 / 2} \frac{\exp \left(-\Delta \psi_{i+1 / 2} / V_{T}\right) n_{i+1}-n_{i}}{\exp \left(-\Delta \psi_{i+1 / 2} / V_{T}\right)-1} \tag{3.25}
\end{equation*}
$$

where $\mu_{n}=D_{n} / V_{T}$.
A definition of Bernoulli function $\mathcal{B}(x)=\frac{x}{\exp (x)-1}$ and a change of variable $z_{i+1 / 2}=\Delta \psi_{i+1 / 2} / V_{T}$ makes the hole and electron current density equations more compact.

$$
\begin{align*}
& J_{p, i+1 / 2}=\frac{q_{e} D_{p}}{h_{i}}\left[-p_{i+1} \mathcal{B}\left(-z_{i+1 / 2}\right)+p_{i} \mathcal{B}\left(z_{i+1 / 2}\right)\right]  \tag{3.26a}\\
& J_{n, i+1 / 2}=\frac{q_{e} D_{n}}{h_{i}}\left[n_{i+1} \mathcal{B}\left(z_{i+1 / 2}\right)-n_{i} \mathcal{B}\left(-z_{i+1 / 2}\right)\right] \tag{3.26b}
\end{align*}
$$

The continuity equation is discretized by inserting (3.26) in (3.16) with some manipulation to get,

$$
\begin{align*}
\frac{\partial p_{i}}{\partial t}= & \frac{2 D_{p}}{h_{i}+h_{i-1}}\left[\frac{\mathcal{B}\left(-z_{i+1 / 2}\right)}{h_{i}} p_{i+1}+\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}} p_{i-1}\right. \\
& \left.-\left(\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}}+\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}}\right) p_{i}\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)}  \tag{3.27a}\\
\frac{\partial n_{i}}{\partial t}= & \frac{2 D_{n}}{h_{i}+h_{i-1}}\left[\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}} n_{i+1}+\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}} n_{i-1}\right. \\
& \left.-\left(\frac{\mathcal{B}\left(-z_{i+1 / 2}\right)}{h_{i}}+\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}}\right) n_{i}\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} \tag{3.27b}
\end{align*}
$$

From (3.26), it is obvious that current densities are linear functions of the respective charge densities. The weighting coefficient of these charge densities is the Bernoulli function and hence this is an exponentially weighted approach. This exponentially weighting adds numerical stability to the method.

Table 3.2.: Bernoulli function values.

| Current | $z_{i+1 / 2}$ | $\mathcal{B}\left(z_{i+1 / 2}\right)$ | $J_{p, i+1 / 2}$ |
| :---: | :---: | :---: | :---: |
| Diffusion dominant | approx. <br> zeros, $\mathcal{E}_{i+1 / 2} \approx 0$ | $\rightarrow 1$ | $-q_{e} D_{p} \frac{p_{i+1}-p_{i}}{h_{i}}$ |
| Drift dominant | large <br> positive, $\mathcal{E}_{i+1 / 2} \ll 0$ | $\rightarrow 0$ | $-q_{e} \frac{D_{p}}{h_{i}} \frac{\Delta \psi_{i+1 / 2}}{V_{T}} p_{i+1}=q_{e} \mu_{p} \mathcal{E}_{i+1 / 2} p_{i+1}$ |
| Drift <br> dominant | large <br> negative $y$, $\mathcal{E}_{i+1 / 2} \gg 0$ | $\rightarrow-y$ | $-q_{e} \frac{D_{p}}{h_{i}} \frac{\Delta \psi_{i+1 / 2}}{V_{T}} p_{i}=q_{e} \mu_{p} \mathcal{E}_{i+1 / 2} p_{i}$ |

The robustness of the Scharfetter-Gummel method is explained with the help of the hole current density equation (3.26a) and Table 3.2. It is apparent from (3.17) that the diffusion component of the current dominates in the regions of zero electric field.


Fig. 3.3.: Charge densities and Péclet number near $p n$-junction using ScharfetterGummel method.

This corresponds to the first row in Table 3.2. In the other extreme where the electric fields are high, the current densities are dominated by the drift components. The drift components are evaluated using either the $i^{t h}$ node or $(i+1)^{t h}$ node charge density rather than the linear interpolation of these two values. This first-order approximation of the drift component, as opposed to a second-order linear interpolated value, tends to damp the numerical oscillations in the charge densities and preserve the boundedness property. For the same reason, the flow direction and node potential's influence is taken into account making Scharfetter-Gummel method transportive.

The $i^{t h}$ node charge density alone affects the solution when the electric field is in the positive $x$ direction, while in the opposite case it is due to $(i+1)^{t h}$ node. This strategy of using either of the neighboring node's information in flux evaluation is known as the upwinding scheme in computational fluid dynamics. Since the Scharfetter-Gummel method takes the flow direction into account, the method also satisfies the transportiveness property. For illustration, the diode with parame-
ters given in Table 3.1 with 100 logarithmically-spaced nodes is discretized using the Scharfetter-Gummel method. The results of a simulation of the diode with a bias of -50 V is shown in Figure 3.3. It can be seen that the charge densities do not become negative implying that the boundedness property is preserved. The reason for the small overshoot in $n$ near the edge of the depletion region is due to the lack of nodes resolving the change in $n$.

### 3.3.3 Boundary charge densities

The continuity equations require the boundary data before a numerical solution is computed. The boundary charge densities given in (2.20) is repeated here for convenience.

$$
\begin{align*}
& n\left(x_{c}\right)=\frac{N_{D}^{+}}{2}+\sqrt{\left(\frac{N_{D}^{+}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{D}^{+}  \tag{3.28a}\\
& p\left(x_{c}\right)=\frac{n_{i e}^{2}}{N_{D}^{+}}  \tag{3.28b}\\
& p\left(x_{a}\right)=\frac{N_{A}^{-}}{2}+\sqrt{\left(\frac{N_{A}^{-}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{A}^{-}  \tag{3.28c}\\
& n\left(x_{a}\right)=\frac{n_{i e}^{2}}{N_{A}^{-}} \tag{3.28d}
\end{align*}
$$

### 3.4 Temporal Discretization

The discretized continuity equations for each control volume can be assembled into a concise form. Let $\mathbf{p}, \mathbf{n}$ and $\boldsymbol{\psi}$ represent vectors of nodal values of $p, n$ and $\psi$ respectively. The equations take the form,

$$
\begin{align*}
\frac{d \mathbf{p}}{d t} & =\mathbf{f}_{\mathbf{p}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) \\
\frac{d \mathbf{n}}{d t} & =\mathbf{f}_{\mathbf{n}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{3.29}\\
\mathbf{0} & =\mathbf{S} \boldsymbol{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)-\mathbf{b}\left(v_{a}, v_{c}\right)=\mathbf{f}_{\psi}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})
\end{align*}
$$

This is a semi-discretized version of the CCT equations, since it is discretized only with respect to space. In order to establish a fully discretized equation, it is necessary
to select a time partition as in Section 3.1 and apply a suitable integration algorithm to (3.29). The system of differential equations (3.29) is commonly referred to as Differential Algebraic Equations (DAEs). The first two systems in (3.29) are systems of differential equations along with the last system of algebraic equations relating the states and electric potential, and hence the name DAE. The algebraically-related electric potential makes this a temporally stiff DAE system.

Stiff DAEs require a stiffly-stable integration algorithm and the vast majority of such algorithms are implicit. A stiffly stable integration algorithm can take larger time steps after fast transients subside without producing numerical instability. The simplest stiffly-stable integration algorithm is backward Euler method. The fully discretized system, after using backward Euler method with a partition as in Section 3.1, is given by

$$
\begin{align*}
\mathbf{p}^{m}-\mathbf{p}^{m-1} & =\Delta t^{m} \mathbf{f}_{\mathbf{p}}\left(\mathbf{p}^{m}, \mathbf{n}^{m}, \boldsymbol{\psi}^{m}\right), \\
\mathbf{n}^{m}-\mathbf{n}^{m-1} & =\Delta t^{m} \mathbf{f}_{\mathbf{n}}\left(\mathbf{p}^{m}, \mathbf{n}^{m}, \boldsymbol{\psi}^{m}\right),  \tag{3.30}\\
\mathbf{0} & =\mathbf{f}_{\boldsymbol{\psi}}\left(\mathbf{p}^{m}, \mathbf{n}^{m}, \boldsymbol{\psi}^{m}, \boldsymbol{N}_{A}^{-}, \boldsymbol{N}_{D}^{+}\right) .
\end{align*}
$$

This is a nonlinear system of equations requiring iterative solution. In Section 3.1, the Newton-Raphson iterator is derived using truncated Taylor's series similar to (3.3). Following the same approach, the Newton-Raphson iterator for (3.30) is derived and given by,

$$
\begin{align*}
& {\left[\begin{array}{ccc}
\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} \\
-\mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} \\
\mathbf{f}_{\psi, \mathbf{p}} & \mathbf{f}_{\boldsymbol{\psi}, \mathbf{n}} & \mathbf{f}_{\boldsymbol{\psi}, \boldsymbol{\psi}}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{p} \\
\Delta \mathbf{n} \\
\Delta \boldsymbol{\psi}
\end{array}\right]}  \tag{3.31}\\
& =\left[\begin{array}{c}
-\left(\mathbf{p}^{m, k}-\mathbf{p}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{p}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right) \\
-\left(\mathbf{n}^{m, k}-\mathbf{n}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{n}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\Psi}^{m, k}\right) \\
-\mathbf{f}_{\boldsymbol{\psi}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}, \boldsymbol{N}_{A}^{-}, \boldsymbol{N}_{D}^{+}\right)
\end{array}\right],
\end{align*}
$$

where the indices are as explained in Section 3.1 along with following notations for compact representation.

$$
\begin{array}{lll}
\mathbf{f}_{\mathbf{p}, \mathbf{p}}=\partial \mathbf{f}_{\mathbf{p}} / \partial \mathbf{p} ; & \mathbf{f}_{\mathbf{p}, \mathbf{n}}=\partial \mathbf{f}_{\mathbf{p}} / \partial \mathbf{n} ; & \mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}}=\partial \mathbf{f}_{\mathbf{p}} / \partial \boldsymbol{\psi} \\
\mathbf{f}_{\mathbf{n}, \mathbf{p}}=\partial \mathbf{f}_{\mathbf{n}} / \partial \mathbf{p} ; & \mathbf{f}_{\mathbf{n}, \mathbf{n}}=\partial \mathbf{f}_{\mathbf{n}} / \partial \mathbf{n} ; & \mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}}=\partial \mathbf{f}_{\mathbf{n}} / \partial \boldsymbol{\psi}  \tag{3.32}\\
\mathbf{f}_{\psi, \mathbf{p}}=\partial \mathbf{f}_{\psi} / \partial \mathbf{p} ; & \mathbf{f}_{\psi, \mathbf{n}}=\partial \mathbf{f}_{\psi} / \partial \mathbf{n} ; & \mathbf{f}_{\psi, \boldsymbol{\psi}}=\partial \mathbf{f}_{\psi} / \partial \boldsymbol{\psi}
\end{array}
$$

There exists a unique solution to (3.29) if $\mathbf{f}_{\psi, \psi}$ is invertible in the neighborhood of the solution [61]. The Jacobian element $\mathbf{f}_{\psi, \psi}$, derived in Appendix A.2.1, turns out to be the stiffness matrix $\mathbf{S}$ of the Poisson's equation. The stiffness matrix arising from discretizing Poisson's equation using the finite element method is typically a symmetric positive definite matrix and hence invertible [59]. The Jacobian elements are derived in Appendix A. 2 and are different for the central difference and ScharfetterGummel methods. Regardless of the method used to spatially discretize $f_{p}$ or $f_{n}$, the general block-matrix structure of the Newton-Raphson iterator remains the same as predicted in (3.4).

### 3.5 Encapsulation

CCT equations describing the dynamics within the device is discretized using the techniques discussed in the previous section. The quantities of interest in a coupled device and electrical network simulation are the device terminal currents. Augmenting device terminal current equations to CCT equations allows the circuit to query the device model for terminal currents which are necessary in the solution of circuit variables. Similarly the device model can query the circuit model for the terminal voltages which affects the device operation. This process of augmenting device terminal current equations to the CCT equations is herein referred to as encapsulation. The necessary derivations for this encapsulation is elaborated here.

The terminal currents are obtained by scaling the current densities by the crosssectional area $a_{c}$ of the diode. The total anode and cathode terminal current densities are denoted by $J_{a}$ and $J_{c}$ respectively. These current densities are sum of drift and
diffusion current densities expressed in terms of scaled charge densities. The scaling factor $N_{M}$ is included in the calculations if scaled charge densities are used. The expressions for the current densities in terms of the charge densities are different depending on spatial discretization strategy used. The vector notations used to represent these currents are given below.

$$
\mathbf{f}_{\mathbf{i}_{d}}=\left[\begin{array}{l}
q_{e} N_{M} a_{c} \tilde{J}_{a}  \tag{3.33}\\
q_{e} N_{M} a_{c} \tilde{J}_{c}
\end{array}\right]=\left[\begin{array}{l}
i_{a} \\
i_{c}
\end{array}\right]=\mathbf{i}_{d}
$$

The equations of the terminal currents based on central difference and ScharfetterGummel discretization methods are given, respectively, in the following sets of equations.

$$
\begin{align*}
\mathbf{i}_{a}= & q_{e} N_{M} a_{c}\left[\frac{\psi_{1 / 2}}{h_{1}}\left(-\mu_{p} \frac{\left(p_{1}+p_{2}\right)}{2}-\mu_{n} \frac{\left(n_{1}+n_{2}\right)}{2}\right)\right. \\
& \left.+D_{n} \frac{\left(n_{2}-n_{1}\right)}{h_{1}}-D_{p} \frac{\left(p_{2}-p_{1}\right)}{h_{1}}\right]  \tag{3.34a}\\
\mathbf{i}_{c}= & q_{e} N_{M} a_{c}\left[\frac{\psi_{N-1 / 2}}{h_{N-1}}\left(-\mu_{p} \frac{\left(p_{N-1}+p_{N}\right)}{2}-\mu_{n} \frac{\left(n_{N-1}+n_{N}\right)}{2}\right)\right. \\
& \left.+D_{n} \frac{\left(n_{N}-n_{N-1}\right)}{h_{N-1}}-D_{p} \frac{\left(p_{N}-p_{N-1}\right)}{h_{N-1}}\right]  \tag{3.34b}\\
& \mathbf{i}_{a}= \\
& \frac{q_{e} N_{M} a_{c}}{h_{1}}\left[D_{p}\left(p_{1} \mathcal{B}\left(z_{3 / 2}\right)-p_{2} \mathcal{B}\left(-z_{3 / 2}\right)\right)\right.  \tag{3.35a}\\
& \left.-D_{n}\left(n_{1} \mathcal{B}\left(z_{3 / 2}\right)-n_{2} \mathcal{B}\left(-z_{3 / 2}\right)\right)\right] \\
\mathbf{i}_{c}= & \frac{q_{e} N_{M} a_{c}}{h_{N-1}}\left[D_{p}\left(p_{N-1} \mathcal{B}\left(z_{N-1 / 2}\right)-p_{N} \mathcal{B}\left(-z_{N-1 / 2}\right)\right)\right.  \tag{3.35b}\\
& \left.-D_{n}\left(n_{N-1} \mathcal{B}\left(z_{N-1 / 2}\right)-n_{N} \mathcal{B}\left(-z_{N-1 / 2}\right)\right)\right]
\end{align*}
$$

The vectorized terminal current equations are augmented to (3.29) resulting in (3.36). The current equations are algebraic and it is apparent that one block-row and blockcolumn needs to be added to (3.32).

$$
\begin{align*}
\frac{d \mathbf{p}}{d t} & =\mathbf{f}_{\mathbf{p}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) \\
\frac{d \mathbf{n}}{d t} & =\mathbf{f}_{\mathbf{n}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{3.36}\\
\mathbf{0} & =\mathbf{S} \boldsymbol{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\mathbf{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)-\mathbf{b}\left(v_{a}, v_{c}\right)=\mathbf{f}_{\psi}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) \\
\mathbf{0} & =\mathbf{i}_{d}-\mathbf{f}_{\mathbf{i}_{d}}=\mathbf{f}_{\mathbf{i}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})
\end{align*}
$$

For a unique solution to exist for (3.36), Jacobian $\mathbf{f}_{\mathbf{i}, \mathbf{i}}$ should be invertible in addition to $\mathbf{f}_{\psi, \psi}[61] . \mathbf{f}_{\psi, \psi}$ is the stiffness matrix and is invertible. $\mathbf{f}_{\mathbf{i}, \mathbf{i}}$ in this case is a $2 \times 2$ identity matrix that is invertible and hence there exists a unique solution. The following Newton iterator can be obtained for (3.36) without repeating derivations utilizing the pattern recognized in the previous section.

$$
\begin{align*}
& {\left[\begin{array}{cccc}
\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} & \mathbf{0} \\
-\mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} & \mathbf{0} \\
\mathbf{f}_{\psi, \mathbf{p}} & \mathbf{f}_{\psi, \mathbf{n}} & \mathbf{f}_{\psi, \boldsymbol{\psi}} & \mathbf{0} \\
\mathbf{f}_{\mathbf{i}, \mathbf{p}} & \mathbf{f}_{\mathbf{i}, \mathbf{n}} & \mathbf{f}_{\mathbf{i}, \boldsymbol{\psi}} & \mathbf{f}_{\mathbf{i}, \mathbf{i}_{d}}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{p} \\
\Delta \mathbf{n} \\
\Delta \boldsymbol{\psi} \\
\Delta \mathbf{i}_{d}
\end{array}\right]}  \tag{3.37}\\
& =\left[\begin{array}{c}
-\left(\mathbf{p}^{m, k}-\mathbf{p}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{p}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right) \\
-\left(\mathbf{n}^{m, k}-\mathbf{n}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{n}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right) \\
-\mathbf{f}_{\psi}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}, \boldsymbol{N}_{A}^{-}, \boldsymbol{N}_{D}^{+}\right) \\
-\mathbf{f}_{\mathbf{i}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right)
\end{array}\right]
\end{align*}
$$

The newly added Jacobian partitions in the last column are zero because the charge densities and electric potential within the device does not vary with a change in terminal current. Nonzero Jacobian partitions $\mathbf{f}_{\mathbf{i}, \mathbf{p}}, \mathbf{f}_{\mathbf{i}, \mathbf{n}}$ and $\mathbf{f}_{\mathbf{i}, \psi}$ due to these formulae has a similar structure and a dimension of $2 \times N$ with following definitions.

$$
\begin{align*}
\mathbf{f}_{\mathbf{i}, \mathbf{p}}=\partial \mathbf{f}_{\mathbf{i}} / \partial \mathbf{p} ; & \mathbf{f}_{\mathbf{i}, \mathbf{n}}=\partial \mathbf{f}_{\mathbf{i}} / \partial \mathbf{n}  \tag{3.38}\\
\mathbf{f}_{\mathbf{i}, \psi}=\partial \mathbf{f}_{\mathbf{i}} / \partial \mathbf{\psi} ; & \mathbf{f}_{\mathbf{i}, \mathbf{i}_{d}}=\partial \mathbf{f}_{\mathbf{i}} / \partial \mathbf{i}_{d}
\end{align*}
$$

Since the terminal currents depend on the charge densities and electrical potentials at the two boundary nodes near the anode and cathode, each of these partitions has only four nonzero entries. These partitions mathematically couple the boundaries of the semiconductor (ohmic contacts) with the external circuit via the boundary condition of the diode model. The nonzero pattern of each partition with a placeholder $\boldsymbol{\chi}$ for $\mathbf{p}$, $\mathbf{n}, \boldsymbol{\psi}$ is shown below and a detailed derivation of the nonzero terms in these partitions are derived in the Appendix A.2.5.

$$
\mathbf{f}_{\mathbf{i}, \mathbf{\chi}}=\left[\begin{array}{ccccc}
\partial i_{a} / \partial \chi_{1} & \partial i_{a} / \partial \chi_{2} & \ldots & 0 & 0  \tag{3.39}\\
0 & 0 & \ldots & \partial i_{c} / \partial \chi_{N-1} & \partial i_{c} / \partial \chi_{N}
\end{array}\right]_{(2 \times N)}
$$

The number of nonzeros in these partitions will be proportional to the number of Dirichlet boundary nodes in general for multi-dimensional device models.

### 3.6 Chapter Summary

The key contribution of this chapter is the derivation of an encapsulated device model with a consistent set of equations. Albeit nontrivial to implement, it is straightforward to extend this methodology to model multi-dimensional multiterminal semiconductor devices including bipolar junction (BJT), field-effect (FET), and insulated-gate bipolar (IGBT) transistors. This encapsulation is conducive to the addition and/or deletion of semiconductor physical processes modeled, spatial dimension and the number of device terminals considered. The external circuit needs to be modeled and the encapsulated device model is added to this circuit model in a consistent way. A procedure for assembling these equations is required which makes the coupled device-circuit simulation possible. Any aforementioned change in encapsulated device model does not change this procedure of assembling complete device-circuit system equations whatsoever.

## 4. COUPLED DEVICE AND ELECTRICAL NETWORK SIMULATION

Semiconductor devices (one or more) in power electronic circuits operate together with external sources and other circuit elements to control the flow of or convert electric power from one form (ac or dc) or voltage level to another. The sources and basic circuit elements such as resistors, inductors, and capacitors are typically modeled by ordinary differential equations (ODEs) using their respective lumped parameters. In order to simulate semiconductor devices and external circuit elements together, these ODEs are augmented to the encapsulated device model set forth in the previous chapter. The device and external electrical circuit models communicate with each other via the boundary conditions. A systematic procedure for coupling devices and external circuit elements is set forth in this chapter.

### 4.1 Diode Resistor Inductor Network



Fig. 4.1.: A simple $R L$ and diode circuit with no ground node.

A simple circuit, where a diode is connected in series with a voltage source, resistor and inductor $(R L)$, is shown in Figure 4.1. This simple circuit is considered for demonstrating the procedure of assembling equations from the netlist bearing in mind
that the proposed strategy is extendable to other complex circuit configurations. The circuit shown in Figure 4.1 is described in a netlist format in Table 4.1. The line number of the netlist is shown to the left of the box. A line-by-line description of the netlist is given below.

1. This line is reserved for the title of the circuit described by the netlist.
2. A $10-\Omega$ resistor R1 is connected between nodes 1 and 2 .
3. A $10-\mu \mathrm{H}$ inductance L 1 is connected between nodes 2 and 3 .
4. Current netlist line consists of a comment.
5. A diode D1 is connected between nodes 3 and 4. It is modeled using Coupled Carrier Transport (CCT) equations and initializing file is given by init.m.
6. A sinusoidal ac voltage source Vs is connected between nodes 1 and 4. It has an offset, amplitude, frequency and phase of $0,5 \mathrm{~V}, 5 \mathrm{kHz}$, and 0 respectively.
7. The end of the netlist file.

Current through two-terminal devices is assumed to be leaving the positive node and entering the negative node. These node numbers are listed in order from left to right after the element's unique name. The syntax for other basic circuit elements and independent sources are similar to that used in conventional SPICE netlists. When a physics-based diode model is used in the simulation, parasitic capacitors

Table 4.1.: Resistor-inductor-diode netlist.

```
simple RL and diode circuit
R 1 2 10
L 2 3 1e-5
! a distributed diode model
D 3 4 CCT init.m
Vs 1 4 ac sin(0 5 5000 0)
. END
```

are used to couple device and circuit as shown in Figure 4.2. The capacitors $C_{a}$ and $C_{c}$ denote the parasitic capacitances at the anode and cathode terminals of the diode. Their values are generally small and difficult to establish analytically; however, accurate knowledge of their values is not essential since the transients associated with these parasitic capacitances are very short-lived. In case an accurate estimate of these capacitors and other circuit parasitics are needed, then solution of the relevant equations from the set of Maxwell's equation is required along with knowledge of the physical placement of the circuit components and their interconnections. The encapsulated device model and external $R L$ circuit is coupled using the boundary conditions of the device model. These boundary conditions, namely the Dirichlet boundary, are obtained from the voltage across the parasitic capacitors.

It is necessary to build an incidence matrix to simulate the circuit in Figure 4.2. The positive-current-direction assumption from first node to the second (listed order in netlist) is used to represent a directed edge from first to the second. Each branch of the circuit is represented by the directed edges connected between nodes. This information is embedded in a matrix, known as incidence matrix, with number of rows and columns corresponding to total number of nodes and branches in the circuit, respectively. When branch $j$ leaves (enters) a node $i$, a nonzero value $+1(-1)$ is entered in $i^{\text {th }}$ row and $j^{\text {th }}$ column of the incidence matrix with rest of the entries being zero. Such an incidence matrix is built using the circuit in Figure 4.2.


Fig. 4.2.: $R L$ and diode circuit with parasitic capacitors and ground nodes added.

In a two-terminal device, current flowing into one terminal should equal the current flowing out of the other. However, the anode and cathode terminal currents of an encapsulated diode model can possibly be different for very brief periods of time (of the order of pico seconds). Consequently, the terminal currents are modeled as two dependent current sources that are functions of charge densities and electrostatic potential within the diode, and the diode terminal voltage. Two edges are used to represent these dependent current sources of an encapsulated diode model in the directed graph. Coupling the diode with the circuit is more amenable with insertion of parasitic capacitors $C_{a}, C_{c}$ and using dependent current sources $i_{a}, i_{c}$ for the terminal currents as illustrated in Figure 4.3.

The directed graph generated using this strategy is shown in Figure 4.4. There are 5 nodes (including the reference node) in the circuit considered. The number of branches including the added parasitic capacitance is 6 , when the diode is represented using single edge. The proposed approach, however, uses 7 branches where a diode

(a) Diode along with parasitic coupling capacitors and ground nodes.

(b) Diode-circuit coupling using dependent current sources.

Fig. 4.3.: Circuit-diode coupling.


Fig. 4.4.: Directed graph of the circuit in Figure 4.2.
is represented using two dependent current sources. Thus, the incidence matrix has a dimension of 5 rows and 7 columns.

Kirchhoff's current law (KCL) states that the algebraic sum of the currents entering any node is zero. Kirchhoff's voltage law (KVL) states that the algebraic sum of potential differences around any closed loop is zero. A KCL equation can be written for each node in the circuit. When all the branch currents are combined into a vector, an appropriate matrix can be obtained such that the matrix-vector product is the KCL equation for all nodes. This appropriate matrix is the incidence matrix described earlier. The resulting incidence matrix is linearly dependent. Any one row can be deleted from this matrix to obtain a linearly independent incidence matrix. Conventionally, the datum, reference, or ground node equation is removed. The resulting incidence matrix is called the reduced incidence matrix. Hereafter, the term incidence matrix (A) is used to refer to the reduced incidence matrix for simplicity.

With this foresight, KCL equations are written only for nodes 1 through 4 as given in the following equations. Each branch current is denoted using $i$ subscripted with the branch name, except for the current source branches where the values are used right away.

$$
\begin{array}{r}
i_{R}+i_{V_{s}}=0 \\
-i_{R}+i_{L}=0  \tag{4.1}\\
-i_{L}+i_{C_{a}}+i_{a}=0 \\
-i_{c}+i_{C_{c}}-i_{V_{s}}=0
\end{array}
$$

This equation is written as matrix-vector product as follows.

$$
\left[\begin{array}{cc|c|c|c|cc}
0 & 0 & 1 & 0 & 1 & 0 & 0  \tag{4.2}\\
0 & 0 & -1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & -1 & 0 & -1
\end{array}\right]\left[\begin{array}{c}
i_{C_{a}} \\
\frac{i_{C_{c}}}{i_{R}} \\
\hline \frac{i_{L}}{i_{V_{s}}} \\
i_{a} \\
i_{c}
\end{array}\right]=\left[\begin{array}{l}
0 \\
0 \\
0 \\
0
\end{array}\right]
$$

These 4 KCL equations relate 7 branch currents. The matrix on the left-hand side is the incidence matrix ( $\mathbf{A}$ ) whose dimension is $4 \times 7$. The branch currents can be grouped by branch type and reordered to form a vector. It is convenient to represent this branch current vector as $\mathbf{i}=\left[\begin{array}{lllllll}i_{C_{a}} & i_{C_{c}} & i_{R} & i_{L} & i_{V_{s}} & i_{a} & i_{c}\end{array}\right]^{T}$ (superscript $T$ denotes transpose). Then KCL is expressed compactly as a partitioned matrix-vector products as follows,

$$
\begin{align*}
{\left[\begin{array}{lllll}
\mathbf{A}_{C} & \mathbf{A}_{R} \mathbf{A}_{L} & \mathbf{A}_{V} & \mathbf{A}_{I}
\end{array}\right]\left[\begin{array}{c}
\mathbf{i}_{C} \\
\mathbf{i}_{R} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V} \\
\mathbf{i}_{I}
\end{array}\right] } & =\left[\begin{array}{l}
\mathbf{0} \\
\mathbf{0} \\
\mathbf{0} \\
\mathbf{0} \\
\mathbf{0}
\end{array}\right],  \tag{4.3}\\
\mathbf{A i} & =\mathbf{0} \tag{4.4}
\end{align*}
$$

The vector of branch currents grouped by branch type is denoted by $\mathbf{i}_{k}$, where $k$ is one of $\{R, L, C, V, I\}$ corresponding to resistors, inductors, capacitors, voltage sources and current sources, respectively. The number of each branch type is denoted by $n_{k}$, which implies each $\mathbf{i}_{k}$ is an $n_{k}$-dimensional vector. A similar notation is used for other partitions of the incidence matrix as shown in Table 4.2.

The voltage across each branch is the difference between the node voltages to which the branches are connected. These equations are explicitly expressed as,

$$
\begin{align*}
v_{C_{a}} & =v_{3}, \\
v_{C_{c}} & =v_{4}, \\
v_{R} & =v_{1}-v_{2}, \\
v_{L} & =v_{2}-v_{3},  \tag{4.5}\\
V_{s} & =v_{1}-v_{4}, \\
v_{I a} & =v_{3}, \\
v_{I c} & =-v_{4} .
\end{align*}
$$

These equations are rewritten to a form similar to the KCL equations with similar notations. The vector of voltages of a specific branch type is denoted as $\mathbf{v}_{k}$, an $n_{k^{-}}$

Table 4.2.: Incidence matrix partitions.

| Variable | Description with dimension in parentheses |
| :---: | :---: |
| $n_{n}$ | number of nodes in the circuit |
| $n_{b}$ | number of branches in the circuit $\left(n_{b}=n_{C}+n_{R}+n_{L}+n_{V}+n_{I}\right)$ |
| $\mathbf{A}_{C}$ | Capacitance incidence matrix $\left(n_{n}-1 \times n_{C}\right)$ |
| $\mathbf{A}_{R}$ | Resistance incidence matrix $\left(n_{n}-1 \times n_{R}\right)$ |
| $\mathbf{A}_{L}$ | Inductance incidence matrix $\left(n_{n}-1 \times n_{L}\right)$ |
| $\mathbf{A}_{V}$ | Voltage source incidence matrix $\left(n_{n}-1 \times n_{V}\right)$ |
| $\mathbf{A}_{I}$ | Current source incidence matrix $\left(n_{n}-1 \times n_{I}\right)$ |
| $\mathbf{A}$ | Incidence matrix $\left(n_{n}-1 \times n_{b}\right)$ |

dimensional vector for all $k$ in $\{R, L, C, V, I\}$. The branch voltages are denoted in a vector form using $\mathbf{v}_{b}$, which is an $n_{b}$-dimensional vector. Matrix-vector-product form or partitioned-matrix-vector-product forms of the voltage equations are given below.

$$
\begin{align*}
{\left[\begin{array}{l}
v_{C_{a}} \\
v_{C_{c}} \\
\hline v_{R} \\
\hline v_{L} \\
\hline V_{s} \\
\hline v_{I a} \\
v_{I c}
\end{array}\right] } & =\left[\begin{array}{llcc}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\hline 1 & -1 & 0 & 0 \\
\hline 0 & 1 & -1 & 0 \\
\hline 1 & 0 & 0 & -1 \\
\hline 0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1
\end{array}\right]\left[\begin{array}{l}
v_{1} \\
v_{2} \\
v_{3} \\
v_{4}
\end{array}\right]  \tag{4.6}\\
{\left[\begin{array}{l}
\mathbf{v}_{C} \\
\mathbf{v}_{R} \\
\mathbf{v}_{L} \\
\mathbf{v}_{s} \\
\mathbf{v}_{I}
\end{array}\right] } & =\left[\begin{array}{l}
\mathbf{A}_{C}^{T} \\
\mathbf{A}_{R}^{T} \\
\mathbf{A}_{L}^{T} \\
\mathbf{A}_{V}^{T} \\
\mathbf{A}_{I}^{T}
\end{array}\right]  \tag{4.7}\\
\mathbf{v}_{b} & =\mathbf{A}^{T} \mathbf{v} \tag{4.8}
\end{align*}
$$

### 4.1.1 Models of basic circuit elements

Current through a capacitor, with a fixed capacitance $C$, is the product of its capacitance and the rate of change of applied voltage. This calculation is done for all capacitors in a given network. Voltage across all capacitors in the circuit is given by $\mathbf{v}_{C}=\mathbf{A}_{C}^{T} \mathbf{v}$. It is convenient to define a capacitance matrix $\mathbf{C}$ that is an $n_{C} \times n_{C}$ diagonal matrix with its nonzero values corresponding to each capacitor's value in the network. Then, the capacitor currents are derived using the capacitance matrix and voltage across the capacitors, i.e.

$$
\begin{equation*}
\mathbf{i}_{C}=\mathbf{C} \frac{d \mathbf{v}_{C}}{d t}=\mathbf{C A}_{C}^{T} \frac{d \mathbf{v}}{d t} \tag{4.9}
\end{equation*}
$$

Similarly, all resistor currents are computed with an $n_{R} \times n_{R}$ diagonal conductance matrix $\mathbf{G}$ and the voltage across the resistors found using $\mathbf{A}_{R}^{T}$. These resistor currents are given by,

$$
\begin{equation*}
\mathbf{i}_{R}=\mathbf{G} \mathbf{v}_{R}=\mathbf{G} \mathbf{A}_{R}^{T} \mathbf{v} \tag{4.10}
\end{equation*}
$$

The voltage across an inductor is the rate of change of flux linking it. This flux linkage $(\boldsymbol{\lambda})$ for all inductive branches is conveniently expressed using an inductance matrix $\mathbf{L}$ and respective branch current vector $\mathbf{i}_{L}$. Then, the inductor voltages are given by,

$$
\begin{equation*}
\mathbf{v}_{L}=\frac{d \boldsymbol{\lambda}}{d t}=\mathbf{L} \frac{d \mathbf{i}_{L}}{d t}=\mathbf{A}_{L}^{T} \mathbf{v} \tag{4.11}
\end{equation*}
$$

Equations (4.4), (4.8), (4.9), (4.10), (4.11) along with voltage and current source equations can be used to build a tableau and simulate the given circuit. The tableau thus derived may not appear to have an algebraic structure. Some structure in these equations is obtained by manipulating (4.4) with the help of basic circuit element equations to obtain the following

$$
\begin{equation*}
\mathbf{A}_{C} \mathbf{C} \mathbf{A}_{C}^{T} \frac{d \mathbf{v}}{d t}+\mathbf{A}_{R} \mathbf{G} \mathbf{A}_{R}^{T} \mathbf{v}+\mathbf{A}_{L} \mathbf{i}_{L}+\mathbf{A}_{V} \mathbf{i}_{V}+\mathbf{A}_{I} \mathbf{i}_{S}=\mathbf{0} \tag{4.12}
\end{equation*}
$$

Equation (4.12) is augmented with inductor and source voltage equations. The resulting set of equations comprise the so-called Modified Nodal Analysis (MNA) equations [62]. The index analysis of this system of differential-algebraic equations and its
solvability is thoroughly analyzed in [63]. The set of equations thus derived becomes,

$$
\begin{align*}
\mathbf{A}_{C} \mathbf{C} \mathbf{A}_{C}^{T} \frac{d \mathbf{v}}{d t}+\mathbf{A}_{R} \mathbf{G} \mathbf{A}_{R}^{T} \mathbf{v}+\mathbf{A}_{L} \mathbf{i}_{L}+\mathbf{A}_{V} \mathbf{i}_{V}+\mathbf{A}_{I} \mathbf{i}_{S} & =\mathbf{0}  \tag{4.13a}\\
\mathbf{L} \frac{d \mathbf{i}_{L}}{d t} & =\mathbf{A}_{L}^{T} \mathbf{v}  \tag{4.13b}\\
\mathbf{A}_{V}^{T} \mathbf{v} & =\mathbf{v}_{s} \tag{4.13c}
\end{align*}
$$

The circuit dynamics are modeled using (4.13). A brief note on the sign of computed source currents $\mathbf{i}_{V}$ is due at this point. The computed source currents will have a negative sign, since current is assumed to flow from positive (node 1) to negative (node 2) in a two-terminal device. If positive source currents are desired, then replacing $\mathbf{A}_{V}$ with $-\mathbf{A}_{V}$ resolves the issue. The encapsulated device model is augmented with (4.13). With some care, augmenting/removing multiple encapsulated device models is straightforward. This is a desirable attribute for implementing variablestructure variable-order simulation strategies discussed in Chapter 7. This procedure of assembling the complete system of equations is delineated in the following section.

### 4.2 Coupled Circuit and Device Simulation

Manually assembling the equations involved in simulating coupled device-circuits is feasible but is time consuming and tedious. Appendix B contains a manual derivation of the device-circuit system equations for the example circuit in Figure 4.1. A similar derivation must be repeated for any change in network topology. On the other hand automatically assembling the necessary equations enables a quick setup of complex circuit simulations. It also requires less effort from the modeling engineer and reduces the derivation/programming errors.

To this end, modified nodal analysis (MNA) is used. A detailed derivation of the MNA equations was discussed in the previous section. The number of unknowns in the MNA approach is $N_{M N A}=n_{n}+n_{L}+n_{V}-1$, which is the sum of dimensions of the vectors $\mathbf{v}, \mathbf{i}_{L}$ and $\mathbf{i}_{V}$. A vector of circuit variables is defined as $\mathbf{y}_{\mathrm{ckt}}=\left[\mathbf{v} \mathbf{i}_{L} \mathbf{i}_{V}\right]^{T}$.

This definition and a rewritten form of (4.13) gives the equations in (4.14). This form helps in coupling the circuit and semiconductor device.

$$
\begin{align*}
{\left[\begin{array}{ccc}
\mathbf{A}_{C} \mathbf{C A}_{C}^{T} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{L} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right] \frac{d}{d t}\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V}
\end{array}\right]=} & {\left[\begin{array}{ccc}
-\mathbf{A}_{R} \mathbf{G} \mathbf{A}_{R}^{T} & -\mathbf{A}_{L} & -\mathbf{A}_{V} \\
\mathbf{A}_{L}^{T} & \mathbf{0} & \mathbf{0} \\
-\mathbf{A}_{V}^{T} & \mathbf{0} & \mathbf{0}
\end{array}\right]\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V}
\end{array}\right] } \\
& +\left[\begin{array}{c}
-\mathbf{A}_{I} \mathbf{i}_{s} \\
\mathbf{0} \\
\mathbf{v}_{s}
\end{array}\right] . \tag{4.14}
\end{align*}
$$

The voltage and current source vectors $\mathbf{i}_{s}$ and $\mathbf{v}_{s}$, respectively, are the inputs to the circuit equations. A combination of independent and dependent current sources can inject currents into the circuit. The current source vector and the respective incidence matrix is partitioned by independent and dependent sources. This partitioning facilitates coupling (4.14) and the encapsulated device model equations.

The current source vector $\mathbf{i}_{s}$ is partitioned such that $\mathbf{i}_{s}=\left[\mathbf{i}_{\text {ind }} \mathbf{i}_{\text {dep }}\right]^{T}$. The vectors of independent and dependent current sources are represented by $\mathbf{i}_{\text {ind }}$ and $\mathbf{i}_{\text {dep }}$, respectively. This leads to a corresponding partitioning in the incidence matrix due to current sources as in $\mathbf{A}_{I}=\left[\mathbf{A}_{\text {ind }} \mathbf{A}_{\text {dep }}\right]$. MNA equations are rewritten using this partitioning in the following form.

$$
\begin{align*}
{\left[\begin{array}{ccc}
\mathbf{A}_{C} \mathbf{C A}_{C}^{T} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{L} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right] \frac{d}{d t}\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V}
\end{array}\right]=} & {\left[\begin{array}{ccc}
-\mathbf{A}_{R} \mathbf{G} \mathbf{A}_{R}^{T} & -\mathbf{A}_{L} & -\mathbf{A}_{V} \\
\mathbf{A}_{L}^{T} & \mathbf{0} & \mathbf{0} \\
-\mathbf{A}_{V}^{T} & \mathbf{0} & \mathbf{0}
\end{array}\right]\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V}
\end{array}\right] } \\
& +\left[\begin{array}{c}
-\mathbf{A}_{\text {dep }} \mathbf{i}_{\text {dep }} \\
\mathbf{0} \\
\mathbf{0}
\end{array}\right]+\left[\begin{array}{c}
-\mathbf{A}_{\text {ind }} \mathbf{i}_{\text {ind }} \\
\mathbf{0} \\
\mathbf{v}_{s}
\end{array}\right] \tag{4.15}
\end{align*}
$$

The discretized diode equations are given in (4.16) for reference. In (4.16c), $\mathbf{b}\left(v_{a}, v_{c}\right)$ is a vector that takes care of the boundary potential due to the zero-spacecharge assumption at the ohmic contacts and diode applied voltages. The diode model should get these voltages from the solution of the circuit equations. Consequently it
is necessary to rewrite this vector as the sum of a boundary potential vector $\mathbf{v}_{b}$ and diode applied voltages $\mathbf{v}_{d}$ as in (4.17). Similarly, the diode current vector $\mathbf{i}_{d}$ is given as an input to the MNA equations via the dependent current source partitions in the right-hand side of (4.15).

$$
\begin{gather*}
\frac{d \mathbf{p}}{d t}=\mathbf{f}_{\mathbf{p}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{4.16a}\\
\frac{d \mathbf{n}}{d t}=\mathbf{f}_{\mathbf{n}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{4.16b}\\
\mathbf{f}_{\psi}=\mathbf{S} \mathbf{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\mathbf{N}_{D}^{+}-\mathbf{N}_{A}^{-}\right)-\mathbf{b}\left(v_{a}, v_{c}\right)=\mathbf{0}  \tag{4.16c}\\
\mathbf{i}_{d}=\mathbf{f}_{\mathbf{i}_{d}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\Psi})  \tag{4.16d}\\
\mathbf{b}\left(v_{a}, v_{c}\right)=\left[\begin{array}{c}
v_{a}-\frac{k T}{q} \log \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right) \\
0 \\
\vdots \\
0 \\
v_{c}+\frac{k T}{q} \log \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right)
\end{array}\right]=\left[\begin{array}{c}
v_{a} \\
0 \\
\vdots \\
0 \\
v_{c}
\end{array}\right]+\left[\begin{array}{c}
-\frac{k T}{q} \log \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right) \\
0 \\
\vdots \\
0 \\
\frac{k T}{q} \log \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right)
\end{array}\right] \tag{4.17}
\end{gather*}
$$

It is convenient to define all of the diode variables using a vector, $\mathbf{y}_{d}=\left[\begin{array}{lll}\mathbf{p} & \mathbf{n} & \boldsymbol{\psi} \\ \mathbf{i}_{d}\end{array}\right]^{T}$. Equations in (4.16) along with (4.17) is rewritten as,

The vectors $\mathbf{v}_{d}$ and $\mathbf{i}_{\text {dep }}$ play a key role in setting up coupled circuit-device simulation. All that is left is to determine matrices that map circuit variables to the appropriate diode $\mathbf{v}_{d}$ and diode currents to the appropriate $\mathbf{i}_{\text {dep }}$. The circuit-variable-to-diode-voltage vector mapping is derived for the simple circuit considered in Fig-
ure 4.2 in the following equations. The nonzero entries in matrix $\mathbf{A}_{d, \mathrm{ckt}}$ are selected such that the matrix maps the circuit variables to diode voltage vector $\mathbf{v}_{d}$.

$$
-\left[\begin{array}{c}
\mathbf{0}  \tag{4.19}\\
\mathbf{0} \\
\mathbf{v}_{d} \\
\mathbf{0}
\end{array}\right]_{(3 N+2) \times 1}=\mathbf{A}_{d, \mathrm{ckt}} \mathbf{u}_{\mathrm{ckt}}
$$

A one-dimensional diode model is considered here. For the mesh-node numbering convention used in this research, Node 1 and Node $N$ are the anode and cathode terminals, respectively. The dimension of $\mathbf{A}_{d, \text { ckt }}$ should be $(3 N+1) \times N_{M N A}$. The nonzero entries appear in Rows $2 N+1$ and $3 N$, and, Columns 3 and 4 of $\mathbf{A}_{d, \mathrm{ckt}}$ for the example circuit. It is given explicitly in the following equation for clarity, with $n_{a}=3$ and $n_{c}=4$.

$$
\begin{gather*}
\mathbf{A}_{d, \mathrm{ckt}}=\begin{array}{ccccccc}
1 & \ldots & n_{a} & \ldots & n_{c} & \ldots & N_{M N A} \\
1 \\
\vdots \\
2 N+1 \\
\vdots \\
3 N \\
3 N+1 \\
3 N+2
\end{array}\left(\begin{array}{ccccccc}
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\
\vdots & & & \ddots & & & \vdots \\
0 & \ldots & -1 & \ldots & 0 & \ldots & 0 \\
\vdots & & & \ddots & & & \vdots \\
0 & \ldots & 0 & \ldots & -1 & \ldots & 0 \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0
\end{array}\right) \tag{4.20}
\end{gather*}
$$

A short-flat border matrix $\mathbf{A}_{\text {ckt }, d}$ is formed that maps the diode currents to the appropriate nodes to which the diode terminals are connected and KCL equations are written. The nonzero entries of this matrix are picked such that,

$$
-\left[\begin{array}{ccc}
\mathbf{A}_{d s} & \mathbf{0} & 0  \tag{4.21}\\
\mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right]\left[\begin{array}{c}
\mathbf{i}_{\mathrm{dep}} \\
\mathbf{0} \\
\mathbf{0}
\end{array}\right]=\mathbf{A}_{\mathrm{ckt}, d} \mathbf{y}_{d}
$$

This implies that the dimension of $\mathbf{A}_{\text {ckt }, d}$ matrix is $N_{M N A} \times(3 N+2)$. The nonzero entries of this matrix will be in the rows $n_{a}$ and $n_{c}$, and, the last two columns. Once again, the values of this matrix are given explicitly for the example circuit considered.

### 4.2.1 Interpretation of interface matrices

It is useful to discuss the relation between interface matrices $\mathbf{A}_{\text {ckt }, d}$, and, $\mathbf{A}_{d, \mathrm{ckt}}$ and different incidence matrix partitions. From (4.21), it is obvious that $\mathbf{A}_{\text {ckt }, d}$ is a slightly reformed version of dependent current source incidence matrix partition $\mathbf{A}_{\text {dep }}$ due to the devices. When a multi-terminal device is modeled, this matrix is to be formed accordingly. The tall $\mathbf{A}_{d, \mathrm{ckt}}$ interface matrix is also a reformed version of an incidence matrix partition. To be specific, the incidence matrix partition due to the coupling capacitors is used to obtain the nonzero entries in this tall matrix. It is interesting to note that the nonzero entries correspond to the Dirichlet boundary nodes of the encapsulated device model equations. From these two observations, it is apparent that, as long as the interface matrices are built consistently, the procedure of assembling equations is opaque to the changes in the spatial dimension of the device models and number of device terminals.

At this point, all the equations, matrices, and vectors are readily available to write the equations that simulates the example circuit considered in Figure 4.2. Several
matrix and vector notations are used that will help the final system of equations to fit the page width. The matrix notations are given below.

$$
\begin{align*}
\mathbf{M}_{d}=\left[\begin{array}{cccc}
\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right] \quad & \mathbf{A}_{d}=\left[\begin{array}{cccc}
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
-\mathbf{D} & \mathbf{D} & \mathbf{S} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}
\end{array}\right]  \tag{4.23}\\
\mathbf{M}_{\mathrm{ckt}}=\left[\begin{array}{cccc}
\mathbf{A}_{C} \mathbf{C} \mathbf{A}_{C}^{T} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{L} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0}
\end{array}\right] & \mathbf{A}_{\mathrm{ckt}}=\left[\begin{array}{cccc}
-\mathbf{A}_{R} \mathbf{G} \mathbf{A}_{R}^{T} & -\mathbf{A}_{L} & -\mathbf{A}_{V} \\
\mathbf{A}_{L}^{T} & \mathbf{0} & \mathbf{0} \\
-\mathbf{A}_{V}^{T} & \mathbf{0} & \mathbf{0}
\end{array}\right]
\end{align*}
$$

The vector notations are as follows:

$$
\begin{align*}
& \mathbf{y}_{d}=\left[\begin{array}{l}
\mathbf{p} \\
\mathbf{n} \\
\boldsymbol{\psi} \\
\mathbf{i}_{d}
\end{array}\right] \quad \mathbf{y}_{\mathrm{ckt}}=\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{i}_{L} \\
\mathbf{i}_{V}
\end{array}\right] \\
& \mathbf{f}_{d}=\left[\begin{array}{c}
\mathbf{f}_{\mathbf{p}} \\
-\mathbf{f}(\mathbf{p}-\mathbf{n})-\mathbf{v}_{B} \\
-\mathbf{f}_{\mathbf{i}}
\end{array}\right] \quad \mathbf{u}_{\mathrm{ckt}}=\left[\begin{array}{c}
-\mathbf{A}_{\mathrm{ind}} \mathbf{i}_{\mathrm{ind}} \\
\mathbf{0} \\
\mathbf{v}_{s}
\end{array}\right] \tag{4.24}
\end{align*}
$$

The dimensions of these matrix and vector notations are given below for reference.

$$
\begin{align*}
\mathbf{M}_{d}, \mathbf{A}_{d} & \in \mathbb{R}^{(3 N+2) \times(3 N+2)} \\
\mathbf{M}_{\text {ckt }}, \mathbf{A}_{\text {ckt }} & \in \mathbb{R}^{N_{M N A} \times N_{M N A}} \\
\mathbf{A}_{\text {ckt }, d} & \in \mathbb{R}^{N_{M N A} \times(3 N+2)}  \tag{4.25}\\
\mathbf{A}_{d, \text { ckt }} & \in \mathbb{R}^{(3 N+2) \times N_{M N A}} \\
\mathbf{y}_{d}, \mathbf{f}_{d} & \in \mathbb{R}^{3 N+2} \\
\mathbf{y}_{\mathrm{ckt}}, \mathbf{u}_{\mathrm{ckt}} & \in \mathbb{R}^{N_{M N A}}
\end{align*}
$$

Equations (4.15),(4.18),(4.19) and (4.21) are combined into a single equation with the help of (4.24).

$$
\left[\begin{array}{cc}
\mathbf{M}_{d} & \mathbf{0}  \tag{4.26}\\
\mathbf{0} & \mathbf{M}_{\mathrm{ckt}}
\end{array}\right] \frac{d}{d t}\left[\begin{array}{c}
\mathbf{y}_{d} \\
\mathbf{y}_{\mathrm{ckt}}
\end{array}\right]=\left[\begin{array}{cc}
\mathbf{A}_{d} & \mathbf{A}_{d, \mathrm{ckt}} \\
\mathbf{A}_{\mathrm{ckt}, d} & \mathbf{A}_{\mathrm{ckt}}
\end{array}\right]\left[\begin{array}{c}
\mathbf{y}_{d} \\
\mathbf{y}_{\mathrm{ckt}}
\end{array}\right]+\left[\begin{array}{c}
\mathbf{f}_{d} \\
\mathbf{u}_{\mathrm{ckt}}
\end{array}\right]
$$

This set of differential-algebraic equations is then integrated with respect to time using any choice of stiffly-stable solver. The procedures and notations are in place for extending this formulation to a circuit with multiple diodes. The procedure is listed below:

1. Parse the given SPICE like netlist.
2. Process the basic circuit elements and sources as usual. Keep track of distributed devices as they are encountered in the netlist and add terminal-parasitic capacitors as needed. These capacitors are processed as any other capacitor in the network.
3. Build $\mathbf{M}_{\mathrm{ckt}}$ and $\mathbf{A}_{\mathrm{ckt}}$ matrices.
4. For each terminal of the distributed device add a dependent current source with appropriate direction and assemble $\mathbf{M}_{d}$ and $\mathbf{A}_{d}$.
5. Build interface matrices $\mathbf{A}_{\text {ckt }, d}$ and $\mathbf{A}_{d, \text { ckt }}$ according to the number of terminals and spatial dimension of device considered. In this thesis, they are 2 and 1 respectively
6. Insert the encapsulated diode equations along with the interface matrices in order.
7. Insert the circuit equations in the end to result in the full system equation.

A single-phase diode bridge rectifier is shown in Figure 4.5. This circuit has 4 diodes and according to the proposed strategy there will be 8 dependent current sources and parasitic capacitors inserted in addition to the existing components.


Fig. 4.5.: Single-phase diode bridge rectifier with resistor-inductor load.

The single diode and external circuit model equation (4.26) is extended to a singlephase diode bridge rectifier. It is relatively straightforward to simulate a complex circuit with multiple diodes using the strategy set forth. Without loss of generality, an abstract and compact equation can be used to describe the dynamics of a coupled semiconductor device and circuit problem as in the following equation.

$$
\begin{equation*}
\mathbf{M} \frac{d \mathbf{y}}{d t}=\mathbf{A y}+\mathbf{f}\left(\mathbf{y}, \mathbf{i}_{\text {ind }}, \mathbf{v}_{s}\right) \tag{4.27}
\end{equation*}
$$

where for a single-phase diode bridge rectifier,

$$
\begin{align*}
\mathbf{M} & =\left[\begin{array}{ccccc}
\mathbf{M}_{d_{1}} & 0 & 0 & 0 & 0 \\
0 & \mathbf{M}_{d_{2}} & 0 & 0 & 0 \\
0 & 0 & \mathbf{M}_{d_{3}} & 0 & 0 \\
0 & 0 & 0 & \mathbf{M}_{d_{4}} & 0 \\
0 & 0 & 0 & 0 & \mathbf{M}_{\mathrm{ckt}}
\end{array}\right]  \tag{4.28a}\\
\mathbf{y} & =\left[\begin{array}{l}
\mathbf{y}_{d_{1}} \\
\mathbf{y}_{d_{2}} \\
\mathbf{y}_{d_{3}} \\
\mathbf{y}_{d_{4}} \\
\mathbf{y}_{\mathrm{ckt}}
\end{array}\right] \tag{4.28b}
\end{align*}
$$

$$
\begin{align*}
& \mathbf{A}=\left[\begin{array}{ccccc}
\mathbf{A}_{d_{1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{d_{1}, \mathrm{ckt}} \\
\mathbf{0} & \mathbf{A}_{d_{2}} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{d_{2}, \mathrm{ckt}} \\
\mathbf{0} & \mathbf{0} & \mathbf{A}_{d_{3}} & \mathbf{0} & \mathbf{A}_{d_{3}, \mathrm{ckt}} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{d_{4}} & \mathbf{A}_{d_{4}, \mathrm{ckt}} \\
\mathbf{A}_{\mathrm{ckt}, d_{1}} & \mathbf{A}_{\mathrm{ckt}, d_{2}} & \mathbf{A}_{\mathrm{ckt}, d_{3}} & \mathbf{A}_{\mathrm{ckt}, d_{4}} & \mathbf{A}_{\mathrm{ckt}}
\end{array}\right]  \tag{4.28c}\\
& \mathbf{f}=\left[\begin{array}{c}
\mathbf{f}_{d_{1}} \\
\mathbf{f}_{d_{2}} \\
\mathbf{f}_{d_{3}} \\
\mathbf{f}_{d_{4}} \\
\mathbf{u}_{\mathrm{ckt}}
\end{array}\right] \tag{4.28d}
\end{align*}
$$

The vector $\mathbf{y}$ is a vector of the combined diode and circuit variables. The independent voltage $\mathbf{v}_{s}$ and current $\mathbf{i}_{\text {ind }}$ are the inputs. This system of differential-algebraic equations is integrated with respect to time using a stiffly-stable integrator. A common choice of stiffly-stable integration algorithm is the backward Euler algorithm. Since the BE algorithm is implicit, a linear system is solved within Newton-Raphson iterations for solving this type of nonlinear initial-value problem. The right-hand side is expressed in a form that helps in deriving the Jacobian as the sum of a linear nonvarying part, and, a nonlinear time-varying part which, in turn, is used in the linear solution. The solution methodology and implementation of the implicit integration algorithm is discussed in detail in the following chapter.

### 4.3 Comparison with Other Formulations

Xyce ${ }^{\mathrm{TM}}$ is a parallel circuit simulator developed by Sandia Laboratories. The circuit simulator is targeted towards coupled device-circuit simulation of VLSI circuits. The underlying mathematical formulation is discussed in [64]. Xyce forces the terminal currents to be equal to that computed using the device physics. That is, a device level KCL is enforced. In addition, in order to deal with nonlinearities in the simulation the device voltage is not allowed to change significantly (by how much is
unspecified). This operation, so-called 'voltage limiting', apparently helps in the convergence of nonlinear iterations. Reference [64] mentions briefly that this operation is inconsistent and adds numerical algorithm implementation difficulties.

A similar coupled device-circuit simulation formulation is the topic of a dissertation [63]. This dissertation is also targeted towards VLSI circuits where KCL is enforced at the device level more so in a mathematical way. For charge conservation, the divergence of hole and electron density less displacement current should be zero. The displacement current is obtained by finding the time derivative of Poisson's equation. In case of diodes, the charge conservation equation is manipulated to make the anode and cathode current equal. Then it is sufficient to include only one terminal current in circuit formulation with similar extension to other multi-terminal devices. In order to compute the displacement current, it is necessary to track the electric field near the semiconductor-metal contacts (terminals). It is apparent that the boundary electric fields are added states, which increase as the spatial dimension of the device model goes from one through three.

The device terminal currents are algebraic functions of the charge densities and electric potentials at the boundaries in the proposed strategy. No additional state needs to be considered because electric field is computed as a function of electric potential. The formulation set forth in this thesis does not force KCL at the device level and also no voltage limiting is used. There are several terminal current equations added to CCT when compared to Xyce and no extra states are added when compared to [63].

### 4.4 Chapter Summary

Circuit model development using Modified Nodal Analysis is reviewed and a systematic procedure for assembling device-circuit system equations is set forth. Extension of the discussed concepts to different, possibly larger, circuit topologies is straightforward. Currently, the implementation supports only resistors, inductors,
capacitors, sources, and diodes. When coupled elements such as transformers or multiterminal circuit elements are used, the formulation requires an appropriate update for the circuit equations. A similar update is required when multi-terminal semiconductor devices are included. As long as the interface matrices are built in consistence with the number of terminals and spatial dimension considered, the procedure set forth herein is equally applicable to any device modeled using device physics.

## 5. TEMPORAL INTEGRATION AND PARAMETER EXTRACTION

A systematic procedure set forth in the previous chapter is used to assemble the encapsulated diode model and circuit equations to result in a system of differential-algebraic equations (DAEs). The coupled device-circuit network is simulated by implementing a temporal integration for the DAE system derived. Implementation of backward Euler method for DAE system is discussed in this chapter. In addition, a procedure to extract the physical parameters of a PIN diode based upon its measured response is set forth.

### 5.1 Backward Euler Temporal Integration

It is useful to consider the system of DAE (5.1b) with mass matrix M, which can be singular. This system is usually obtained after the spatial discretization of the semiconductor devices and augmenting the circuit equations as derived in previous chapter. The vector $\mathbf{u}$ consists of semiconductor species densities, electric potential and circuit variables combined as depicted in (4.26). The backward Euler algorithm is applied to (5.1b) to get (5.1c).

$$
\begin{align*}
\mathbf{M} \frac{d \mathbf{y}}{d t} & =\mathbf{A} \mathbf{u}+\mathbf{f}\left(\mathbf{u}, \mathbf{i}_{\text {ind }}, \mathbf{v}_{s}\right)  \tag{5.1a}\\
\mathbf{M} \frac{d \mathbf{y}}{d t} & =\mathbf{g}\left(\mathbf{u}, \mathbf{i}_{\text {ind }}, \mathbf{v}_{s}\right)  \tag{5.1b}\\
\mathbf{M}\left(\mathbf{y}^{m}-\mathbf{y}^{m-1}\right) & =\Delta t^{m} \mathbf{g}\left(\mathbf{y}^{m}\right) \tag{5.1c}
\end{align*}
$$

Equation (5.1c) represents a nonlinear system of equations. A Newton-Raphson (NR) iteration is commonly used to solve nonlinear systems of this form. The NR iterator is obtained by expanding (5.1c) using a Taylor's series about $\mathbf{u}^{m}$ up to firstorder terms and equating the resulting expression to zero. Instead of solving for
$\mathbf{y}^{m}$ directly, the increment $\mathbf{z}=\mathbf{y}^{m}-\mathbf{y}^{m-1}$ is solved. The right-hand side gradient $\mathbf{g}$ is rewritten as $\mathbf{g}_{\mathbf{z}}(\mathbf{z})=\mathbf{g}\left(\mathbf{y}^{m-1}+\mathbf{z}\right)$. This implementation is based on implicit Runge-Kutta methods explained in [61].

$$
\begin{equation*}
\mathbf{M z}-\Delta t^{m} \mathbf{g}_{\mathbf{z}}(\mathbf{z})+\left[\mathbf{M}-\Delta t^{m} \frac{\partial \mathbf{g}_{\mathbf{z}}(\mathbf{z})}{\partial \mathbf{z}}\right] \Delta \mathbf{z}=\mathbf{0} \tag{5.2}
\end{equation*}
$$

Several sparse matrix-vector products are avoided in this implementation compared with direct solutions of $\mathbf{y}^{m}$. Although the computational savings is small, this implementation is more efficient nonetheless. The two equivalent versions of the NR iterator that can be implemented are given in the following set of equations.

$$
\begin{align*}
& {\left[\mathbf{M}-\Delta t^{m} \frac{\partial \mathbf{g}_{\mathbf{z}}(\mathbf{z})}{\partial \mathbf{z}}\right] \Delta \mathbf{z}=-\mathbf{M} \mathbf{z}+\Delta t^{m} \mathbf{g}_{\mathbf{z}}(\mathbf{z})}  \tag{5.3a}\\
& {\left[\frac{\mathbf{M}}{\Delta t^{m}}-\frac{\partial \mathbf{g}_{\mathbf{z}}(\mathbf{z})}{\partial \mathbf{z}}\right] \Delta \mathbf{z}=-\mathbf{M} \frac{\mathbf{z}}{\Delta t^{m}}+\mathbf{g}_{\mathbf{z}}(\mathbf{z})} \tag{5.3b}
\end{align*}
$$

The linear system above is solved for the iterative update $\mathbf{z}^{k+1}=\mathbf{z}^{k}+\Delta \mathbf{z}$ until $\|\Delta \mathbf{z}\|$ is less than a specified tolerance. The Jacobian $\frac{\partial \mathbf{g}_{\mathbf{z}}(\mathbf{z})}{\partial \mathbf{z}}$, which is approximated as $\left.\frac{\partial \mathbf{g}(\mathbf{y})}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}^{m-1}}$ has a linear fixed part and a nonlinear varying part as shown in the equation below.

$$
\begin{equation*}
\left.\frac{\partial \mathbf{g}(\mathbf{y})}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}^{m-1}}=\mathbf{A}+\left.\frac{\partial \mathbf{f}(\mathbf{y})}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}^{m-1}} \tag{5.4}
\end{equation*}
$$

The linear fixed part is $\mathbf{A}$ and the nonlinear time-varying part of the Jacobian is $\partial \mathbf{f} / \partial \mathbf{y}$. Each of these parts has partitions due to diodes and circuit elements as discussed in the previous chapter.

In a quasi-NR method, the Jacobian is computed only once per time step as opposed to every iteration as in full-NR. A decision to choose one of these approaches is based on the comparison of computational performance of the complete simulation. Full-NR tends to converge faster and takes fewer time steps to complete the simulation than the quasi-NR. A flag is used to choose between these NR implementations. The computational performance for these two cases is compared in Chapter 6 to decide the most suitable implementation for the type of simulation problem considered in this thesis. The backward Euler algorithm is depicted as a flowchart in Figure 5.1.


Fig. 5.1.: Backward Euler algorithm flowchart.

The temporal error in the simulation is controlled by keeping the local truncation error (LTE) at each time step less than a user-specified tolerance. The time step is also chosen such that the truncation error is acceptable within user-specified tolerance limits. The steps are chosen using an automatic step control mechanism discussed
in [65]. Usually, the LTE is computed by constructing an inexpensive higher-order solution $\hat{\mathbf{y}}^{m}$ and approximating the LTE as $\hat{\mathbf{y}}^{m}-\mathbf{y}^{m}$. In each time-step calculation, the initial iterate $\left.\mathbf{z}^{k}\right|_{k=1}$ is assumed to be zero. Consequently, the first gradient computation is the gradient at $t^{m-1}$. The gradient computed during converging iterations of the Newton method corresponds to that at $t^{m}$. With no further computational overhead, these values are used in trapezoidal integration rule to compute $\hat{\mathbf{y}}^{m}$.

$$
\begin{equation*}
\hat{\mathbf{y}}^{m}=\mathbf{y}^{m-1}+\frac{\Delta t^{m}}{2}\left[\mathbf{g}\left(\mathbf{y}^{m-1}\right)+\mathbf{g}\left(\mathbf{y}^{m}\right)\right] \tag{5.5}
\end{equation*}
$$

The approximate LTE and the error norm [61] are given by,

$$
\begin{align*}
\mathrm{LTE} & \approx \hat{\mathbf{y}}^{m}-\mathbf{y}^{m},  \tag{5.6a}\\
\operatorname{LTE} & \approx \frac{\Delta t^{m}}{2}\left[\mathbf{g}\left(\mathbf{y}^{m-1}\right)-\mathbf{g}\left(\mathbf{y}^{m}\right)\right],  \tag{5.6b}\\
\mathrm{ERR} & =\sqrt{\frac{1}{n} \sum_{i=1}^{n}\left(\frac{\operatorname{LTE}_{i}}{\operatorname{Atol}_{i}+\operatorname{Rtol}_{i} \cdot \max \left(\left|y_{i}^{m-1}\right|,\left|y_{i}^{m}\right|\right)}\right)^{2}} \tag{5.6c}
\end{align*}
$$

A weighted root-mean-square norm is used to compute the error using this LTE. User-specified relative and absolute error tolerances (scalar or vector) are the weights in the weighted root-mean-square norm. When the solution at the current time step satisfies the error tolerance norm ERR $\leq 1$, since LTE $\leq$ tol, the next time step is chosen larger or smaller depending on this ERR. The time-step control formula (5.7) employs ERR and exponent $q=(\min (o, \hat{o})+1)$, where $o$ and $\hat{o}$ are the orders of backward Euler and trapezoidal integration methods.

$$
\begin{equation*}
\Delta t^{m+1}=\Delta t^{m} \cdot\left(\frac{1}{\mathrm{ERR}}\right)^{1 / q} \tag{5.7}
\end{equation*}
$$

This factor, which multiplies the last time step size $\left(\Delta t^{m}\right)$, makes the subsequent step sizes to grow or shrink automatically. When the current step solution does not meet the tolerance criterion, the step is rejected and repeated with the smaller step size chosen by the formula. The same formula is used to grow the next step size when the step is accepted. If the NR iteration fails to converge, the step is rejected and step size scaled by 0.1 . In order to improve the likelihood that the next step is accepted,
this factor is scaled by fac and limited between $\mathrm{fac}_{\text {min }}$ and $\mathrm{fac}_{\text {max }}$. In a simulation environment, it is also advisable to limit the step size between $\Delta t_{\min }$ and $\Delta t_{\max }$.

$$
\begin{align*}
\Delta t_{\text {new }} & =\Delta t^{m} \cdot \min \left(\text { fac }_{\max }, \max \left(\mathrm{fac}_{\min }, \text { fac } \cdot(1 / \mathrm{ERR})^{1 / q}\right)\right)  \tag{5.8a}\\
\Delta t^{m+1} & =\min \left(\Delta t_{\max }, \max \left(\Delta t_{\min }, \Delta t_{\text {new }}\right)\right) \tag{5.8b}
\end{align*}
$$

The usual choices of the factors are made as fac $\min =0.1,1.5 \leq \mathrm{fac}_{\max } \leq 5$ and $0.8 \leq$ fac $\leq 0.9$. A time step of $10^{-6} \mathrm{~s}$ is a common choice in the simulation of power electronic circuits. The time scale of the fast dynamics within the semiconductor diode are of the order of $10^{-12} \mathrm{~s}$. Consequently, a choice for the minimum and maximum step sizes are $\Delta t_{\min }=10^{-16} \mathrm{~s}$ and $\Delta t_{\max }=10^{-6} \mathrm{~s}$. A flow chart of the simulation algorithm is shown in Figure 5.2.

The tolerances used in the error-norm calculation, namely the absolute and relative tolerances, can be either vectors or scalars. A brief and clear explanation on the choices of these tolerances is given in [66]. It suggests the use of a vector-valued tolerance when the solution scales are too different. As a rule of thumb, these tolerances are chosen such that when $d$ decimal places of accuracy are sought in $u_{i}^{m}$, then $\operatorname{Rtol}_{i}=10^{-(d+1)}$. A choice of $\mathrm{Atol}_{i}$ is made for which $\left|u_{i}^{m}\right|$ is insignificant. In this thesis, only the charge densities and circuit variables are used in the error norm computation. The electrostatic potential is a function of charge densities within the respective device and in an implicit integration algorithm, this relation is satisfied within the error due to computer's finite precision. Consequently, these can be neglected in the LTE computations.

### 5.2 Parameter Extraction

The system of equations for an example diode resistor inductor network are as derived in the previous chapter. The partitions in (5.3) for the system of equations


Fig. 5.2.: Simulation algorithm flowchart.
in (4.26) are shown explicitly here. On a block level, the iterator is given by the following equation.

$$
\frac{\mathbf{M}}{\Delta t^{m}}-\left(\mathbf{A}+\frac{\partial \mathbf{g}(\mathbf{y})}{\partial \mathbf{y}}\right)=\left[\begin{array}{cc}
\frac{\mathbf{M}_{d}}{\Delta t^{m}}-\left(\mathbf{A}_{d}+\frac{\partial \mathbf{f}_{d}}{\partial \mathbf{y}_{d}}\right) & -\mathbf{A}_{d, \mathrm{ckt}}  \tag{5.9}\\
-\mathbf{A}_{\mathrm{ckt}, d} & \frac{\mathbf{M}_{\mathrm{ckt}}}{\Delta t^{m}}-\left(\mathbf{A}_{\mathrm{ckt}}+\frac{\partial \mathbf{u}_{\mathrm{ckt}}}{\partial \mathbf{y}_{\mathrm{ckt}}}\right)
\end{array}\right]
$$

The top left block matrix is due to the encapsulated diode model. It is block representation of the left-hand side matrix derived in (3.37). The partitions within this block matrix are derived in Appendix A.2. Of the two diagonal block matrices,


Fig. 5.3.: Circuit used for diode characterization.
the right bottom one is due to the circuit. The partition $\mathbf{f}_{\text {ckt,ckt }}$ is zero (matrix) because the $\mathbf{f}_{\text {ckt }}$ vector is due to independent voltage and current sources. If nonlinear circuit elements are present, then this partition will be nonzero. The matrix $\mathbf{A}_{\text {ckt }}$ is already derived and readily available once a given netlist is parsed. The flow chart in Figure 5.2 can be used to implement a coupled device and network simulation.

Physical parameters of a diode are required for a system-level simulation. A commercially available PIN diode is considered here to demonstrate a procedure for approximating these physical parameters. The procedure given in [67] is modified according to the assumptions and operating conditions of the diode. PIN diode S1A (Fairchild) is a relatively slow diode with a recovery time of $1.8 \mu \mathrm{~s}$ when the forward current is 0.5 A . The forward voltage drop of the diode is 1.1 V at a rated current of 1 A and the reverse breakdown voltage is 50 V .

The circuit in Figure 5.3 is used for diode characterization. The voltage source used for quasi-steady state, and, switching measurements are sinusoidal, and, square wave voltage sources respectively. The frequency of the voltage source is chosen to be approximately 100 Hz assuming that the diode dynamics, one of the fastest dynamics in the systems, settles down within a half-period of the source. A commercial function generator (Agilent 33120A) is used to generate both sinusoidal and square wave source voltages.

Internal resistance of function generator together with the wire resistance is used as the resistive load. The inductance of the connecting wire is approximated using the

Table 5.1.: Estimated Circuit parameters.

| Part | Description | value |
| :---: | :--- | :---: |
| $V_{s}$ | voltage source (square or sinu- <br> soidal) | 10 Vpp 100 Hz sine or square |
| $R_{s}$ | Source resistance and wire resis- <br> tance | $51.6 \Omega$ |
| $L_{s}$ | inductance of wire running between <br> the source and diode | 950 nH |
| $C_{a}, C_{c}$ | coupling capacitors | 1 pF |

formulae for parasitic inductances as presented in [68]. The estimated circuit parameters are given in Table.5.1. The formulation developed in this thesis for simulating coupled device-circuit simulation problem is both tested and used to characterize the commercial diode.

The $p^{+}$and $n^{+}$regions in a PIN diode are heavily doped. The doping densities in these heavily doped regions are assumed to be equal as in [67]. This assumption implies that most of the forward voltage $\left(V_{F}\right)$ is dropped across the $p^{+} i$ and $i n^{+}$ junctions equally. There will be some voltage dropped across the intrinsic region of the diode during normal operation. It is assumed that this voltage drop is small compared to the voltages dropped across the junctions. The doping in the $p^{+}$and $n^{+}$ regions can be adjusted to match the forward voltage drop across the diode close to that given in the datasheet. When $N_{\text {dop }}$ is used to represent the doping densities in $p^{+}$and $n^{+}$regions, the built-in voltage of these junctions can be expressed as,

$$
\begin{align*}
V_{F} & \approx 2 \times V_{b i}, \\
V_{b i} & =\frac{k T}{q} \ln \left(\frac{N_{\mathrm{dop}}}{n_{i}}\right) . \tag{5.10}
\end{align*}
$$

The intrinsic carrier density for $\mathrm{Si}, n_{i}$, is constant. A table of $V_{b i}$ versus $N_{d o p}$ is given in Table.5.2 for immediate reference. Assuming a small voltage is dropped
across the intrinsic region, this table is used for ascertaining the doping levels. The doping density of heavily doped region is assumed to be $10^{18} \mathrm{~cm}^{-3}$ for a nominal voltage drop of 1.1 V .

Table 5.2.: Doping density versus forward voltage drop $V_{F}$.

| $N_{\text {dop }}\left(\mathrm{cm}^{-3}\right)$ | $V_{b i}(\mathrm{~V})$ | $V_{F}(\mathrm{~V})$ |
| :---: | :---: | :---: |
| $10^{20}$ | 0.6 | 1.2 |
| $10^{19}$ | 0.5 | 1.0 |
| $10^{18}$ | 0.4816 | 0.9632 |
| $10^{17}$ | 0.4221 | 0.8442 |

The reverse breakdown voltage of the diode under consideration is 50 V . For a lightly doped $i$ region, the doping density is determined from breakdown voltage versus doping density chart in [69], and approximated as $10^{16} \mathrm{~cm}^{-3}$. The doping density in $p^{+}$region is assumed to vary as $\cos \left(\frac{\pi x}{2 X_{j p}}\right)$. A similar variation $\cos \left(\frac{\pi x}{2 X_{j n}}\right)$ is assumed in the $n^{+}$region. The corresponding carrier life time is approximately $\tau_{n}=\tau_{p}=10^{-4} \mathrm{~s}$. The diode current density is approximated as $\frac{q_{e}\left(\bar{p}-n_{i}\right) w_{i}}{\tau}$ [69]. The total current density is expressed using the notations of this thesis as,

$$
\begin{equation*}
J=\frac{I}{a_{c}}=\frac{q_{e}\left(p-n_{i}\right) W_{d}}{\tau_{p}}, \tag{5.11}
\end{equation*}
$$

where $I$ is the current through the diode with cross-sectional area $a_{c}$. Given a doping density for the heavily doped region and $n_{i}$ of Si , the approximate width of the intrinsic layer $\left(W_{d}\right)$ and $a_{c}$ are estimated.

The physical parameters are fine-tuned by performing several quasi-steady state and switching simulations starting with their approximate estimates. The comparison of measured I-V characteristics and diode reverse-recovery with simulation results are used as a guidance to perturb the parameters. The perturbation is done judiciously and iteratively to improve the agreement between measurements and simulations. Only a few iterations of these simulations runs are required to converge to a set of
physical parameters. The iteratively refined parameter set is given in Table 5.3. The comparison of the simulation of $R L$ and diode circuit with the fine-tuned parameters and measurements is presented in the next section.

### 5.3 Diode Resistor Inductor Circuit Example

A $10-\mathrm{V}$ zero-to-peak and $100-\mathrm{Hz}$ voltage is applied to an example network of $R L$ and diode with converged physical parameter estimates. The simulated diode (with Table 5.3) variables for two cycles of the source voltage are shown in Figure 5.4. The diode current $i_{d}$ is nonzero and follows the applied voltage due to positive half-cycle of the voltage source. The inset plot shows the reverse recovery current as the diode becomes reverse-biased. After the diode turns off, the voltage across it is same as negative half-cycle of applied source voltage.

The simulated and measured quasi-steady state I-V characteristics of the diode is shown in the Figure 5.5. The simulated and measured characteristics agree reasonably well. Even with a $100-\mathrm{Hz}$ voltage supply, there is a small reverse recovery as shown in the inset plot of Figure 5.4 and this is the reason for splitting in the I-V characteristics near the cut-in voltage of the diode. The curve is highly nonlinear near the cut-

Table 5.3.: Extracted physical parameters of S1A PIN diode.

| Parameter | Description | Value |
| :---: | :---: | :---: |
| $X_{j p}$ | $p^{+}$region length | $29 \mu \mathrm{~m}$ |
| $X_{j n}$ | $n^{+}$transition length | $2 \mu \mathrm{~m}$ |
| $X_{x}$ | $n^{+}$region length | $29 \mu \mathrm{~m}$ |
| $W_{d}$ | intrinsic layer width | $80 \mu \mathrm{~m}$ |
| $a_{c}$ | cross-sectional area | $110 \times 10^{-4} \mathrm{~cm}^{2}$ |
| $N_{d o p}$ | doping density in $p^{+}$and $n^{+}$regions | $10^{18} \mathrm{~cm}^{-3}$ |
| $N_{D}$ | doping density in intrinsic layer | $10^{16} \mathrm{~cm}^{-3}$ |



Fig. 5.4.: Simulated voltage $v_{d}$ across (top) and current $i_{d}$ (bottom) through the sample diode.
in voltage and becomes almost linear thereafter. The function generator used to characterize the device is not capable of reaching reverse-breakdown voltage levels of the diode sample. Hence, the typical knee and increase in reverse-bias current found in a typical I-V characteristic is not present in the figure shown.

A square wave source is used to characterize a diode using the reverse-recovery response. The duration for which the reverse recovery current is nonzero depends on the forward current, the doping level, width, area and the carrier life times [70]. When fine-tuning the physical parameters, comparing the simulated and measured reverserecovery transients and using the above cues enables one to establish the physical parameters with reasonable certainty.

The simulated and measured reverse-recovery or switching dynamics are compared in Figure 5.6. The oscilloscope is triggered for a negative edge in the source voltage. There is a negative transition at approximately $10 \mu$ s into the simulation.

The measured data is shifted in time so that the negative transition measured and simulated occur at same time for the purposes of comparison. The simulated diode currents agree with the measurements with reasonable accuracy. The magnitude of the simulated diode current is under-estimated, for a brief time for which the diode current is negative. The reason for this could be due to the transmission-line effects of a meter-long coaxial cable used as an interconnecting wire in the measurements. There is a small difference in the current dynamics in the tail of the reverse-recovery current. This is due to the approximation in the doping density and its distribution near the 2 junctions in the PIN diode. The doping density distribution affects the excess charge in the intrinsic region which plays a major role in the reverse recovery dynamics.

The measured and simulated steady-state voltage drop across the diode also agree with reasonable accuracy. There is some mismatch at the corners of the diode voltage after the source voltage goes negative. This is attributed due to the difference in the


Fig. 5.5.: comparison of I-V characteristics of the diode.


Fig. 5.6.: PIN diode switching current (top) and voltage (bottom) dynamics.
current dynamics, specifically the transmission line effect. The simulated voltage stays close to the on-state voltage as long as the negative current is almost constant but the measured reverse current is already varying while the simulated current is
constant. This results in the difference in the voltages near the negative transition. As the diode current increases towards zero, the diode voltage increases in the negative direction. The small differences in the voltage near the tail of the reverse recovery is due to similar difference in the current dynamics. When the voltage has reached close to the steady-state reverse-bias conditions, the diode turns off, whereupon both measured and simulated voltages are the same.

### 5.4 Chapter Summary

An implicit integration algorithm is implemented to simulate the coupled devicecircuit problem. The backward Euler algorithm used in this research is a first-order algorithm. It is possible to use higher order algorithms like backward differentiation formulae due to Gear [71], implicit Runge-Kutta or semi-implicit Rosenbrock methods [61]. A sophisticated step-size controller may be of use to avoid repetitive step rejections. However, when using multi-step methods, it is common to use the lowest order integration after every step-restart due to a step rejection. In a variable-structure variable-order simulation, step-restarts are more likely to happen due to varying structure and so it is prudent to choose single-step higher-order over multi-step algorithms. The diode parameters are approximated by iteratively refining the initial parameter estimates from the datasheet of a device. It is possible to convert this procedure into an optimization problem and to use more sophisticated optimization algorithms to find the physical parameters of the diode [72].

## 6. EXPERIMENTAL VALIDATION

The full system simulation and experimental measurements for an example singlephase full-bridge rectifier are presented in this chapter. Close attention is paid to the diode losses, one of the main motivations of this research. The computational performance of the full system simulation is reported that lays the groundwork for the variable-structure variable-order strategy to be described in the next chapter.

### 6.1 Full System Simulation

The circuit diagram of single-phase diode bridge rectifier is shown in Figure 6.1. A high-power linear operational amplifier fed by a signal generator is used as the source. The source-side resistor and inductor are that of the interconnecting wires between the source and the rectifier input nodes $n 3$ and $n 7$. The rectifier is built using S1A PIN diodes and the load-side circuit elements are commercially available components. The circuit parameters are given in Table 6.1 described using netlist as given in Table 6.2.


Fig. 6.1.: Single-phase diode bridge rectifier.

Table 6.1.: Single-phase diode bridge rectifier circuit parameters.

| Part name | Description | Value |
| :---: | :---: | :---: |
| $v_{s}$ | Source voltage | $9.3 \mathrm{~V}_{\mathrm{pk}-\mathrm{pk}}, 5 \mathrm{kHz}$ |
| $R_{s}$ | Source resistance | $0.1 \Omega$ |
| $L_{s}$ | Source inductance | 500 nH |
| $R_{L}$ | Load resistance | $35.2 \Omega$ |
| $L_{L}$ | Load inductance | $290 \mu \mathrm{H}$ |

Table 6.2.: Single-phase diode bridge rectifier netlist.

```
single-phase diode bridge rectifier
Rs 1 2 0.01
Ls 2 3 5e-7
RL 4 5 35.2
LL 5 6 290e-6
D1 3 4 cct init.m
D2 7 4 cct init.m
D3 6 3 cct init.m
D4 6 7 cct init.m
Vs 1 7 ac sin(0 4.65 5000 0)
. END
```

The incidence graph is generated after adding parasitic capacitors and ground nodes as shown in Figure 6.2. The procedure proposed in Chapter 4 is used to setup the system equations for simulation. The derived DAE system is integrated using the backward Euler integration algorithm. Starting with (5.3), a Newton iterator for the single-phase diode bridge rectifier system equation, as in (4.28), is explicitly expressed. Identifying the block-matrix structure in this iterator is useful for the strategy proposed in the next chapter.

$$
\begin{equation*}
\left[\frac{\mathbf{M}}{\Delta t^{m}}-\left(\mathbf{A}+\frac{\partial \mathbf{g}(\mathbf{y})}{\partial \mathbf{y}}\right)\right] \Delta \mathbf{z}=-\mathbf{M} \frac{\mathbf{z}}{\Delta t^{m}}+\mathbf{g}_{\mathbf{z}}(\mathbf{z}) \tag{6.1}
\end{equation*}
$$



Fig. 6.2.: Incidence graph for the circuit with parasitic capacitors and ground node added.

$$
\left[\begin{array}{ccccc}
\mathbf{D}_{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{1}  \tag{6.2}\\
\mathbf{0} & \mathbf{D}_{2} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{2} \\
\mathbf{0} & \mathbf{0} & \mathbf{D}_{3} & \mathbf{0} & \mathbf{E}_{3} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{4} & \mathbf{E}_{4} \\
\mathbf{F}_{1} & \mathbf{F}_{2} & \mathbf{F}_{3} & \mathbf{F}_{4} & \mathbf{C}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{z}_{d_{1}} \\
\Delta \mathbf{z}_{d_{2}} \\
\Delta \mathbf{z}_{d_{3}} \\
\Delta \mathbf{z}_{d_{4}} \\
\Delta \mathbf{z}_{c k t}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{r}_{d_{1}} \\
\mathbf{r}_{d_{2}} \\
\mathbf{r}_{d_{3}} \\
\mathbf{r}_{d_{4}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]
$$

where

$$
\begin{align*}
& \mathbf{D}_{i}=\mathbf{M}_{d_{i}}-\left(\mathbf{A}_{d_{i}}+\frac{\partial \mathbf{f}_{d_{i}}}{\partial \mathbf{y}_{d_{i}}}\right) \\
& \mathbf{C}_{i}=\mathbf{M}_{\mathrm{ckt}}-\left(\mathbf{A}_{\mathrm{ckt}}+\frac{\partial \mathbf{u}_{\mathrm{ckt}}}{\partial \mathbf{y}_{\mathrm{ckt}}}\right)  \tag{6.3}\\
& \mathbf{E}_{i}=-\mathbf{A}_{\mathrm{ckt}, d_{i}} \\
& \mathbf{F}_{i}=-\mathbf{A}_{d_{i}, \mathrm{ckt}}
\end{align*}
$$

The indices 1 through 4 correspond to the diode number in the circuit. Partitions $\mathbf{D}_{i}$ and $\mathbf{C}_{i}$ are Jacobian block-matrices attributed to diode and circuit, respectively. The matrices that map diode variable $\mathbf{y}_{d_{i}}$ to circuit variable $\mathbf{y}_{\mathrm{ckt}}$ and vice versa are $\mathbf{A}_{\text {ckt }, d_{i}}$ and $\mathbf{A}_{d_{i}, \mathrm{ckt}}$, respectively. Left-hand side $\Delta \mathbf{z}$ vectors are the incremental updates for the diode and circuit partitions. Right-hand side $\mathbf{z}$ is the residual vector with an appropriate diode or circuit partition identifier. The equation above has a block-matrix arrow structure, that is the nonzero block partitions on the diagonal,
last row, and column, form an arrow shape. This algebraic structure is conducive to the domain decomposition method of solving linear systems [73] that is exploited in the variable-structure variable-order strategy. Results presented in this chapter use solutions of the linear system without decomposition as in,

$$
\begin{equation*}
\mathbf{J} \Delta \mathbf{z}=\mathbf{r} . \tag{6.4}
\end{equation*}
$$

It is essential to discuss the measurement equipment capabilities before comparing the simulated and measured data. Yokogawa DL850 scope with 720210 analog voltage input modules are used for measurements. The module has a maximum sample rate of $100 \mathrm{MS} / \mathrm{s}$ which means the data is acquired every $10^{-8} \mathrm{~s}$. The current is measured using Tektronix current probe TCP312 along with its amplifier TCPA300. TCP312 is rated for 30 A dc and its bandwidth is 100 MHz . From the manufacturer's datasheet [74], for small currents that are used in this experiment, full $100-\mathrm{MHz}$ bandwidth is available. The $1 \mathrm{~A} / \mathrm{V}$ output of the current probe amplifier TCPA 300 is connected to the voltage input module of Yokogawa DL850. Thus, voltage and current measurements have a uniform data sampling rate of $100 \mathrm{MS} / \mathrm{s}$.

### 6.1.1 Computational performance

For the purposes of run-time comparison all simulations are executed on the same computer with specifications as listed in Appendix D. The one-dimensional mesh carefully chosen for this performance evaluation has 72 nodes and hence 218 equations in the encapsulated model. Four such device models together with 10-circuit variables results in an 882-dimensional DAE for the single-phase diode bridge rectifier system. Commands tic and toc are used to measure run time of system simulation for four cycles of the source waveform and are 53.10 and 81.8 s for full and quasi-NR methods, respectively. The same simulation study is performed with the MATLAB profiler to pinpoint the computational bottlenecks. The simulation times with and without full-NR enabled are 55.92 and 83.32 s , respectively, which is greater than toc-tic times indicating the small computational burden due to the profiler. A list

Table 6.3.: Profiler output listing computational bottlenecks.

| Computational task | Quasi-NR time (s) | Full-NR time (s) |
| :---: | :---: | :---: |
| Jacobian | 8.61 | 18.58 |
| Right-hand side residual | 24.83 | 12.59 |
| Linear solve in NR iteration | 47.67 | 22.55 |

of computationally expensive tasks, shown in Table 6.3 , are obtained from the profiler results.

It is clear that for the same error tolerance as discussed in the next subsection, the full-NR is faster than quasi-NR method. The former takes 1441 time steps, which is 900 steps fewer than that used by the latter. In both NR implementations, the linear solve stage is the most expensive computationally . The next most expensive stage is the residual computation stage in the quasi-NR method and Jacobian computation stage in the full-NR method, respectively. The run times and numbers reported corroborates the points made earlier in Chapter 5. Hereafter, all simulations use the full-NR-based BE algorithm that is suitable for the problem considered in this thesis.

### 6.1.2 Step size and sample rate

The step size taken by the simulation and measurement interval is plotted in Figure 6.3. The data is sampled at a uniform rate and hence measurement interval plot is a straight line at $\log \left(10^{-8}\right)$. The step size of the simulation is chosen by the step-size controller discussed in Chapter 5. A relative tolerance of $5 \times 10^{-4}$ and absolute tolerance of $5 \times 10^{-6}$ is used in the error estimator. Tightening the tolerances further exhibits no improvement in simulation accuracy and increases the simulation run time. The step size controller aggressively tries to increase the step size and whenever there is a step rejection due to either divergence in the Newton iteration or LTE not meeting the tolerance, the step size is reduced by a factor of ten. This
effect can be seen as dips near $5 \times 10^{-4}$ where the step size is drastically reduced due to repetitive step rejections.


Fig. 6.3.: Measured and simulated step-size comparison.

Step rejections due to exceeding the error tolerance is attributed to the limitation of the controller where the step size changes only by a factor within $\mathrm{fac}_{\text {min }}$ and $\mathrm{fac}_{\text {max }}$. Whenever a step size beyond the limits is required, the size saturates at these limits and results in a step rejection. An implementation of a sophisticated step-size controller that looks at the history of LTE to predict the next step size for probable step acceptance $[61,75]$ is one possible solution. The step-size plot is zoomed in near $5 \times 10^{-4} \mathrm{~s}$ as shown in Figure 6.4. Voltage, current, and power inset plots that follow approximately span this interval. It is seen that near $5 \times 10^{-4} \mathrm{~s}$, the step size is within an order of measurement bandwidth. This observation is kept in mind to judiciously ascertain the reasons for the discrepancies between simulated and measured data.


Fig. 6.4.: Step-size comparison zoomed near $5 \times 10^{-4} \mathrm{~s}$.

### 6.1.3 ac bus waveforms

The ac-voltage $v_{a c}$, measured between nodes n 3 and n 7 as $v_{\mathrm{n} 3}-v_{\mathrm{n} 7}$, and accurrent through $L_{s}$ are plotted in Figure 6.5. The simulated low-frequency voltage and current waveforms are in close agreement with measured ones. There are three small voltage spikes in the inset plot where the measured voltage somewhat differs from simulated. Coincidentally, these are the times where the step-size is close to $10^{-8} \mathrm{~s}$. The measurement bandwidth limitation may be a cause for this difference. The current waveform is qualitatively consistent with the measured current, especially towards the later half in the inset plot. The differences in the peak value of current is more readily seen in the inset plot. This is the duration of the commutation interval where all four diodes conduct and current rises accordingly. The peak value of current spikes are a function of the short-circuit impedance with conducting diodes in the loop. The conductance of the diodes vary due to parameter variation resulting in different current peaks.


Fig. 6.5.: Voltage (top) and current (bottom) waveforms on ac-side.

### 6.1.4 dc bus waveforms

The dc bus voltage $v_{d c}\left(v_{\mathrm{n} 4}-v_{\mathrm{n} 6}\right)$ and the bus or inductor current are plotted in Figure 6.6. The first spike in voltage as shown in the inset plot is simulated accurately. This suggests that there should be appropriate spikes in voltage across all diodes. Albeit uncertain, the small mismatch in the second voltage spike is likely due to measurement bandwidth limitations. There is a small difference in voltage waveforms during the slow transients. It is convenient to recall from Chapter 5 that a similar trend was seen during the parameter-extraction procedure. Hence, the approximation errors in parameter extraction stage is responsible for this voltage difference. The slowly varying current transients are in reasonable agreement even during reverse recovery intervals. The fact that measured currents near $4 \times 10^{-4}$ and $6 \times 10^{-4} \mathrm{~s}$ are slightly different evince the non-identical physical parameters of the diodes.


Fig. 6.6.: Voltage (top) and current (bottom) waveforms on dc-side.

### 6.1.5 Diode 1 variables

The diode D1 voltage and current plots are shown in Figure 6.7. The measured and simulated currents match well and the reverse-recovery current is predicted with reasonable accuracies. The small oscillation near the tail of the reverse-recovery current is related to the spike in voltage across the diode. The simulated and measured diode voltage $v_{d_{1}}$ are in reasonable agreement with the largest discrepancy occurring during the voltage spike as shown in Figure 6.8. It may be tempting to attribute this difference to measurement bandwidth limitation but its not the case. Using the notation $v_{i j}=v_{\mathrm{n} i}-v_{\mathrm{n} j}$, dc bus voltage $v_{d c}=v_{46}=v_{43}+v_{36}=v_{d_{1}}+v_{d_{2}}$. Since there are spikes in $v_{d c}$ and not in $v_{d_{1}}$, they must appear in $v_{d_{2}}$ to satisfy the KVL equation derived above. Consequently, the reason for difference in voltage spikes measured may be due to the physical placement of diode and measurement probes or physical parameter variations. This experimental measurement demonstrates the difficulty in modeling the relevant physical phenomena in the example circuit. The


Fig. 6.7.: Diode voltage (top) and current (bottom) waveforms.
instantaneous power loss in the diode is computed by finding the $v_{d_{1}} i_{d_{1}}$ product as shown in Figure 6.9. The second spike (positive) in power is due to the mismatch in voltage spike. The average power then can be calculated by evaluating the average of this instantaneous power loss curve.

### 6.1.6 Power and energy calculations

The ac and dc-side power is computed using the voltage-current product on the respective sides. The instantaneous powers thus computed are shown in Figure 6.10. The ac-side power agrees reasonably with the simulation. The differences in the dcside power is mainly due to the differences in measured and simulated $v_{d c}$. The energy into and out of the rectifier is the area under one-cycle of the $P_{\mathrm{ac}}$ and $P_{\mathrm{dc}}$ waveforms, respectively. Similarly, energy loss in Diode 1 is the area under one-cycle of the $P_{d 1}$ waveform as in Figure 6.9. The converter loss is estimated as the difference in input and output energy, that is ac-side less dc-side energy.


Fig. 6.8.: Diode voltage (top) and current (bottom) waveforms near $4 \times 10^{-4}$ and $5 \times 10^{-4} \mathrm{~s}$, respectively


Fig. 6.9.: Diode instantaneous power loss waveform.

The per-cycle-energy is computed up to four decimal places, truncated and rounded up, with appropriate metric prefix for measured and simulated waveforms as summarized in Table 6.4. The relative error is computed as the ratio of absolute error to the measured quantity and is expressed in percentage up to two decimal places. Simulated and measured ac-side energies are almost same with a relative error of $0.05 \%$. The error in dc-side side energy is commensurate with the difference in power plots which, in turn, is due to difference in the dc-bus voltage. Simulated and measured Diode 1 loss computation match exactly. The difference in ac and dc-side energies is used in converter loss calculation and hence approximately a $12 \%$ error is found. When the diode bridge rectifier loss is approximated as four times the loss in single diode, a $0 \%$ error is obtained for the loss calculation. The accuracy of energy calculations, mindful of measurement errors, is encouraging.


Fig. 6.10.: Instantaneous power waveforms on ac (top) and dc (bottom) sides.

Table 6.4.: Simulated and measured energy comparisons.

| Quantity | Simulated energy (mJ) | Measured energy (mJ) | Relative error \% |
| :---: | :---: | :---: | :---: |
| ac | 0.1695 | 0.1696 | 0.05 |
| dc | 0.1016 | 0.1088 | 6.65 |
| Diode 1 | 0.0170 | 0.0170 | 0.00 |
| diff(ac,dc) | 0.0679 | 0.0608 | 11.68 |

### 6.2 Chapter Summary

The proposed equation assembly procedure is used to establish a detailed simulation of single-phase diode bridge rectifier. Simulation results obtained using MATLAB predict high-frequency phenomena, energies, and losses in the circuit with reasonable detail and accuracy. The deviations are due to physical parameter variation in diodes (that is, not all diodes are identical) and errors due to measurement. The doping profile that is assumed in this work need not match exactly with that in the actual diode. The same reasoning applies to the physical dimensions of the diode. Albeit thorough knowledge of the physical parameter improves the simulation accuracy, there is still some uncertainty due to the variations among non-identical diodes.

## 7. VARIABLE-STRUCTURE VARIABLE-ORDER SIMULATION PARADIGM

In this chapter, methods of improving the simulation speed are set forth and analyzed. The principal contribution is a variable-structure variable-order (VSVO) simulation paradigm in which inactive devices are assumed to be disconnected from the circuit, thus modifying the structure of the tableau equation and reducing the number of unknown to be solved. The general structure of the tableau equation is preserved thus enabling the use of block Gaussian elimination as a means of improving simulation speed through parallelism. Other factors that improve simulation speed are also considered. In particular, recommendations as to the choice of mesh for PIN diodes are made. Additionally, a comparison is made between direct and indirect (iterative) approaches for solving the linear equations that are a part of the NR or quasi-NR iterations performed at each time step. Another important aspect of VSVO-based simulations is deciding when to disconnect or reconnect devices back into the circuit. A device activity monitor is set forth that accomplishes this task accurately and reliably.

### 7.1 Variable-Structure Variable-Order Paradigm

Although the coupled device-circuit simulation provides an accurate portrayal of the transient performance of power electronic circuits, this approach is seldom used by power electronic circuit designers and analysts due to the prohibitively long times needed to solve the corresponding equations, which are of large dimension and numerically stiff. Observation of the species distributions in a diode during typical transients reveals that their spatial distributions are well behaved and resemble two sigmoid-like functions for each of the $p$ and $n$ densities. It would appear that only a
few dozen nodes in a spatial discretization would be more than adequate to portray these distributions accurately. A good adaptive spatial mesh is capable of placing these handful of nodes at appropriate places. However, due to the large drift-todiffusion current ratio in the depletion region, additional nodes are needed to satisfy the Péclet constraint if a central differencing approach is used. This constraint must be satisfied for purposes of numerical stability. Thus, a relatively large number of nodes are needed inside and in the vicinity of the depletion region even though the species densities do not vary significantly in this region. Moreover, if an adaptive mesh is used the sparse tableau equations have to be completely rebuilt and refactored symbolically and numerically several times for each time step. Thus, adaptive meshing does not provide a significant computational advantage. The results of an initial study, attached in Appendix C, for a single-diode problem with adaptive mesh corroborates this conclusion. The Scharfetter-Gummel differencing method permits a somewhat coarser mesh and, when applied to a reasonably resolved fixed mesh, appears to provide the most efficient solution of an individual device.

For circuits containing multiple devices, however, maintaining a fixed tableau structure does not appear to be the most efficient approach. At any given instant of time, only a subset of the devices are active (e.g. diodes are forward biased). Inactive devices such as reverse-biased diodes do not contribute to the circuit losses that are of paramount concern to power electronic circuit designers. It would appear that a variable-structure and hence variable-order simulation structure, in which inactive devices are removed from the tableau formulation and subsequent solution, would offer significant computational advantage since the dimension would be reduced correspondingly. However, if this is done, it is important to monitor the voltage across the inactive devices to anticipate their becoming active (e.g. diodes becoming forward biased), at which point the structure of the tableau must change to accommodate the soon-to-become-active device. Moreover, by excluding inactive devices, the number of nodes required in the corresponding device models is significantly reduced since the depletion region is small for forward-biased conditions.

Each of these considerations lead to the proposed variable-structure variable-order simulation paradigm. The key attributes are listed below.

- A fixed graded mesh of the semiconductor sufficient to capture the required transients.
- Exploit the knowledge of the fixed block-matrix structure to speed up solution of the linear system of equations.
- Automatically and seamlessly change to and from dimensionally different tableau structures, due to part of the circuit that is active because of the active devices.

Each of these attributes are delineated in the following subsections.

### 7.1.1 Fixed graded mesh

The mesh is chosen with a finer mesh near the two junctions in a PIN diode. During forward-bias and weak-reverse-bias operation of the diode, species densities vary significantly near the junctions. Weak reverse bias is used to refer to negative voltages that are a fraction of the rated or peak operating voltage, which in turn, must be less than the reverse breakdown voltage. In PIN diodes, the $p^{+}$and $n^{+}$ regions are heavily doped and $i$ region is lightly doped. This results in a depletion region that widens more in the $i$ region compared to the $p^{+}$and $n^{+}$regions. The depletion region width varies according to the applied reverse bias voltage. In order to adequately resolve the charge densities under full reverse bias, a fine mesh over a wide region surrounding the two junctions (actually wider in $i$ region) is required when compared to weak-reverse-bias conditions. Consequently, for forward- and weak-reverse-bias conditions, the CCT equations can be solved with a smaller number of nodes compared with a model that is needed to simulate the full range of operation and still satisfy the error tolerance limits.

### 7.1.2 Linear solver

A linear system of equations needs to be solved when an implicit integration algorithm is used. The solution of a large linear system computed in each iteration of the backward Euler algorithm in this research is a computational bottleneck. The two most general approaches of solving a large linear system of equations are iterative methods $[73,76]$ and direct methods [77-81]. Sparse direct methods like UMFPACK [78], PARDISO [80], and SuperLU [81] are used to solve sparse linear systems and consists of several stages. These stages are (1) symbolic factorization, (2) numeric factorization, and (3) forward and backward substitutions. While forward and backward substitution are the fastest and most straight-forward stages, the differences in the aforementioned algorithms lies in the strategies used in the symbolic and numeric factorization stages.

The linear system computed during each iteration and/or step is in general a non-symmetric linear system and is invertible. A non-stationary iterative method such as the Generalized Minimum Residual (GMRES) algorithm is suitable for such systems. This method projects the original problem on to a smaller linear subspace (Krylov space), computes the solution in this reduced space and projects it back to the original space. The solution is improved iteratively until a certain tolerance criteria is satisfied. The linear system should be well conditioned in order for the algorithm to converge in a few iterations. When the eigenvalues of a linear system are spread across a wide range of real and imaginary values, convergence requires a large number of iterations. The eigenvalue spectrum for a typical Newton iterator matrix computed at various operating points in Figure 7.1 are shown in Figure 7.2.

These figures imply that the Newton iterator has a very wide spectrum of eigenvalues. Moreover, the spectrum varies according to the operating conditions and is widest at peak reverse-bias voltage (operating point 3). When iterative solution of such an ill-conditioned linear system is desired, then preconditioning is used to bring these widely spread eigenvalues together and make the iterative method converge


Fig. 7.1.: Four Newton iterator matrices are sampled at the marked instances on current and voltage response of a diode in RLD network.
faster. The best preconditioner is the inverse of the matrix that is to be preconditioned but the computation of such a preconditioner is very expensive. A more efficient way to compute a preconditioner is to use incomplete-LU factors.

A numerical experiment to compare the performance of iterative and direct solvers is devised in MATLAB. The iterator matrices and right-hand sides are computed at certain points of the response of a resistor-inductor-diode example circuit as shown in Figure 7.1. The backslash operator used in MATLAB invokes a sparse direct solver, UMFPACK [78] to be specific. MATLAB also has in-built iterative solvers such as gmres which takes incomplete-LU factors that are computed using ilu, as inputs for preconditioning. In order to reduce the uncertainties in total execution time, the linear system is solved repeatedly so that the cumulative total execution time is several seconds. Moreover, the experiment is repeated several times. The average time from these experiments, when divided by the number of solves, gives the per

Eigen spectrum at different operating points


Fig. 7.2.: Eigen spectrum of respective Newton iterator matrices in clockwise starting at top left for operating points marked left to right as in Figure 7.1.
linear system solution time. The execution time of iterative methods does not include the computation time of the incomplete-LU factors.

The linear system solve times for the example $R L$ and diode example is shown in Table 7.1. This example network at each iteration solves a $224 \times 224$ linear system.

Table 7.1.: Average time for solving $224 \times 224$ linear system (smaller the better).

| Operating point | Direct $\left(10^{-3} \mathrm{~s}\right)$ | Iterative $\left(10^{-3} \mathrm{~s}\right)$ |
| :---: | :---: | :---: |
| $\# 1$ | 0.37 | 46.27 |
| $\# 2$ | 0.37 | 42.11 |
| $\# 3$ | 0.39 | 43.88 |
| $\# 4$ | 0.37 | 44.01 |

It can be seen that sparse direct methods are consistently faster than the iterative methods by two orders of magnitude. It is concluded that for the typical problem size considered in this research, direct methods are preferable. The computational complexity of sparse direct methods in general is super-linear. That is, the computational time of direct methods increases super-linearly as problem size gets bigger and there may exist a break-even point beyond which iterative methods with an effective preconditioner may prove to be efficient. It is possible to validate this claim, starting from $[82,83]$ and references listed therein, and, using a mathematically rigorous analysis and development of iterative method specially tailored for the problem considered. However, such mathematical rigor and development is beyond the scope of this research.

The linear system obtained in a multiple device-circuit simulation has a blockmatrix structure which can be derived from (4.28). Block Gaussian elimination (BGE) can be used to exploit this structure and reduce the linear solve time like in the partitioned finite element method developed in [84]. The short-hand notations created using (5.3) for a single-phase bridge rectifier in Chapter 6 is repeated here for convenience.

$$
\begin{align*}
& {\left[\frac{\mathbf{M}}{\Delta t^{m}}-\left(\mathbf{A}+\frac{\partial \mathbf{g}(\mathbf{y})}{\partial \mathbf{y}}\right)\right] \Delta \mathbf{z}=-\mathbf{M} \frac{\mathbf{z}}{\Delta t^{m}}+\mathbf{g}_{\mathbf{z}}(\mathbf{z}) }  \tag{7.1}\\
& {\left[\begin{array}{ccccc}
\mathbf{D}_{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{1} \\
\mathbf{0} & \mathbf{D}_{2} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{2} \\
\mathbf{0} & \mathbf{0} & \mathbf{D}_{3} & \mathbf{0} & \mathbf{E}_{3} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{4} & \mathbf{E}_{4} \\
\mathbf{F}_{1} & \mathbf{F}_{2} & \mathbf{F}_{3} & \mathbf{F}_{4} & \mathbf{C}
\end{array}\right]\left[\begin{array}{l}
\Delta \mathbf{z}_{d_{1}} \\
\Delta \mathbf{z}_{d_{2}} \\
\Delta \mathbf{z}_{d_{3}} \\
\Delta \mathbf{z}_{d_{4}} \\
\Delta \mathbf{z}_{c k t}
\end{array}\right]=\left[\begin{array}{l}
\mathbf{r}_{d_{1}} \\
\mathbf{r}_{d_{2}} \\
\mathbf{r}_{d_{3}} \\
\mathbf{r}_{d_{4}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right] } \tag{7.2}
\end{align*}
$$

where

$$
\begin{align*}
& \mathbf{D}_{i}=\mathbf{M}_{d_{i}}-\left(\mathbf{A}_{d_{i}}+\frac{\partial \mathbf{f}_{d_{i}}}{\partial \mathbf{y}_{d_{i}}}\right) \\
& \mathbf{C}_{i}=\mathbf{M}_{\mathrm{ckt}}-\left(\mathbf{A}_{\mathrm{ckt}}+\frac{\partial \mathbf{u}_{\mathrm{ckt}}}{\partial \mathbf{y}_{\mathrm{ckt}}}\right)  \tag{7.3}\\
& \mathbf{E}_{i}=-\mathbf{A}_{\mathrm{ckt}, d_{i}} \\
& \mathbf{F}_{i}=-\mathbf{A}_{d_{i}, \mathrm{ckt}}
\end{align*}
$$

```
Algorithm 7.1 Block Gaussian Elimination.
    procedure \(\operatorname{BGE}(\mathbf{J}, \Delta \mathbf{w}, \mathbf{r})\)
        \(\mathbf{E}_{i}^{\prime} \leftarrow \mathbf{D}_{i} \backslash \mathbf{E}_{i}, \mathbf{r}_{d_{i}}^{\prime} \leftarrow \mathbf{D}_{i} \backslash \mathbf{r}_{d_{i}}\) for \(i \in\{1,2,3,4\} \quad \triangleright\) Parallelizable
        \(\mathbf{r}_{\mathrm{ckt}}^{\prime} \leftarrow \mathbf{r}_{\mathrm{ckt}}-\mathbf{F}_{i} \mathbf{r}_{d_{i}}^{\prime}\)
            \(\mathbf{S} \leftarrow \mathbf{C}-\sum_{i=1}^{4} \mathbf{F}_{i} \mathbf{E}_{i}^{\prime}\)
            \(\Delta \mathbf{z}_{\mathrm{ckt}} \leftarrow \mathbf{S} \backslash \mathbf{r}_{\mathrm{ckt}}^{\prime}\)
        \(\Delta \mathbf{z}_{d_{i}} \leftarrow \mathbf{r}_{d_{i}}^{\prime}-\mathbf{E}_{i}^{\prime} \Delta \mathbf{z}_{\mathrm{ckt}}\) for \(i \in\{1,2,3,4\} \quad \triangleright\) Parallelizable
```

The BGE algorithm descried in Saad [73] is applied to this linear system. It is convenient to recall from Chapter 6 that matrix partitions $\mathbf{D}_{i}, \mathbf{C}, \mathbf{E}_{i}$, and $\mathbf{F}_{i}$ are attributed to diodes, circuit, circuit-diode interface, and diode-circuit interface, respectively. The notations used originally in [73] is modified with the above notations and is presented in Algorithm 7.1. The possible computations that can be performed in parallel are the Steps 2 and 6 in Algorithm 7.1. Computing the right-hand side of the reduced system and Schur complement as in the Steps 3 through 5 uses accumulation. Hence, these steps cannot be parallelized or else they may result in data race conditions. Data race in parallel computations usually results in numbers that are not reliable and hence such sections are better left in their sequential form. Two of the popular parallel programming paradigms [85] are Message Passing Interface (MPI) [86] and OpenMP [87]. MPI is a specification for message passing libraries that moves data from one address space of a process to the other in a coordinated way, possibly over a computer network. MPI is primarily a parallel programming paradigm on distributed memory architecture since data can be distributed among different processes running on different hardware with physically separated memory. OpenMP is an Application Program Interface (API) that is capable of spawning multiple threads in a shared memory architecture, like in a desktop computer with multiple CPU cores that are common these days.

The computational performance of a parallel programming paradigm relies on a number of factors such as data locality, cost of data transfer and its volume, spawn-
ing thread/process, etc. MPI tries to spawn multiple processes on multiple CPUs. A process in turn can spawn a number of threads, so spawning a thread is computationally cheaper than a process. Most of the data fits in the local memory for the typical problem sizes of this research. This suggests OpenMP is more suitable, for the problem considered in this research.

Either (7.2) or (7.1) can be solved to compute the next step solution $\mathbf{y}^{m}$ in a typical system simulation. The solution of (7.2) is referred to as full solve and that of Algorithm 7.1 as BGE solve. Full and BGE solve use UMFPACK as its kernel linear solver. These solution methodologies are implemented in C programming language by calling appropriate functions from UMFPACK and OpenMP libraries. The computational performance of each strategy executed on a computer with specifications as in Appendix D is compared in Table 7.2. The overall dimension of the full system for a full bridge rectifier simulation is 882 . There are 10 circuit variables and each diode consists of 218 equations. Each $\mathbf{D}_{i}$ is a $218 \times 218$ matrix and $\mathbf{r}_{d i}$ a 218-dimensional vector. The last two columns gives the measured and calculated speed-up. Amdahl's law [88] is used to calculate the speed-up by parallelizing the BGE algorithm Steps 2 and 6 . The speed-up is calculated using the time taken by the serial $\left(t_{s}\right)$ and parallelizable $\left(t_{p}\right)$ part of the BGE algorithm running on a single thread. The times obtained for the sample 10000 solves are $t_{s}=0.49 \mathrm{~s}, t_{p}=30.57 \mathrm{~s}$. Their sum is not exactly equal to the total run time ( 30.9 s ) because these times are obtained from a consecutive run. The sum of times taken by BGE algorithm Step 2 and 6 is $t_{p}$ of which Step 2 alone takes 30.54 s . This reveals that the computation in Step 2 of BGE is the costliest. The time taken by the Steps 3 through 5 is $t_{s}$. Informations $t_{s}$ and $t_{p}$ are used to extrapolate speed-up using $N_{\text {thrds }}$ threads [89],

$$
\begin{equation*}
\text { Speed-up }=\frac{t_{s}+t_{p}}{t_{s}+\frac{t_{p}}{N_{\mathrm{thrds}}}} \tag{7.4}
\end{equation*}
$$

Even though speed-up calculations are optimistic, they are close to the measured ones. This gives confidence that this metric can be used to calculate the speed-up in the parallelization of the problems considered in this research. The measured speed-

Table 7.2.: Performance comparison of 10000 solves of a linear system of 882 equations.

|  |  |  | Speed-up |  |
| :---: | :---: | :---: | :---: | :---: |
| Threads | Full solve time (s) | BGE solve time (s) | measured | calculated |
| 1 | 33.5 | 30.9 | 1.1 | 1.0 |
| 2 | 33.2 | 16.0 | 2.1 | 2.12 |
| 3 | 33.2 | 16.0 | 2.1 | 2.12 |
| 4 | 33.4 | 8.8 | 3.8 | 4.12 |

ups are discussed starting with the single-thread case. There is a speed improvement of $10 \%$ over the full solve for this case. As mentioned earlier, sparse direct solvers have a super-linear computational complexity and hence, when a sequence of smaller linear systems are solved and solutions are assembled, there is a possibility to get some additional computational gain.

The results for other cases are explained qualitatively with the help of Figure 7.3. Main thread signifies the entry and exit point of the BGE algorithm. The width (not to scale) of other shaded parts shows the time taken by the corresponding algorithm states. It is obvious that when only one thread is used, the BGE solve time is the longest corresponding to the figure length. The main thread forks into multiple threads as they are spawned and one of them is the main thread itself executing the parallel task. The OS kernel decides which thread (inclusive of main thread) executes the parallel task. For the example in the Figure 7.3, either of the top or bottom fork in the 2 thread case could be the main thread. Only one thread computes the right-hand side of the reduced system and Schur complement to prevent data racing and all other threads do no work during this interval. Once the Schur system is solved by the active thread, the other solution variables are computed in parallel. It can be seen that when two threads are used, BGE solve is 2 X faster.

In the three-thread scenario, while the active thread is working on the last parallel task, the other two threads wait on this active thread. This is the reason that the 3- and 2-thread scenarios both have a 2 X speed-up and essentially the computer resource (thread) is wasted here. When the number of threads match the number of diode partitions, the speed-up is maximum, and for this example it is close to 4X as expected. There will not be a speedup with increase in the number of threads beyond this point as excessively spawned threads will do no work and wait on the other active threads. The figure gives a cue that when the number of parallel tasks are unevenly matched with threads, the percentage of 'No work' portion shown in dots increases which brings down the speed-up. The problem division has only one level in the example explained thus far, that is one big problem is divided into 5 tasks of which 4 can be executed in parallel. It is possible to increase the number of levels, not to arbitrarily many, and follow a strategy like in [84]. Dividing the problem


Fig. 7.3.: Qualitative figure to explain BGE solver's performance.
into arbitrary levels and hence a large number of parallel tasks will bring down the speed-up because of the volume of data that needs to be exchanged between these tasks.

### 7.1.3 Structurally varying tableau

The circuit variable vector $\mathbf{u}_{\mathrm{ckt}}$ consists of all node voltages followed by the inductor and source currents. Step 6 of Algorithm 7.1 computes the change in circuit variables $\left[\begin{array}{llll}\Delta \mathbf{v} & \Delta \mathbf{i}_{L} & \Delta \mathbf{i}_{V}\end{array}\right]^{T}$. The corresponding right-hand side residual vector $\mathbf{r}_{\mathrm{ckt}}$ then has three partitions. The first partition corresponds to the mismatch in KCL equations, the second to the mismatch in inductor voltage equations and finally the third with the source voltage equations. This implies that the residual vector can be partitioned as $\left[\Delta \mathbf{i} \Delta \mathbf{v}_{L} \Delta \mathbf{v}_{s}\right]^{T}$. These two vectors are related to each other via the Schur complement S. The Schur complement is partitioned accordingly as,

$$
\left[\begin{array}{ccc}
\mathbf{Y} & \mathbf{U}_{1} & \mathbf{U}_{2}  \tag{7.5}\\
\mathbf{W}_{1} & \mathbf{Z} & \mathbf{0} \\
\mathbf{W}_{2} & \mathbf{0} & \mathbf{0}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{v} \\
\Delta \mathbf{i}_{L} \\
\Delta \mathbf{i}_{V}
\end{array}\right]=\left[\begin{array}{c}
\Delta \mathbf{i} \\
\Delta \mathbf{v}_{L} \\
\Delta \mathbf{v}_{s}
\end{array}\right]
$$

This indicates that when the Schur complement $\mathbf{S}$ is partitioned according to variables mentioned, the top-left partition in it is an admittance matrix. The $\mathbf{U}$ and $\mathbf{W}$ partitions correspond to current and voltage-gains respectively. The $\mathbf{F} \mathbf{E}^{\prime}$ product or $\sum \mathbf{F}_{i} \mathbf{E}_{i}^{\prime}$ in Step 4 of the BGE represents the contribution to admittance matrix $\mathbf{Y}$ partition of the Schur complement by the diode connected between the respective nodes.

This key insight allows the addition and removal of a particular diode partition easily to a simple block-column and row deletion in (7.1). Apart from removal of appropriate $\mathbf{D}_{i}$ matrix, deletion of $\mathbf{F}_{i}$ is equivalent to physically not applying voltages to the diode and that of $\mathbf{E}_{i}$ is equivalent to diode not injecting currents into the circuit. Hence the removal of block-column and row is equivalent to physically disconnecting the particular diode from the circuit that is simulated. There will be only $\mathbf{D}, \mathbf{E}$ and
$\mathbf{F}$ in the Newton iterator for single resistor-inductor-diode network example. The block-matrix structures of the Newton iterator for the on and off-state of the diode is given below.

$$
\begin{align*}
{\left[\begin{array}{ll}
\mathbf{D} & \mathbf{E} \\
\mathbf{F} & \mathbf{C}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{z}_{d} \\
\Delta \mathbf{z}_{\mathrm{ckt}}
\end{array}\right] } & =\left[\begin{array}{c}
\mathbf{r}_{d} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]  \tag{7.6}\\
\mathbf{C} \Delta \mathbf{z}_{\mathrm{ckt}} & =\mathbf{r}_{\mathrm{ckt}} \tag{7.7}
\end{align*}
$$

When the diode is not conducting, the admittance or conductance between the nodes to which it is connected is near-zero or zero. BGE applied to the system (7.6) results in a Schur complement that is simply $\mathbf{C}-\mathbf{F E}^{\prime}$. The physical observation of zero/near-zero admittance is mathematically equivalent to stating that $\mathbf{F E}^{\prime} \approx \mathbf{0}$. It is computationally advantageous to solve the small-dimensioned system in (7.7) for the diode-off condition than to solve the entire system using BGE. As the simulation of this example network proceeds, switching the Newton iterator structure between (7.6) and (7.7) seamlessly results in a structurally varying tableau. A metric or measurement needs to be devised so that the decision to switch between these 2 structures can be made.

When the diode voltage crosses zero going positive, the diode starts to conduct. Switching the structure of the tableau from (7.7) to (7.6) at this time is favorable. The reason being precomputing the diode states for zero applied voltage and using it is easier than computing the diode states for the diode cut-in voltage and initializing the initial conditions of the DAE consistently [90]. A consistent set of initial conditions is very important in DAE based simulations for the results to be meaningful.

The metric used to decide when to switch from (7.6) to (7.7) is more involved. A metric based on the measurement of difference of anode and cathode terminal currents used in the preliminary research is updated to a metric that measures spatial-average of the time rate of change of the carrier densities within the diode. At the end of each time-step, approximate $\partial \mathbf{p} / \partial t$ and $\partial \mathbf{n} / \partial t$ are computable. Spatial-numerical integration of the rate $\partial \mathbf{p} / \partial t$ over the body of the diode as in (7.8) is equivalent to finding the difference in the terminal current densities due to holes (Divergence
theorem). For small device currents, that in effect is the device current itself because the hole current density is zero at the anode terminal and is the device current density at the cathode terminal towards the end of the reverse recovery.

$$
\begin{equation*}
D_{\mathrm{mon}}=\int_{\Omega} \frac{\partial \mathbf{p}}{\partial t} d \Omega=a_{c} \int_{X} \mathbf{f}_{\mathbf{p}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) d x \tag{7.8}
\end{equation*}
$$

Terminal current densities and hence currents are found using the spatial derivatives of carrier densities and electric potential. The current-based metric then will involve difference between these nearly equal currents and may be polluted with finite precision error. Integration is often preferred over differentiation in computations owing to its numerical stability. The minority carriers in the $i$ and $n$ regions, which constitutes the bulk of the device, are holes and it is sufficient to integrate the rate of change of holes within the body of the diode to monitor the device activity. When $D_{\text {mon }}$ is lesser than a tolerance of $5 \%$, the diode is deemed to be off. This tolerance that relies on scaled hole density may seem large but when scaled by appropriate parameters results in $65 \mu \mathrm{~A}$. That is, an equivalent current-based metric tracks the device-terminal current and triggers whenever it crosses $65 \mu \mathrm{~A}$. Other circuit conditions such as applied diode voltage and whether it is increasing or decreasing needs to be tracked to ascertain the validity of trigger which is unlikely to happen with the proposed activity monitor.

As the diode starts to conduct, the $i$-region is gradually flooded with holes and the activity monitor will be much larger than the tolerance (0.05). Hence, the monitor will not trigger while the diode is turning on. The monitor will not trigger when the applied diode voltage goes negative briefly due to voltage spikes in the circuit while the diode is conducting. This due to the fact that the excess holes in the $i$ region needs to removed before the diode can get back to its reverse-blocking or off state. The activity monitor value is well above the tolerance when the diode is on and conducting. When the diode goes through reverse-recovery, the excess holes from the $i$-region are removed to bring the diode to off state. The $\partial \mathbf{p} / \partial t$ rate is nonzero close to the ends of the depletion region formed at the $p^{+} i$ junction towards the end
of the reverse-recovery. The spatial-integration of this rate approaches zero and the activity monitor triggers promptly as the diode turns off.

The performance of this metric for the example $R L$ and diode ( $R L D$ ) network is shown in Figure 7.4. The frequency of the source voltage is 100 Hz and the source inductance is small. Due to these reasons the diode has a very small reverse recovery and it is on for $50 \%$ of the time and off for the rest. The diode activity metric based on the carrier density rates is shown in the lower subplot of Figure 7.4. The diode off signal is triggered only after reverse recovery subsides. This results in a crisp diode onoff signal. The diode partition does not have to be solved, and, the associated residual and Jacobian calculations do not have to be performed during the $50 \%$ diode off duration. This implies that qualitatively variable-structure variable-order simulation will result in a little over 2X speed-up. Extending this idea to a single-phase bridge rectifier shown in Figure 6.1 is straightforward. It is useful to consider an example operating instance where diodes D1 and D4 are not conducting while D2 and D3


Fig. 7.4.: Diode activity detector output for the example $R L D$ network.
are conducting. For this case, $\mathbf{F}_{1} \mathbf{E}_{1}^{\prime} \approx \mathbf{F}_{4} \mathbf{E}_{4}^{\prime} \approx \mathbf{0}$. It is computationally wasteful to evaluate this zero or near-zero matrix in the BGE approach. The deletion of columns and rows with $\left\{\mathbf{D}_{1}, \mathbf{D}_{4}\right\}$ in (7.1) for the case considered, shown in the equation below, computes only the relevant nonzero variables.

$$
\left[\begin{array}{ccc}
\mathbf{D}_{2} & \mathbf{0} & \mathbf{E}_{2}  \tag{7.9}\\
\mathbf{0} & \mathbf{D}_{3} & \mathbf{E}_{3} \\
\mathbf{F}_{2} & \mathbf{F}_{3} & \mathbf{C}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{z}_{\mathrm{d}_{2}} \\
\Delta \mathbf{z}_{\mathrm{d}_{3}} \\
\Delta \mathbf{z}_{\mathrm{ckt}}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{r}_{\mathrm{d}_{2}} \\
\mathbf{r}_{\mathrm{d}_{3}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]
$$

As the current commutes from one diode pair to the other, for a brief period all diodes are on. Consequently, the three possible matrix structures that needs to be solved in a bridge rectifier simulation are (7.1), (7.9) when Diodes 2, 3 are on, and replacing indices $\{2,3\}$ with $\{1,4\}$ in (7.9) when Diodes 1,4 are on. This is the basis for the variable-structure variable-order (VSVO) simulation paradigm. The paradigm is variable-structure due to the change in structure of the tableau, and variable-order because of the change in the number of state variables as the simulation progresses.

### 7.2 Single-Phase Bridge Rectifier Simulation

All diode and circuit variables at each time step in a bridge rectifier simulation for four cycles are saved. The diode carrier densities are then used to evaluate the gradient functions which is approximately the rate of change of carrier densities. This rate is distributed over the body of the diode and is integrated over the diode length as in (7.8). When $D_{\text {mon }}$ is less than the $5 \%$ tolerance ( $65 \times 10^{-6} \mathrm{~A}$ ), the corresponding diode state is made zero. As soon as the voltage across the diode becomes positive the corresponding diode state is set to one. The computed diode states of all four diodes in the bridge rectifier is shown in Figure 7.5.

It can be seen that the diode pair D1 and D4 or D2 and D3 conduct most of the time as expected. All four diodes are on for a brief duration between the change of states of these pairs. The computed diode switch state is crisp without chattering indicating that the metric is immune to false triggers. Diode 1 current and Diode 1


Fig. 7.5.: Diode switch on-off states versus time.
state scaled by 0.1 are plotted together for illustration in Figure 7.6. At every even integer multiple of $10^{-4} \mathrm{~s}$, the diode briefly turns on and off. Hence the diode current increases, reaches its peak, decreases, and, reaches its negative peak before turning fully on. During this interval, the current crosses $65 \mu \mathrm{~A}$ two times for which a current metric may trigger falsely but not $D_{\text {mon }}$. At each odd integer multiple of $10^{-4} \mathrm{~s}$, the diode goes through its reverse-recovery. This interval also has two instances where a current metric may trigger but not $D_{\text {mon }}$. Diode off state is calculated accurately using $D_{\text {mon }}$ in an elegant, straightforward way.

This metric is readily implemented in the simulation to determine when a diode is to be connected or disconnected. From Figure 7.5, it is obvious that only for a fraction of the entire simulation time the full system needs to be simulated. At other times, its enough to include the conducting diode pair equations in the tableau. The diode states computed as shown in Figure 7.5 is further processed to estimate the percentage of the computation time with 2 and 4 diodes on. Diode pairs D1 and


Fig. 7.6.: Diode 1 current and its on-off state versus time.

D4 or D2 and D3 are on and needs to be in the tableau equations for $47.72 \%$ of the computation time. All four diode equations needs to be in the tableau for $52.23 \%$ of the computation time. Depending on the source and load-side inductance, and the current levels, the ratio of these percentages may change. VSVO simulation has an overhead of keeping track of the diode states using (7.8) and varying the dimension of the Newton iterator. The different tableau structures that needs to be solved in Newton iteration of single-phase diode bridge rectifier simulation is given below.

$$
\begin{align*}
& {\left[\begin{array}{ccc}
\mathbf{D}_{2} & \mathbf{0} & \mathbf{E}_{2} \\
\mathbf{0} & \mathbf{D}_{3} & \mathbf{E}_{3} \\
\mathbf{F}_{2} & \mathbf{F}_{3} & \mathbf{C}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{z}_{d_{2}} \\
\Delta \mathbf{z}_{d_{3}} \\
\Delta \mathbf{z}_{\mathrm{ckt}}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{r}_{d_{2}} \\
\mathbf{r}_{d_{3}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]}  \tag{7.10}\\
& {\left[\begin{array}{ccc}
\mathbf{D}_{1} & \mathbf{0} & \mathbf{E}_{1} \\
\mathbf{0} & \mathbf{D}_{4} & \mathbf{E}_{4} \\
\mathbf{F}_{1} & \mathbf{F}_{4} & \mathbf{C}
\end{array}\right]\left[\begin{array}{l}
\Delta \mathbf{z}_{d_{1}} \\
\Delta \mathbf{z}_{d_{4}} \\
\Delta \mathbf{z}_{\mathrm{ckt}}
\end{array}\right]=\left[\begin{array}{l}
\mathbf{r}_{d_{1}} \\
\mathbf{r}_{d_{4}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]} \tag{7.11}
\end{align*}
$$

$$
\left[\begin{array}{ccccc}
\mathbf{D}_{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{1}  \tag{7.12}\\
\mathbf{0} & \mathbf{D}_{2} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{2} \\
\mathbf{0} & \mathbf{0} & \mathbf{D}_{3} & \mathbf{0} & \mathbf{E}_{3} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{4} & \mathbf{E}_{4} \\
\mathbf{F}_{1} & \mathbf{F}_{2} & \mathbf{F}_{3} & \mathbf{F}_{4} & \mathbf{C}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{z}_{d_{1}} \\
\Delta \mathbf{z}_{d_{2}} \\
\Delta \mathbf{z}_{d_{3}} \\
\Delta \mathbf{z}_{d_{4}} \\
\Delta \mathbf{z}_{\mathrm{ckt}}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{r}_{d_{1}} \\
\mathbf{r}_{d_{2}} \\
\mathbf{r}_{d_{3}} \\
\mathbf{r}_{d_{4}} \\
\mathbf{r}_{\mathrm{ckt}}
\end{array}\right]
$$

The varying tableau structure allows changes in the circuit topology; however, the circuit partition $\mathbf{C}$ does not change. This is not a coincidence but is only possible due to the system assembly procedure set forth in Chapter 4. The assembly procedure distributes certain portion of circuit equations, related to diode to be specific, across the tableau in the form of $\mathbf{E}_{i}$ and $\mathbf{F}_{i}$. This is a very desirable feature that ensues a varying tableau structure with fixed circuit partition. Such a flexibility is not available if the circuit is modeled using a state-space approach.

### 7.3 VSVO Simulation Results

The results of a VSVO simulation with a full linear system solve is presented and discussed in this section. The VSVO simulation results are compared with reference simulation results labeled as 'ref' in the plots and are phase-shifted for clarity. These are the same results that were presented in Chapter 6. A carefully chosen coarse mesh with 57 nodes is used to discretize one-dimensional diode model. Four such diode models together with 10 circuit variables results in a $702 \times 702$ system. Since only two sets of diode model need to be simulated when two diodes are on, the system dimension then becomes $352 \times 352$. The toc-tic timing is 37.42 s compared to 53.10 s of the full system simulation which gives a 1.4 X speed-up. VSVO diode power calculation is identical to the reference simulation waveform as shown in Figure 7.7. Switch-out (swo) refers to the action of removing diode equations from the simulation at time $t_{\text {swo }}$. Similarly, switch-in (swi) is used to refer to the action of inserting diode model equations back in to the system at time $t_{\text {swi }}$. There are no discernible glitches in the power waveform due to swo or swi.

The voltage and current waveforms of Diode 1 are shown in Figure 7.8 where again no visible glitches appear due to swo or swi. These waveforms are zoomed near the negative half-cycle of the source voltage as in Figure 7.9. The current waveform is zoomed on the $y$-axis in the range $\pm 0.1 \mathrm{~mA}$ to illustrate the change in $i_{d_{1}}$ due to swo and swi near $5.0 \times 10^{-4}$ and $6.0 \times 10^{-4} \mathrm{~s}$, respectively. As the current crosses $-6.5 \times 10^{-5} \mathrm{~A}$, the diode model is switched out and the diode current is forced to become zero near $5.0 \times 10^{-4} \mathrm{~s}$. The diode voltage crosses zero with a positive slope near $6.0 \times 10^{-4} \mathrm{~s}$ when the diode equations are inserted back in to the simulation. It can be seen that the diode current settles within two time-step computations after initializing the diode with zero-voltage-biased solution because it is close to the consistent initial condition. It is evident that $D_{\text {mon }}$ performs very well in monitoring diode activity and inserting diode equations with zero-voltage-biased solution at $t_{\text {swi }}$ works satisfactorily.

Plots of ac and dc variables are included for the sake of completeness. The differences in ac quantities are really small. The dc quantities exhibit some difference


Fig. 7.7.: Instantaneous power waveforms of Diode 1 using VSVO.


Fig. 7.8.: Voltage (top) and current (bottom) waveforms of Diode 1 comparison.


Fig. 7.9.: Zoomed voltage and current waveforms of Diode 1 using VSVO.


Fig. 7.10.: Voltage (top) and current (bottom) waveforms on ac-side.


Fig. 7.11.: Voltage (top) and current (bottom) waveforms on dc-side.

Table 7.3.: Simulated and measured energy comparisons.

| Quantity | VSVO energy (mJ) | Measured energy (mJ) | Relative error \% |
| :---: | :---: | :---: | :---: |
| ac | 0.1690 | 0.1696 | 0.35 |
| dc | 0.1012 | 0.1088 | 6.98 |
| Diode 1 | 0.0170 | 0.0170 | 0.00 |
| diff(ac,dc) | 0.0678 | 0.0607 | 11.70 |

in voltage spikes. The dc bus voltage is the sum of $v_{d_{1}}$ and $v_{d_{2}}$, which pinpoints the source for this difference to that in diode voltages and the mesh used for the two simulations are different. The energy calculations done in Table 6.4 is repeated in Table 7.3. The relative error is computed as the ratio of absolute error and measured quantity expressed as percentage. The errors are not too far from Table 6.4 indicating that by using a coarser mesh for the diode discretization and including the diode model on an as-needed basis does not affect the simulation and energy calculations significantly.

### 7.4 Speed-up calculations

MATLAB's parallel computing toolbox can be used to parallelize certain sections of the VSVO simulation program. The inbuilt parfor and/or spmd commands work on the concept of shared-memory programming. The main MATLAB client instantiates so-called MATLAB workers and a large computational task referred as job is split into tasks and sent to these workers [91]. The solution is reassembled by the client after the workers complete their tasks. A computer with multi-core central processing units can use parallel computing toolbox on so-called local cluster. The overhead due to instantiation of MATLAB client and workers, and communication between them overwhelms the speed-up achievable for the problem size considered in this research. A low-level $\mathrm{C} / \mathrm{C}++$ program with multi-thread capability is rec-

Table 7.4.: VSVO profile results.

| Computation |  | VSVO Time (s) |  |
| :---: | :---: | :---: | :---: |
|  | Description | on | off |
|  | Compute Jacobian once per step | 15.61 | 18.62 |
| Residual | Compute the right-hand side residual | 9.46 | 10.92 |
| Linear solve | Solve the linear system inside Newton iteration | 10.61 | 13.30 |
| Diode monitor | Monitor diode activity | 1.38 | 1.50 |

ommended for realizing the simulation speed-up for the problem considered in this thesis. A theoretical calculation of achievable speed-up is deduced from the profile results of MATLAB-based VSVO implementation.

It is convenient to recall the full system simulation performance analysis from Chapter 6. The toc-tic timing is 53.10 s and the profiler timing is 55.92 s . Jacobian, residual computation, and, linear system solve time are all parallelizable. The serial portion of the linear solve stage is small and neglected in the calculation. A coarser diode mesh with 57 nodes make the dimension of encapsulated diode model to be 171. The dimension of the full system becomes $702 \times 702$. The toc-tic timing is 37.42 s and the profiler timing is 40.22 s . The dimension of the linear system solved is $702 \times 702$ and $352 \times 352$, when four and two diodes conduct, respectively. The percentage of these durations with respect to computation time is found to be $52 \%$ and $48 \%$, respectively.

A speed-up calculation formula is derived based on VSVO simulation profile results in MATLAB as shown in Table 7.4. Solution of the full set of linear equations is replaced with BGE for a speed-up and it has a sequential part that cannot be parallelized. The sequential portion of BGE solve requires only a few milli-seconds on average for the circuit considered and is neglected. The time taken for the computational tasks are given for two cases where VSVO is enabled (on) and disabled (off). VSVO-enabled simulation uses a structurally-varying tableau whereas VSVO-
disabled simulation uses a fixed tableau since all diodes are always included. It can be seen that the times taken by each computation for VSVO-on case is on average $83 \%$ that of the VSVO-off case or takes $17 \%$ lesser time.

The reference-simulation-profile time is denoted as $t_{\text {ref }}=55.92 \mathrm{~s}$. In VSVO-off scenario, the serial portion is takes 2.13 s and parallelizable part takes 45.42 s . This information together with the fact that VSVO-on computations takes $83 \%$ of the time taken by VSVO-off case is used in (7.4) to calculate speed-up,

$$
\begin{equation*}
\text { Speed-up }=\frac{t_{\mathrm{ref}}}{t_{s}+\frac{0.83 t_{p}}{N_{\mathrm{thrds}}}}, \tag{7.13}
\end{equation*}
$$

where, $t_{s}$ and $t_{p}$ are the respective times taken by serial and parallel portion of the code. This calculation is applicable when 1,2 or 4 threads are used. When the number of diodes and threads are not evenly matched, computational resources are wasted and speed-up for 2 and 3 thread case will be same. Using VSVO strategy, two diodes are on for $52 \%$ and four diodes are on for $48 \%$ of the computation time. Figure 7.12, an adaptation of Figure 7.3, is used to depict the speed-up achieved due to VSVO strategy qualitatively. It can be seen that one-step calculations with 2 diodes on will be twice as fast than when 4 diodes are on. This implies that the speed-up for the 3 -thread scenario is the weighted average of 2 - and 4 -thread cases with weights $52 \%$ and $48 \%$, respectively. As a sanity check for the speed-up calculation (7.13), the speed-up for single-thread case is computed and compared.

$$
\begin{equation*}
\frac{55.92}{40.22}=1.39 \approx \frac{55.92}{2.13+\frac{0.83 \times 45.42}{1}}=1.40 \tag{7.14}
\end{equation*}
$$

It is apparent that the measured and calculated speed-ups are reasonably close. This equation is used to estimate the speed-up for different number of threads as in Table 7.5 with weighted-averaging for three-thread case. The achievable speed with 4 parallel computing threads is $484 \%$.

Table 7.5.: Estimated speed-up.

| Thread | Speed-up |
| :---: | :---: |
| One | 1.40 |
| Two | 2.66 |
| Three | 3.71 |
| Four | 4.84 |

### 7.5 Chapter Summary

The numerical solution of PDEs that model drift-diffusion is still an active area of research [92]. Spatial adaptivity for such a problem is in its infancy. The coupled charge transport (CCT) equation involves coupled drift-diffusion equations along with a Poisson's equation, that comprise a system of PDEs. Until a reliable, robust, spatial adaptive strategy is developed, it is argued that a carefully chosen fixed mesh is preferred for the CCT equation. The matrix structure of the system tableau is exploited to speed up the linear solve stage using sparse direct solvers. VSVO strategy utilizes less computational effort for a given system compared with full system simulation. The activity metric needs a revision when other devices like IGBT, FET and BJT are


Fig. 7.12.: Thread utilization with 3 threads for 2 and 4 -diodes on cases.
considered. A VSVO simulation using 4 threads is capable of simulating the example single-phase diode bridge rectifier 4.84 times faster than the conventional simulation. This speed-up is further expected to increase with the dimension of the device model and number of devices in the simulated circuit.

## 8. SUMMARY, CONCLUSIONS, AND FUTURE RESEARCH

The primary contributions, main conclusions, and suggestions for future research are summarized herein. First, an extensive survey of coupled device-circuit simulation approaches was provided in Chapter 1. Next, the relevant physical phenomena associated with power semiconductor diodes were reviewed in Chapter 2 culminating in the well-established coupled drift, diffusion, and continuity (CDDC) equations.

Practicable discretization procedures for numerically solving the CDDC equations were identified and analyzed in Chapter 3. Therein, an encapsulated device model was set forth setting the stage for its inclusion into potentially complex power electronic circuits. The hole and electron densities at selected nodes represent independent state variables whereas the corresponding electric potentials represent dependent states. The inputs to the encapsulated device model are the applied voltages at the anode and cathode, which are used as Dirichlet boundary conditions in the solution of the CCDC equations, while the outputs are the net anode and cathode currents. Although a one-dimensional spatial discretization was used, the model and discretization approach is readily extended to two dimensions, if needed. Possible extensions include the incorporation of other potentially relevant physical phenomena such as impact ionization or Auger recombination, for example.

Next, a systematic procedure of incorporating the encapsulated device models into power electronic circuits containing voltage sources, current sources, resistors, inductors, and/or capacitors was set forth in Chapter 4. The circuit equations establish and supply the anode and cathode voltages to the encapsulated device models while the device models establish and supply the anode and cathode currents to the circuit equations via dependent current sources. The circuit equations constitute border matrices of a sparse block-structured system of equations (i.e. tableau). Extension to
include multi-dimensional device models and/or multi-terminal devices will not affect the block structure of the tableau equations.

In order to validate the device models and simulation methodology set forth in Chapter 4, a simple power electronic circuit, consisting of a single-phase square-wave voltage source, resistor, inductor, and diode, was both simulated and constructed using a commercially available PIN diode. Using this test circuit, a parameter extraction procedure was set forth in Chapter 5 to establish the diode's physical parameters needed for circuit-level simulations. Therein, it was shown that, once the physical diode parameters are established, the simulated and measured reverse-recovery characteristics of the diode are in excellent agreement.

To evaluate and validate the simulation approach for a more practical power electronic circuit, a single-phase full-bridge rectifier was both simulated and constructed. The results of an extensive set of experimental and computer studies were presented in Chapter 6. Therein, the time-step requirements and overall computational performance was discussed for a conventional single-threaded solution of the discretized equations. Comparisons of the simulated and measured responses revealed excellent agreement in terms of voltages, currents, and power losses.

After establishing the validity of the modeling and simulation approach, attention then focused on analyzing and improving its computational performance. Computational bottlenecks in the simulation were identified in Chapter 7 and a block Gaussian elimination (BGE) technique was used to solve a block-structured system of linear equations that must be solved, typically several times per time step. BGE allows computationally intensive tasks to be performed in parallel using multi-threaded programming techniques thereby improving simulation speed. Other parallizable tasks such as Jacobian and gradient evaluation, and residual computation were also identified.

To further improve computational performance, a variable-structure variable-order (VSVO) simulation paradigm was also set forth in Chapter 7. In this paradigm, inactive devices (e.g., reverse-biased diodes) are assumed to be disconnected from
the circuit (hence variable structure) resulting in a system of equations of much lower order (hence variable order). An indirect and not-so-obvious benefit of a VSVO-based simulation is that a much coarser mesh can be used for all diodes within the circuit since the depletion region is physically much wider when a diode is reverse-biased or inactive and since a relatively fine mesh is required throughout the deletion region. Thus, VSVO-based simulations will involve significantly fewer variables.

Essential to a VSVO simulation is the determination of when to disconnect an active device or reconnect an inactive one. For this purpose, a simple-to-implement and reliable (immune to false triggers) diode activity monitor was set forth. The observed variations of species densities within the device can be used to implement activity monitors for other devices such as IGBTs, MOSFETs, and BJTs. The activity monitor is used to vary the tableau matrix structure seamlessly between different possible circuit topologies and is capable of utilizing the available computational resources (threads) to the maximum extent possible. In the example simulation of a full-bridge rectifier with four diodes, it was shown that a VSVO simulation is capable of achieving a $484 \%$ improvement over a conventional single-threaded simulation with no observable loss of accuracy.

### 8.1 Future Work

The areas in which the research can be extended are discussed in the following subsections.

### 8.1.1 Multi-dimensional/multi-terminal devices

IGBTs can be modeled using its inherent bipolar $p^{+} n^{-} p^{+}$structure [93]. The currents entering and leaving the two $p^{+}$-regions are used to model the device current. Following the approaches described in references [93, 94], it is possible to readily extend this thesis to include IGBTs in circuit simulations. In three-terminal devices, three dependent-current sources are needed to couple the device and circuit equations.

In some devices, it may be needed to represent spatial effects in two-dimensional. Accordingly, the procedure set forth in Chapter 4 must modified to build the necessary border matrices. A two dimensional discretization based upon the ScharfetterGummel approach is described in [95] and can be readily applied without affecting the other procedures or block structure of the tableau equation.

### 8.1.2 Coupled Electro-thermal modeling

Since many of the physical parameters, hence device terminal behaviors, are temperature dependent, it is desirable to incorporate a thermal model into the CDDC equations. The heat conduction equation is very similar to Poisson's equation (one of the CDDC equations) with right-hand side being the heat source term divided by the thermal conductivity. A model for heat source terms needs can be derived starting with the equations in references [96, 97]. At a macroscopic level in power semiconductor devices, thermal phenomena are temporally slower than the transport phenomena. Consequently, a relaxation technique can be used wherein a coarser mesh is used for the thermal system, which is solved at fixed intervals, whose solution is then used to update the temperature-dependent semiconductor material properties. In this case, the coupling matrix representing the coupling between thermal and transport phenomena is zero. It is also possible to include a fully coupled electro-thermal simulation with more computational effort. A time-dependent heat equation augmented with the transport model in [98] is capable of predicting the electro-thermal behavior more accurately than the aforementioned techniques.

### 8.1.3 Numerical aspects

In this thesis, MATLAB was used to implement both the full and VSVO simulations. A C/C++ implementation using shared memory programming can result in additional improvements in speed because of the fundamental difference between interpreted and compiled languages. However, for more complex circuits involving
multi-dimensional device models, a combination of shared and distributed (OpenMP and MPI) programming approaches may prove to be more appropriate [99]. It is prudent to consider freely available libraries, with large user base, and with long term support such as Trilinos [100] and PETSc [101] for development to include such advanced capabilities. The $\mathrm{C} / \mathrm{C}++$ implementation with the capability of simulating multi-terminal devices appears to be the most logical extension of this thesis.

Dividing the overall simulation hierarchically into devices and then further into smaller problems can achieve further improvements using thread numbers greater than the number devices in the circuit. A strategy similar to that used in the dissertation [84] can be used to geometrically divide the device into smaller geometric domains to increase simulation speed. An algebraic partitioning of the matrix as in [102] can also be used to achieve similar results.

Finally, for systems with a very large number of variables, there exists a class of iterative methods that do not require explicit formation of the full Jacobian matrix. Specifically, the so-called Jacobian-free Newton-Krylov [82] method may prove useful for simulating complex circuits with multi-dimensional devices. Such circuits will involve a tableau with a very large number of variables. If the computational gain of not having to explicitly form the Jacobian exceeds the computational effort needed to precondition and iteratively solve the linear system of equations, iterative methods can be more competitive than direct methods. Such a problem-specific iterative method can be formulated using the references $[76,82,83]$.

REFERENCES

## REFERENCES

[1] P. Antognetti and G. Massobrio, Semiconductor Device Modeling with Spice. New York, NY, USA: McGraw-Hill, Inc., 1990.
[2] A. Corp, "Medici," Palo Alto, 2000.
[3] A. Jüngel, Transport Equations for Semiconductors, 1st ed. Berlin Heidelberg: Springer, 2009.
[4] J. W. Jerome, "Semiconductor models: Their mathematical study and approximation," 1992.
[5] A. Jüngel, "Transport equations for semiconductors," Johannes GutenbergUniversität Mainz, 2004.
[6] W. van Roosbroeck, "Theory of flow of electron and holes in germanium and other semiconductors," Bell Syst. Tech. J., vol. 29, pp. 560-607, 1950.
[7] M. Zláal, "Finite element solution of the fundamental equations of semiconductor devices," Math. Comput., vol. 46, pp. 27-43, January 1986. [Online]. Available: http://portal.acm.org/citation.cfm?id=21230.21232
[8] R. Verfürth, "Adaptive finite element methods lecture notes winter term 2007/08."
[9] -_, "Computaional fluid dynamics lecture notes summer term 2007."
[10] L. F. Shampine and M. W. Reichelt, "The matlab ode suite."
[11] L. R. Petzold, "A description of dassl: a differential/algebraic system solver," in International Conference on Scientific Computing, 1983.
[12] J. Lang, Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Springer, 2001.
[13] H. Gummel, "A self-consistent iterative scheme for one-dimensional steady state transistor calculations," Electron Devices, IEEE Transactions on, vol. 11, no. 10, pp. $455-465$, oct 1964 .
[14] E. M. Buturla, P. E. Cottrell, B. M. Grossman, and K. A. Salsburg, "Finiteelement analysis of semiconductor devices: The fielday program," IBM Journal of Research and Development, vol. 25, no. 4, pp. 218-231, july 1981.
[15] A. Brown, A. Asenov, S. Roy, and J. Barker, "Development of a parallel 3d finite element power semiconductor device simulator," in Physical Modelling of Semiconductor Devices, IEE Colloquium on, apr 1995, pp. 2/1-2/6.
[16] O. Heinreichsberger, S. Selberherr, M. Stiftinger, and K. P. Traar, "Fast iterative solution of carrier continuity equations for three-dimensional device simulation," SIAM J. Sci. Stat. Comput., vol. 13, pp. 289-306, January 1992. [Online]. Available: http://portal.acm.org/citation.cfm? $\mathrm{id}=132131.132155$
[17] M. Davis and G. Carey, "Multilevel solution of augmented drift-diffusion equations," COMPEL: International Journal of Computations and Mathematics in Electrical, vol. 15, pp. 4-18(15), 1 February 1996.
[18] D. Scharfetter and H. Gummel, "Large-signal analysis of a silicon read diode oscillator," Electron Devices, IEEE Transactions on, vol. 16, no. 1, pp. 64-77, jan 1969.
[19] G. Heiser, C. Pommerell, J. Weis, and W. Fichtner, "Three-dimensional numerical semiconductor device simulation: algorithms, architectures, results," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 10, no. 10, pp. 1218-1230, oct 1991.
[20] M. Reiser, "A two-dimensional numerical fet model for dc, ac, and large-signal analysis," Electron Devices, IEEE Transactions on, vol. 20, no. 1, pp. 35-45, jan 1973.
[21] C. M. Snowden, "Semiconductor device modelling," Reports on Progress in Physics, vol. 48, no. 2, p. 223, 1985. [Online]. Available: http: //stacks.iop.org/0034-4885/48/i=2/a=002
[22] R. Bank, D. Rose, and W. Fichtner, "Numerical methods for semiconductor device simulation," Electron Devices, IEEE Transactions on, vol. 30, no. 9, pp. 1031-1041, sep 1983.
[23] M. Fournié, "High order conservative difference methods for 2d driftdiffusion model on non-uniform grid," Applied Numerical Mathematics, vol. 33, no. 1-4, pp. $381-392$, 2000. [Online]. Available: http: //www.sciencedirect.com/science/article/pii/S0168927499001051
[24] J. Trattles and C. Johnson, "Discretisation techniques for the physical simulation of power semiconductor devices," in Power Semiconductor Devices, IEE Colloquium on, dec 1994, pp. 8/1-8/6.
[25] B. Gokhale, "Numerical solutions for a one-dimensional silicon n-p-n transistor," Electron Devices, IEEE Transactions on, vol. 17, no. 8, pp. 594-602, aug 1970.
[26] R. Bank, W. Coughran, W. Fichtner, E. Grosse, D. Rose, and R. Smith, "Transient simulation of silicon devices and circuits," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 4, no. 4, pp. 436 451, october 1985.
[27] M. Burger and R. Pinnau, "Fast optimal design of semiconductor devices," SIAM Journal on Applied Mathematics, vol. 64, no. 1, pp. pp. 108-126, 2003. [Online]. Available: http://www.jstor.org/stable/4096032
[28] J. Machek and S. Selberherr, "A novel finite-element approach to device modeling," Electron Devices, IEEE Transactions on, vol. 30, no. 9, pp. 1083 - 1092, sep 1983.
[29] N. Mohamed and M. Z. Sujod, "Finite elements in semiconductor devices," Information Management and Engineering, International Conference on, vol. 0, pp. 108-110, 2009.
[30] R. Lomax and S. Y. Oh, "Finite elements for semiconductor device simulation," in Circuits, Systems and Computers, 1977. Conference Record. 1977 11th Asilomar Conference on, nov 1977, pp. $35-40$.
[31] H. Sakata, K. Sugimoto, S. Isomura, and E. Masada, "One and two dimensional device simulation of gto using finite element method," in Power Conversion Conference, 1993. Yokohama 1993., Conference Record of the, apr 1993, pp. 569-574.
[32] M. Sharma and G. Carey, "Semiconductor device simulation using adaptive refinement and flux upwinding," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 8, no. 6, pp. 590-598, jun 1989.
[33] Z. Chen, "Finite element analysis of the one-dimensional full drift-diffusion semiconductor model," SIAM J. Numer. Anal., vol. 32, pp. 455-483, April 1995. [Online]. Available: http://portal.acm.org/citation.cfm?id=207475.207484
[34] G.-L. Tan, X.-L. Yuan, Q.-M. Zhang, W. Ku, and A.-J. Shey, "Two-dimensional semiconductor device analysis based on new finite-element discretization employing the s-g scheme," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 8, no. 5, pp. 468 -478, may 1989.
[35] B. Schmithusen, K. Gartner, and W. Fichtner, "A novel grid adaption procedure for stationary 2d device simulation," in Proc. Tech. Nanotechnol. Conf. Trade Show, Feb 2003, pp. 504-507.
[36] R. Guerrieri and M. Rudan, "Optimization of the finite-element solution of the semiconductor-device poisson equation," Electron Devices, IEEE Transactions on, vol. 30, no. 9, pp. 1097 - 1103, sep 1983.
[37] J. Burgler, J. Coughran, W.M., and W. Fichtner, "An adaptive grid refinement strategy for the drift-diffusion equations," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 10, no. 10, pp. 1251 -1258, oct 1991.
[38] I. Babuška and W. C. Rheinboldt, "A-posteriori error estimates for the finite element method," International Journal for Numerical Methods in Engineering, vol. 12, no. 10, pp. 1597-1615, 1978. [Online]. Available: http://dx.doi.org/10.1002/nme. 1620121010
[39] K. Tseng, S. Pan, and C. Foo, "Issues in power diode modelling for circuit simulation," in Power Electronics and Drive Systems, 1997. Proceedings., 1997 International Conference on, vol. 1, may 1997, pp. 78-84 vol.1.
[40] P. Leturcq, "Power semiconductor device modelling dedicated to circuit simulation," in Power Semiconductor Devices and ICs, 1999. ISPSD '99. Proceedings., The 11th International Symposium on, 1999, pp. 19-26.
[41] A. Strollo, "A new spice model of power p-i-n diode based on asymptotic waveform evaluation," Power Electronics, IEEE Transactions on, vol. 12, no. 1, pp. $12-20$, jan 1997.
[42] R. Chibante, A. Araujo, and A. Carvalho, "A fem punch-through igbt model using an efficient parameter extraction method," in Industrial Electronics Society, 2005. IECON 2005. 31st Annual Conference of IEEE, nov. 2005, p. 6 pp.
[43] R. Kolessar and H.-P. Nee, "A new physics-based circuit model for 4h-sic power diodes implemented in saber," in Applied Power Electronics Conference and Exposition, 2001. APEC 2001. Sixteenth Annual IEEE, vol. 2, 2001, pp. 989 -994 vol.2.
[44] R. Chibante, A. Araujo, and A. Carvalho, "Finite-element modeling and optimization-based parameter extraction algorithm for npt-igbts," Power Electronics, IEEE Transactions on, vol. 24, no. 5, pp. 1417-1427, may 2009.
[45] R. Kolessar, "Physical study of the power diode turn-on process," in Industry Applications Conference, 2000. Conference Record of the 2000 IEEE, vol. 5, 2000 , pp. $2934-2940$ vol. 5.
[46] C. Schmeiser, "A model for the transient behaviour of long-channel mosfets," SIAM J. Appl. Math., vol. 54, pp. 175-194, February 1994. [Online]. Available: http://dx.doi.org/10.1137/S0036139992225759
[47] S. Schröder and R. De Doncker, "Physically based models of high power semiconductor devices for pspice," in Industry Applications Conference, 1999. Thirty-Fourth IAS Annual Meeting. Conference Record of the 1999 IEEE, vol. 1, 1999, pp. $379-384$ vol.1.
[48] C. Ma, P. Lauritzen, P.-Y. Lin, I. Budihardjo, and J. Sigg, "A systematic approach to modeling of power semiconductor devices based on charge control principles," in Power Electronics Specialists Conference, PESC '94 Record., 25th Annual IEEE, jun 1994, pp. 31-37 vol.1.
[49] N. Krihely and S. Ben-Yaakov, "Modeling and evaluation of diode reverse recovery in discrete-transition simulators," in Energy Conversion Congress and Exposition (ECCE), 2010 IEEE, 2010, pp. 4514-4520.
[50] J. Tichenor, S. Sudhoff, and J. Drewniak, "Behavioral igbt modeling for predicting high frequency effects in motor drives," Power Electronics, IEEE Transactions on, vol. 15, no. 2, pp. $354-360$, mar 2000.
[51] Z. John Shen, Y. Xiong, X. Cheng, Y. Fu, and P. Kumar, "Power mosfet switching loss analysis: A new insight," in Industry Applications Conference, 2006. 41st IAS Annual Meeting. Conference Record of the 2006 IEEE, vol. 3, oct. 2006, pp. $1438-1442$.
[52] A. Q. Huang and B. Zhang, "Comparing sic switching power devices: Mosfet, npn transistor and gto thyristor," Solid-State Electronics, vol. 44, no. 2, pp. 325 - 340, 2000. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S0038110199002397
[53] M. Chinthavali, L. Tolbert, and B. Ozpineci, "Sic gto thyristor model for hvdc interface," in Power Engineering Society General Meeting, 2004. IEEE, june 2004, pp. 680 - 685 Vol.1.
[54] M. Aghdam and T. Thiringer, "Comparison of sic and si power semiconductor devices to be used in 2.5 kw dc/dc converter," in Power Electronics and Drive Systems, 2009. PEDS 2009. International Conference on, nov. 2009, pp. 1035 -1040.
[55] R. Pierret, Advanced Semiconductor Fundamentals, 2nd ed. New Jersey: Prentice Hall, 2002.
[56] B. Baliga, Fundamentals of Power Semiconductor Devices. New York: Springer, 2008.
[57] S. Sze, Physics of Semiconductor Devices, 3rd ed. New Jersey: John Wiley, 2007.
[58] P. Deuflhard, Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms. Springer Publishing Company, Incorporated, 2011.
[59] S. Brenner and L. Scott, The Mathematical Theory of Finite Element Methods, 3rd ed. New York: Springer, 2008.
[60] H. Versteeg and W. Malalasekara, An Introduction to Computational Fluid Dynamics. Pearson Education Limited, 2007.
[61] E. Hairer and G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems. Berlin: Springer Series in Computational Mathematics, 1991.
[62] R. Riaza, Differential-Algebraic Systems: Analytical Aspects and Circuit Applications. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 2008.
[63] C. Tischendorf, "Coupled systems of differential algebraic and partial differential equations in circuit and device simulation," Ph.D. dissertation, HumboldtUniversität zu Berlin, 2003.
[64] E. Keiter, S. Hutchinson, R. Hoekstra, E. Rankin, T. Russo, and L. Waters, "Computational algorithms for device-circuit coupling," Sandia Report, no. SAND2003-0080, 2003.
[65] E. Hairer, S. Norsett, and G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems. Berlin: Springer-Verlag, 1993.
[66] K. Brenan, S. Campbell, and L. Petzold, Numerical Solution of Inital-Value Problems in Differential-Algebraic Equations. Philadelphia: SIAM, 1996.
[67] H. Garrab, B. Allard, H. Morel, K. Ammous, S. Ghedira, A. Amimi, K. Besbes, and J.-M. Guichon, "On the extraction of pin diode design parameters for validation of integrated power converter design," Power Electronics, IEEE Transactions on, vol. 20, no. 3, pp. 660-670, 2005.
[68] P. Krein, Elements of Power Electronics, 1st ed. New York: Oxford, 1998.
[69] S. Linder, Power Semiconductors, 1st ed. Switzerland: EPFL Press, 2006.
[70] R. Pierret, Semiconductor Device Fundamentals, 2nd ed. Addison Wesley, 1996.
[71] U. M. Ascher and L. R. Petzold, Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations, 1st ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1998.
[72] P. Bochev and D. Ridzal, "Large-scale scientific computing," I. Lirkov, S. Margenov, and J. Waśniewski, Eds. Berlin, Heidelberg: SpringerVerlag, 2008, ch. Finite Element Solution of Optimal Control Problems Arising in Semiconductor Modeling, pp. 235-242. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-78827-0_25
[73] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2003.
[74] "Tektronix current probe and amplifiers," http://www.tek.com/sites/tek.com/ files/media/media/resources/TCPA300_TCP312_TCP305_TCP303_TCPA400_ TCP404XL_Current_Probes_and_Amplifiers_Datasheet_60W-16458-6.pdf, accessed: February,2015.
[75] K. Gustafsson, "Control theoretic techniques for stepsize selection in explicit runge-kutta methods," ACM Trans. Math. Softw., vol. 17, no. 4, pp. 533-554, Dec. 1991. [Online]. Available: http://doi.acm.org/10.1145/210232.210242
[76] R. Barrett, M. W. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, Templates for the solution of linear systems: building blocks for iterative methods. Siam, 1994, vol. 43.
[77] G. H. Golub and C. F. Van Loan, Matrix computations (3rd ed.). Baltimore, MD, USA: Johns Hopkins University Press, 1996.
[78] T. A. Davis, "Algorithm 832: Umfpack v4.3-an unsymmetric-pattern multifrontal method," ACM Trans. Math. Softw., vol. 30, no. 2, pp. 196-199, Jun. 2004. [Online]. Available: http://doi.acm.org/10.1145/992200.992206
[79] -_, Direct Methods for Sparse Linear Systems (Fundamentals of Algorithms 2). Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2006.
[80] O. Schenk and K. Gärtner, "Solving unsymmetric sparse systems of linear equations with pardiso," Future Gener. Comput. Syst., vol. 20, no. 3, pp. 475-487, Apr. 2004. [Online]. Available: http://dx.doi.org/10.1016/j.future. 2003.07.011
[81] J. W. Demmel, S. C. Eisenstat, J. R. Gilbert, X. S. Li, and J. W. H. Liu, "A supernodal approach to sparse partial pivoting," SIAM J. Matrix Analysis and Applications, vol. 20, no. 3, pp. 720-755, 1999.
[82] D. A. Knoll and D. E. Keyes, "Jacobian-free newton-krylov methods: A survey of approaches and applications," J. Comput. Phys., vol. 193, no. 2, pp. 357-397, Jan. 2004. [Online]. Available: http://dx.doi.org/10.1016/j.jcp.2003.08.010
[83] C. T. Kelley, Iterative Methods for Linear and Nonlinear Equations, ser. Frontiers in Applied Mathematics. SIAM, 1995, no. 16. [Online]. Available: http://www.siam.org/books/textbooks/fr16_book.pdf
[84] J. H. Cole, "Optimum partitioning strategies in partitioned finite element analysis of electromechanical devices," Ph.D. dissertation, Purdue University, 2011.
[85] B. Barney, "Introduction to parallel computing," https://computing.llnl.gov/ tutorials/parallel_comp/, accessed: February,2015.
[86] _—, "Message passing interface," https://computing.llnl.gov/tutorials/mpi/, accessed: February,2015.
[87] __, "Openmp," https://computing.llnl.gov/tutorials/mpi/, accessed: February,2015.
[88] G. M. Amdahl, "Validity of the single processor approach to achieving large scale computing capabilities," in Proceedings of the April 1820, 1967, Spring Joint Computer Conference, ser. AFIPS '67 (Spring). New York, NY, USA: ACM, 1967, pp. 483-485. [Online]. Available: http://doi.acm.org/10.1145/1465482.1465560
[89] J. L. Gustafson, "Reevaluating amdahl's law," Commun. ACM, vol. 31, no. 5, pp. 532-533, May 1988. [Online]. Available: http://doi.acm.org/10.1145/ 42411.42415
[90] P. N. Brown, A. C. Hindmarsh, and L. R. Petzold, "Consistent initial condition calculation for differential-algebraic systems," SIAM J. Sci. Comput., vol. 19, no. 5, pp. 1495-1512, Sep. 1998. [Online]. Available: http://dx.doi.org/10.1137/S1064827595289996
[91] "How parallel computing products run a job," http://www.mathworks.com/ help/distcomp/how-parallel-computing-products-run-a-job.html, accessed: April, 2015.
[92] M. Stynes, "Numerical methods for convection-diffusion problems or the 30 years war," arXiv preprint arXiv:1306.5172, 2013.
[93] R. Chibante, A. Araujo, and A. Carvalho, "A new physics based spice model for npt igbts," in Industrial Electronics Society, 2003. IECON '03. The 29th Annual Conference of the IEEE, vol. 2, Nov 2003, pp. 1156-1161 Vol.2.
[94] X. Kang, E. Santi, J. Hudgins, P. Palmer, and J. Donlon, "Parameter extraction for a physics-based circuit simulator igbt model," in Applied Power Electronics Conference and Exposition, 2003. APEC '03. Eighteenth Annual IEEE, vol. 2, Feb 2003, pp. 946-952 vol.2.
[95] P. Bochev, K. Peterson, and X. Gao, "A new control volume finite element method for the stable and accurate solution of the drift-diffusion equations on general unstructured grids," Computer Methods in Applied Mechanics and Engineering, vol. 254, pp. 126-145, 2013.
[96] P. Gough, P. Walker, and K. Whight, "Electrothermal simulation of power semiconductor devices," in Power Semiconductor Devices and ICs, 1991. ISPSD '91., Proceedings of the 3rd International Symposium on, Apr 1991, pp. 89-94.
[97] B. Du, J. Hudgins, E. Santi, A. Bryant, P. Palmer, and H. Mantooth, "Transient electrothermal simulation of power semiconductor devices," Power Electronics, IEEE Transactions on, vol. 25, no. 1, pp. 237-248, Jan 2010.
[98] K. Kells, S. Mller, G. Wachutka, and W. Fichtner, "Simulation of self-heating effects in a power p-i-n diode," in Simulation of Semiconductor Devices and Processes, S. Selberherr, H. Stippel, and E. Strasser, Eds. Springer Vienna, 1993, pp. 41-44. [Online]. Available: http://dx.doi.org/10.1007/978-3-7091-6657-4_9
[99] T. Hoefler, J. Dinan, D. Buntinas, P. Balaji, B. Barrett, R. Brightwell, W. Gropp, V. Kale, and R. Thakur, "Mpi+ mpi: a new hybrid approach to parallel programming with mpi plus shared memory," Computing, vol. 95, no. 12, pp. 1121-1136, 2013.
[100] M. Heroux, R. Bartlett, V. H. R. Hoekstra, J. Hu, T. Kolda, R. Lehoucq, K. Long, R. Pawlowski, E. Phipps, A. Salinger, H. Thornquist, R. Tuminaro, J. Willenbring, and A. Williams, "An Overview of Trilinos," Sandia National Laboratories, Tech. Rep. SAND2003-2927, 2003.
[101] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. F. Smith, and H. Zhang, "PETSc users manual," Argonne National Laboratory, Tech. Rep. ANL-95/11 - Revision 3.5, 2014. [Online]. Available: http://www.mes.anl.gov/petsc
[102] E. Polizzi and A. H. Sameh, "A parallel hybrid banded system solver: The spike algorithm," Parallel Comput., vol. 32, no. 2, pp. 177-194, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1016/j.parco.2005.07.005

APPENDICES

## A. DERIVATIONS IN ONE DIMENSION

## A. 1 Poisson's Equation

A one-dimensional partition of the space, $0=x_{1}<x_{2} \ldots<x_{i}<\ldots<x_{N}=X$, with piecewise linear functions is shown in Figure A.1. In Figure A. 1 it can be seen that $i^{\text {th }}$ basis function overlaps with $\{i-1, i, i+1\}$ basis functions. The Poisson's equation after integration by parts is shown here for convenience.

$$
\begin{equation*}
\int_{0}^{X} \varepsilon \psi^{\prime}(x) \sum_{i=1}^{N} \varphi_{i}^{\prime} d x=\int_{0}^{X} \rho(x) \sum_{i=1}^{N} \varphi_{i} d x \tag{A.1}
\end{equation*}
$$



Fig. A.1.: One-dimensional discretization.

The piecewise linear function $\varphi_{i}$ is used in (A.1). It is shown again here for reference.

$$
\varphi_{i}(x)= \begin{cases}\frac{x-x_{i-1}}{x_{i}-x_{i-1}}, & x \in\left[x_{i-1}, x_{i}\right]  \tag{A.2}\\ \frac{x_{i+1}-x}{x_{i+1}-x_{i}}, & x \in\left[x_{i}, x_{i+1}\right] \\ 0, & \text { otherwise }\end{cases}
$$

For a given $i$, this integration over the entire domain is split into intervals. The integral evaluation is nonzero for only for the interval $\left[x_{i-1}, x_{i+1}\right]$. That is,

$$
\begin{equation*}
\int_{0}^{X} \varepsilon \psi^{\prime}(x) \varphi_{i}^{\prime}(x) d x=\int_{x_{i-1}}^{x_{i}} \varepsilon \psi_{i-1} \varphi_{i-1}^{\prime} \varphi_{i}^{\prime} d x+\int_{x_{i-1}}^{x_{i+1}} \varepsilon \psi_{i} \varphi_{i}^{\prime} \varphi_{i}^{\prime} d x+\int_{x_{i}}^{x_{i+1}} \varepsilon \psi_{i+1} \varphi_{i+1}^{\prime} \varphi_{i}^{\prime} d x \tag{A.3}
\end{equation*}
$$

It is convenient to work with canonical elements for the derivations in the intervals $\left[x_{i-1}, x_{i+1}\right]$. The elements in interval $h_{i-1}$ with $\left[x_{i-1}, x_{i}\right]$ are mapped onto canonical element with $\hat{x} \in[0,1]$ (refer Figure A.2).

$$
\begin{aligned}
x & =x_{i-1}+\hat{x} h_{i-1} \\
\hat{x} & =\frac{x-x_{i-1}}{h_{i-1}} \\
d x & =d \hat{x} h_{i-1}
\end{aligned}
$$

In interval $h_{i-1}$, the nonzero basis functions are $\varphi_{i-1}$ and $\varphi_{i}$. These functions transformed to $\hat{x}$ coordinate are given as,

$$
\begin{aligned}
\varphi_{i}(\hat{x}) & =\frac{\hat{x}}{h_{i-1}}, \\
\frac{d \varphi_{i}(\hat{x})}{d \hat{x}} & =\frac{1}{h_{i-1}}, \\
\varphi_{i-1}(\hat{x}) & =\frac{1-\hat{x}}{h_{i-1}}, \\
\frac{d \varphi_{i-1}(\hat{x})}{d \hat{x}} & =\frac{-1}{h_{i-1}}
\end{aligned}
$$



Fig. A.2.: Canonical element.

The terms in (A.3) are evaluated with the help of coordinate transformation as in the following equations.

$$
\begin{align*}
\int_{x_{i-1}}^{x_{i}} \varphi_{i-1}^{\prime} \varphi_{i}^{\prime} d x & =\int_{0}^{1} \frac{-1}{h_{i-1}} \frac{1}{h_{i-1}} h_{i-1} d \hat{x}=-\frac{1}{h_{i-1}}  \tag{A.5}\\
\int_{x_{i-1}}^{x_{i+1}} \varphi_{i}^{\prime} \varphi_{i}^{\prime} d x & =\int_{x_{i-1}}^{x_{i}} \varphi_{i}^{\prime} \varphi_{i}^{\prime} d x+\int_{x_{i}}^{x_{i+1}} \varphi_{i}^{\prime} \varphi_{i}^{\prime} d x \\
& =\int_{0}^{1} \frac{1}{h_{i-1}} \frac{1}{h_{i-1}} h_{i-1} d \hat{x}+\int_{0}^{1} \frac{-1}{h_{i}} \frac{-1}{h_{i}} h_{i} d \hat{x}
\end{align*}
$$

$$
\begin{align*}
& =\frac{1}{h_{i-1}}+\frac{1}{h_{i}}  \tag{A.6}\\
\int_{x_{i}}^{x_{i+1}} \varphi_{i+1}^{\prime} \varphi_{i}^{\prime} d x & =\int_{0}^{1} \frac{1}{h_{i}} \frac{-1}{h_{i}} h_{i} d \hat{x}=-\frac{1}{h_{i}} \tag{A.7}
\end{align*}
$$

So the $i^{\text {th }}$ equation is written as,

$$
\begin{equation*}
\int_{0}^{X} \varepsilon \psi^{\prime}(x) \varphi_{i}^{\prime}(x) d x=-\frac{\varepsilon}{h_{i-1}} \psi_{i-1}+\left(\frac{\varepsilon}{h_{i-1}}+\frac{\varepsilon}{h_{i}}\right) \psi_{i}-\frac{\varepsilon}{h_{i}} \psi_{i+1} \tag{A.8}
\end{equation*}
$$

It is assumed that right-hand side function $\rho$ can be approximated by piecewise linear functions similar to the electrical potential. The $i^{\text {th }}$ equation is derived similar to the left-hand side with nodal values of the electric charge $\rho_{i}=\rho\left(x_{i}\right)$.

$$
\begin{equation*}
\int_{0}^{X} \rho(x) \varphi_{i}(x) d x=\int_{x_{i-1}}^{x_{i}} \rho_{i-1} \varphi_{i-1} \varphi_{i} d x+\int_{x_{i-1}}^{x_{i+1}} \rho_{i} \varphi_{i} \varphi_{i} d x+\int_{x_{i}}^{x_{i+1}} \rho_{i+1} \varphi_{i+1} \varphi_{i} d x \tag{A.9}
\end{equation*}
$$

Using the canonical element as derived above the individual integrals are evaluated.

$$
\begin{align*}
\int_{x_{i-1}}^{x_{i}} \varphi_{i-1} \varphi_{i} d x & =\int_{0}^{1} \hat{x}(1-\hat{x}) h_{i-1} d \hat{x} \\
& =\left.\frac{h_{i-1} \hat{x}^{2}}{2}\right|_{0} ^{1}-\left.\frac{h_{i-1} \hat{x}^{3}}{3}\right|_{0} ^{1} \\
& =\frac{h_{i-1}}{6}  \tag{A.10}\\
\int_{x_{i-1}}^{x_{i+1}} \varphi_{i} \varphi_{i} d x & =\int_{x_{i-1}}^{x_{i}} \varphi_{i} \varphi_{i} d x+\int_{x_{i}}^{x_{i+1}} \varphi_{i} \varphi_{i} d x \\
& =\int_{0}^{1} \hat{x}^{2} h_{i-1} d \hat{x}+\int_{0}^{1}(1-\hat{x})^{2} h_{i} d \hat{x} \\
& =\left.\frac{h_{i-1} \hat{x}^{3}}{3}\right|_{0} ^{1}-\left.\frac{h_{i}(1-\hat{x})^{3}}{3}\right|_{0} ^{1} \\
& =\frac{h_{i-1}}{3}+\frac{h_{i}}{3}  \tag{A.11}\\
& =\frac{h_{i}}{6}
\end{align*}
$$

The $i^{\text {th }}$ equation of the right-hand side is then written as,

$$
\begin{equation*}
\int_{0}^{X} \rho(x) \varphi_{i}(x) d x=\frac{h_{i-1}}{6} \rho_{i-1}+\left(\frac{h_{i-1}}{3}+\frac{h_{i}}{3}\right) \rho_{i}+\frac{h_{i}}{6} \rho_{i+1}, \tag{A.13}
\end{equation*}
$$

where $\rho_{i}=q_{e} N_{M}\left(p_{i}-n_{i}+N_{D, i}^{+}-N_{A, i}^{-}\right)$.
The source vector can be rewritten as a product of a matrix-vector product $\rho=$ Ď $\check{\text {. }}$ Equations (A.8) and (A.13) are assembled into a matrix structure as follows,

$$
\begin{align*}
& {\left[\begin{array}{cccccc}
\frac{\varepsilon}{h_{1}} & -\frac{\varepsilon}{h_{1}} & 0 & \cdots & 0 & 0 \\
-\frac{\varepsilon}{h_{1}} & \frac{\varepsilon}{h_{1}}+\frac{\varepsilon}{h_{2}} & -\frac{\varepsilon}{h_{2}} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{\varepsilon}{h_{N-2}}+\frac{\varepsilon}{h_{N-1}} & -\frac{\varepsilon}{h_{N-1}} \\
0 & 0 & 0 & \cdots & -\frac{\varepsilon}{h_{N-1}} & \frac{\varepsilon}{h_{N-1}}
\end{array}\right]\left[\begin{array}{c}
\psi_{1} \\
\psi_{2} \\
\vdots \\
\psi_{N-1} \\
\psi_{N}
\end{array}\right]}  \tag{A.14}\\
& =q_{e} N_{M}\left[\begin{array}{cccccc}
\frac{h_{1}}{3} & \frac{h_{1}}{6} & 0 & \cdots & 0 & 0 \\
\frac{h_{1}}{6} & \frac{h_{1}+h_{2}}{3} & \frac{h_{2}}{6} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{h_{N-2}+h_{N-1}}{3} & \frac{h_{N-1}}{6} \\
0 & 0 & 0 & \cdots & \frac{h_{N-1}}{6} & \frac{h_{N-1}}{3}
\end{array}\right]\left[\begin{array}{c}
\rho_{1} \\
\rho_{2} \\
\vdots \\
\rho_{N-1} \\
\rho_{N}
\end{array}\right] .
\end{align*}
$$

The matrix on the left-hand side is known as stiffness matrix $\check{\mathbf{S}}$ and that on the right-hand side is mass matrix $\check{\mathbf{D}}$. These matrices are further modified to take boundary conditions into account.

## A. 2 Derivation of Jacobian Elements

The Newton-Raphson iterator (3.31) is given here for reference. This is rewritten directly to compute the next iterate solution ( $\mathbf{p}^{m, k+1}, \mathbf{n}^{m, k+1}, \boldsymbol{\psi}^{m, k+1}$ ) as in (A.16).

$$
\begin{align*}
& {\left[\begin{array}{ccc}
\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} \\
-\mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} \\
\mathbf{f}_{\psi, \mathbf{p}} & \mathbf{f}_{\psi, \mathbf{n}} & \mathbf{f}_{\psi, \boldsymbol{\psi}}
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{p} \\
\Delta \mathbf{n} \\
\Delta \boldsymbol{\psi}
\end{array}\right]}  \tag{A.15}\\
& =\left[\begin{array}{c}
-\left(\mathbf{p}^{m, k}-\mathbf{p}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{p}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right) \\
-\left(\mathbf{n}^{m, k}-\mathbf{n}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{n}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}\right) \\
-\mathbf{f}_{\psi}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\psi}^{m, k}, \boldsymbol{N}_{A}^{-}, \boldsymbol{N}_{D}^{+}\right)
\end{array}\right]
\end{align*}
$$

$$
\begin{align*}
& {\left[\begin{array}{ccc}
\mathbf{I}-\Delta t^{m} \mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\Delta t^{m} \mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\Delta t^{m} \mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} \\
-\Delta t^{m} \mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I}-\Delta t^{m} \mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\Delta t^{m} \mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} \\
\mathbf{f}_{\psi, \mathbf{p}} & \mathbf{f}_{\psi, \mathbf{n}} & \mathbf{f}_{\boldsymbol{\psi}, \boldsymbol{\psi}}
\end{array}\right]\left[\begin{array}{l}
\mathbf{p}^{m, k+1} \\
\mathbf{n}^{m, k+1} \\
\boldsymbol{\psi}^{m, k+1}
\end{array}\right]}  \tag{A.16}\\
& =\left[\begin{array}{c}
\mathbf{p}^{m, k}+\Delta t^{m}\left(\mathbf{f}_{\mathbf{p}}{ }^{m, k}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} \mathbf{p}^{m, k}-\mathbf{f}_{\mathbf{p}, \mathbf{n}} \mathbf{n}^{m, k}-\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} \boldsymbol{\psi}^{m, k}\right) \\
\mathbf{n}^{m, k}+\Delta t^{m}\left(\mathbf{f}_{\mathbf{n}}{ }^{m, k}-\mathbf{f}_{\mathbf{n}, \mathbf{p}} \mathbf{p}^{m, k}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} \mathbf{n}^{m, k}-\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} \boldsymbol{\psi}^{m, k}\right) \\
-\mathbf{f}_{\boldsymbol{\psi}}^{m, k}+\mathbf{f}_{\boldsymbol{\psi}, \mathbf{p}} \mathbf{p}^{m, k}+\mathbf{f}_{\psi, \mathbf{n}} \mathbf{n}^{m, k}+\mathbf{f}_{\psi, \psi} \boldsymbol{\psi}^{m, k}
\end{array}\right]
\end{align*}
$$

The iterator (A.15) or (A.16) can be used in the simulation. It can be seen that in (A.16) there are a couple of sparse Jacobian matrix-vector products as opposed to (A.15). Hence, the implementation using the Newton-Raphson iterator in (A.15) is more efficient. The derivation of the top two block-row Jacobian partitions in (A.16) are due to two different discretization methods and are different. The last block-row of the Jacobian in (A.16) is common to both discretization methods and it is derived first in the next subsection.

## A.2.1 Poisson's equation

The last row of the iterator above is written explicitly along with $\mathbf{f}_{\psi}$ as in the following set of equations.

$$
\begin{align*}
\mathbf{f}_{\psi, \mathbf{p}} \mathbf{p}^{m, k+1}+\mathbf{f}_{\psi, \mathbf{n}} \mathbf{n}^{m, k+1}+\mathbf{f}_{\psi, \boldsymbol{\psi}} \boldsymbol{\psi}^{m, k+1} & =-\mathbf{f}_{\boldsymbol{\psi}}+\mathbf{f}_{\psi, \mathbf{p}} \mathbf{p}^{m, k}+\mathbf{f}_{\psi, \mathbf{n}} \mathbf{n}^{m, k}+\mathbf{f}_{\psi, \psi} \boldsymbol{\psi}^{m, k} \\
\mathbf{f}_{\psi} & =\mathbf{S} \boldsymbol{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)-\mathbf{b}\left(v_{a}, v_{c}\right) \tag{A.17}
\end{align*}
$$

The Jacobian elements are derived in order as follows,

$$
\begin{align*}
& \mathbf{f}_{\psi, \mathbf{p}}=\frac{\partial \mathbf{f}_{\psi}}{\partial \mathbf{p}}=-\mathbf{D}  \tag{A.18}\\
& \mathbf{f}_{\psi, \mathbf{n}}=\frac{\partial \mathbf{f}_{\psi}}{\partial \mathbf{n}}=+\mathbf{D}  \tag{A.19}\\
& \mathbf{f}_{\psi, \psi}=\frac{\partial \mathbf{f}_{\psi}}{\partial \boldsymbol{\psi}}=\mathbf{S} \tag{A.20}
\end{align*}
$$

These values plugged in (A.17) and manipulated to establish the following equations.

$$
\begin{align*}
-\mathbf{D} \mathbf{p}^{m, k+1}+\mathbf{D} \mathbf{n}^{m, k+1}+\mathbf{S} \boldsymbol{\psi}^{m, k+1}= & -\mathbf{S} \boldsymbol{\psi}^{m, k}+\mathbf{D} \mathbf{p}^{m, k}-\mathbf{D} \mathbf{n}^{m, k} \\
& +\mathbf{D}\left(\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)+\mathbf{b}\left(v_{a}, v_{c}\right) \\
& -\mathbf{D} \mathbf{p}^{m, k}+\mathbf{D} \mathbf{n}^{m, k}+\mathbf{S} \boldsymbol{\psi}^{m, k}  \tag{A.21}\\
-\mathbf{D} \mathbf{p}^{m, k+1}+\mathbf{D} \mathbf{n}^{m, k+1}+\mathbf{S} \boldsymbol{\psi}^{m, k+1}= & \mathbf{D}\left(\boldsymbol{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)+\mathbf{b}\left(v_{a}, v_{c}\right) \tag{A.22}
\end{align*}
$$

In implementation either (A.17) or (A.22) is used depending on whether the NewtonRaphson iteration seeks $\boldsymbol{\chi}^{m, \infty}$ or the increment $\boldsymbol{\chi}^{m, \infty}-\boldsymbol{\chi}^{m-1, \infty}$ respectively.

## A.2.2 Central difference method

The right-hand side of the continuity equation is discretized in the following manner under central difference method. This equation is used as basis for the derivation of the Jacobian block matrices.

$$
\begin{align*}
f_{p}(i)= & -\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\mu_{p}\left(\frac{p_{i}+p_{i+1}}{2} \frac{\psi_{i}-\psi_{i+1}}{h_{i}}-\frac{p_{i-1}+p_{i}}{2} \frac{\psi_{i-1}-\psi_{i}}{h_{i-1}}\right)\right. \\
& \left.-D_{p}\left(\frac{p_{i+1}-p_{i}}{h_{i}}-\frac{p_{i}-p_{i-1}}{h_{i-1}}\right)\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)}  \tag{A.23a}\\
f_{n}(i)= & \frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\mu_{n}\left(\frac{n_{i}+n_{i+1}}{2} \frac{\psi_{i}-\psi_{i+1}}{h_{i}}-\frac{n_{i-1}+n_{i}}{2} \frac{\psi_{i-1}-\psi_{i}}{h_{i-1}}\right)\right. \\
& \left.+D_{n}\left(\frac{n_{i+1}-n_{i}}{h_{i}}-\frac{n_{i}-n_{i-1}}{h_{i-1}}\right)\right]-\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} \tag{A.23b}
\end{align*}
$$

The right-hand side of these equations involve $\{i-1, i, i+1\}$ nodal values of $(p, n, \psi)$ distributions. Then the nonzero entries of the block matrices correspond to $\{i-1, i, i+1\}$ columns of the $i^{\text {th }}$-row. The term common to $f_{p}$ and $f_{n}$ is the R-G term and it involves only $i^{\text {th }}$ nodal values. Defining a notation for R-G $\left(\mathcal{R}_{i}\right)$ term makes the derivation of this term's contribution to the Jacobian entries easy.

$$
\begin{align*}
\mathcal{R}_{i} & =\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)}  \tag{A.24a}\\
\frac{\partial \mathcal{R}_{i}}{\partial p_{i}} & =\frac{n_{i}}{\tau_{n}\left(p_{i}+p_{1}\right)+\tau_{p}\left(n_{i}+n_{1}\right)}-\frac{\tau_{n}\left(n_{i} p_{i}-n_{i e}^{2}\right)}{\left(\tau_{n}\left(p_{i}+p_{1}\right)+\tau_{p}\left(n_{i}+n_{1}\right)\right)^{2}} \tag{A.24b}
\end{align*}
$$

$$
\begin{equation*}
\frac{\partial \mathcal{R}_{i}}{\partial n_{i}}=\frac{p_{i}}{\tau_{n}\left(p_{i}+p_{1}\right)+\tau_{p}\left(n_{i}+n_{1}\right)}-\frac{\tau_{p}\left(n_{i} p_{i}-n_{i e}^{2}\right)}{\left(\tau_{n}\left(p_{i}+p_{1}\right)+\tau_{p}\left(n_{i}+n_{1}\right)\right)^{2}} \tag{A.24c}
\end{equation*}
$$

## A.2.2.1 $f_{p, p}, f_{p, n}, f_{p, \psi}$ matrices

The entries of matrices $\mathbf{f}_{\mathbf{p}, \mathbf{p}}, \mathbf{f}_{\mathbf{p}, \mathbf{n}}, \mathbf{f}_{\mathbf{p}, \psi}$ are derived in the following set of equations with the use of (A.23). The $i^{t h}$-row and $j^{\text {th }}$-column entry of the Jacobian partition is referred with the help $\mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, j)$.

$$
\begin{align*}
& \mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i-1)= \frac{\partial f_{p}(i)}{\partial p_{i-1}}=\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{p}}{2} \frac{\Delta \psi_{i-1 / 2}}{h_{i-1}}+D_{p} \frac{1}{h_{i-1}}\right]  \tag{A.25a}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i)= \frac{\partial f_{p}(i)}{\partial p_{i}} \\
&=-\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{p}}{2}\left(\frac{-\Delta \psi_{i+1 / 2}}{h_{i}}-\frac{-\Delta \psi_{i-1 / 2}}{h_{i-1}}\right)\right. \\
&\left.+D_{p}\left(\frac{1}{h_{i-1}}+\frac{1}{h_{i}}\right)\right]-\frac{\partial \mathcal{R}_{i}}{\partial p}(i)  \tag{A.25b}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i+1)=\frac{\partial f_{p}(i)}{\partial p_{i+1}}=\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{p}}{2} \frac{\Delta \psi_{i+1 / 2}}{h_{i}}+D_{p} \frac{1}{h_{i}}\right]  \tag{A.25c}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i-1)= \frac{\partial f_{p}(i)}{\partial n_{i-1}}=0  \tag{A.25d}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i)= \frac{\partial f_{p}(i)}{\partial n_{i}}=-\frac{\partial \mathcal{R}_{i}}{\partial n_{i}}  \tag{A.25e}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i+1)=\frac{\partial f_{p}(i)}{\partial n_{i+1}}=0  \tag{A.25f}\\
& \mathbf{f}_{\mathbf{p}, \mathbf{\psi}}(i, i-1)=\frac{\partial f_{p}(i)}{\partial \psi_{i-1}}=\frac{\mu_{p}}{\left(h_{i-1}+h_{i}\right)} \frac{\left(p_{i-1}+p_{i}\right)}{h_{i-1}}  \tag{A.25~g}\\
& \mathbf{f}_{\mathbf{p}, \psi}(i, i)=\frac{\partial f_{p}(i)}{\partial \psi_{i}}=-\frac{\mu_{p}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{p_{i-1}+p_{i}}{h_{i-1}}+\frac{p_{i+1}+p_{i}}{h_{i}}\right]  \tag{A.25h}\\
& \mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}}(i, i+1)=\frac{\partial f_{p}(i)}{\partial \psi_{i+1}}=\frac{\mu_{p}}{\left(h_{i-1}+h_{i}\right)} \frac{\left(p_{i+1}+p_{i}\right)}{h_{i}} \tag{A.25i}
\end{align*}
$$

## A.2.2.2 $f_{n, p}, f_{n, n}, f_{n, \psi}$ matrices

The entries of matrices $\mathbf{f}_{\mathbf{n}, \mathbf{p}}, \mathbf{f}_{\mathbf{n}, \mathbf{n}}$ and $\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}}$ are derived with the use of (A.23). The same matrix index notation from previous section is followed here.

$$
\begin{equation*}
\mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i-1)=\frac{\partial f_{n}(i)}{\partial p_{i-1}}=0 \tag{A.26a}
\end{equation*}
$$

$$
\begin{align*}
& \mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i)=\frac{\partial f_{n}(i)}{\partial p_{i}}=-\frac{\partial \mathcal{R}_{i}}{\partial p_{i}}  \tag{A.26b}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i+1)=\frac{\partial f_{n}(i)}{\partial p_{i+1}}=0  \tag{A.26c}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i-1)=\frac{\partial f_{n}(i)}{\partial n_{i-1}}=\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{n}}{2} \frac{-\Delta \psi_{i-1 / 2}}{h_{i-1}}+D_{n} \frac{1}{h_{i-1}}\right]  \tag{A.26d}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i)=\frac{\partial f_{n}(i)}{\partial n_{i}} \\
&=\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{n}}{2}\left(\frac{-\Delta \psi_{i+1 / 2}}{h_{i}}-\frac{-\Delta \psi_{i-1 / 2}}{h_{i-1}}\right)\right. \\
&\left.\mathbf{- D}_{n}\left(\frac{1}{h_{i-1}}+\frac{1}{h_{i}}\right)\right]-\frac{\partial \mathcal{R}_{i}}{\partial n}(i)  \tag{A.26e}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i+1)=\frac{\partial f_{n}(i)}{\partial n_{i+1}}=\frac{2}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mu_{n}}{2} \frac{-\Delta \psi_{i+1 / 2}}{h_{i}}+D_{n} \frac{1}{h_{i}}\right]  \tag{A.26f}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{\psi}}(i, i-1)=\frac{\partial f_{n}(i)}{\partial \psi_{i-1}}=-\frac{\mu_{n}}{\left(h_{i-1}+h_{i}\right)} \frac{\left(n_{i-1}+n_{i}\right)}{h_{i-1}}  \tag{A.26g}\\
& \mathbf{f}_{\mathbf{n}, \mathbf{\psi}}(i, i)=\frac{\partial f_{n}(i)}{\partial \psi_{i}}=\frac{\mu_{n}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{n_{i-1}+n_{i}}{h_{i-1}}+\frac{n_{i}+n_{i+1}}{h_{i}}\right]  \tag{A.26h}\\
& \mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}}(i, i+1)=\frac{\partial f_{n}(i)}{\partial \psi_{i+1}}=-\frac{\mu_{n}}{\left(h_{i-1}+h_{i}\right)} \frac{\left(n_{i}+n_{i+1}\right)}{h_{i}} \tag{A.26i}
\end{align*}
$$

## A.2.3 Scharfetter-Gummel method

The equations used for the Jacobian derivations is given in (A.27). It is necessary to discuss the derivatives of the repeated terms before the derivation of Jacobian elements. Then, compact equations can be derived with definition of notations for these repeated terms.

$$
\begin{align*}
f_{p}(i)= & \frac{2 D_{p}}{h_{i}+h_{i-1}}\left[\frac{\mathcal{B}\left(-z_{i+1 / 2}\right)}{h_{i}} p_{i+1}\right. \\
& \left.-\left(\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}}+\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}}\right) p_{i}+\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}} p_{i-1}\right] \\
& -\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} \tag{A.27a}
\end{align*}
$$

$$
\begin{align*}
f_{n}(i)= & \frac{2 D_{n}}{h_{i}+h_{i-1}}\left[\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}} n_{i+1}\right. \\
& \left.-\left(\frac{\mathcal{B}\left(-z_{i+1 / 2}\right)}{h_{i}}+\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}}\right) n_{i}+\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}} n_{i-1}\right] \\
& -\frac{n_{i} p_{i}-n_{i e}^{2}}{\tau_{p}\left(n_{i}+n_{1}\right)+\tau_{n}\left(p_{i}+p_{1}\right)} \tag{A.27b}
\end{align*}
$$

The term to be considered is the derivative of the Bernoulli function $\mathcal{B}\left(z_{i \pm 1 / 2}\right)$ with respect to $\psi_{i-1}, \psi_{i}$ and $\psi_{i+1}$. The variable $z_{i \pm 1 / 2}$ is defined in (A.28).

$$
\begin{align*}
& z_{i+1 / 2}=\frac{\Delta \psi_{i+1 / 2}}{V_{T}}=\frac{\psi_{i+1}-\psi_{i}}{V_{T}}  \tag{A.28a}\\
& z_{i-1 / 2}=\frac{\Delta \psi_{i+1 / 2}}{V_{T}}=\frac{\psi_{i}-\psi_{i-1}}{V_{T}} \tag{A.28b}
\end{align*}
$$

Using (A.28) the following derivatives can be derived.

$$
\begin{align*}
& \frac{\partial z_{i \pm 1 / 2}}{\partial \psi_{i+1}}= \pm \frac{1}{V_{T}}  \tag{A.29a}\\
& \frac{\partial z_{i \pm 1 / 2}}{\partial \psi_{i}}=\mp \frac{1}{V_{T}} \tag{A.29b}
\end{align*}
$$

The definition of $z_{i \pm 1 / 2}$ and its derivative is used to express the derivatives of the Bernoulli function $\left(\mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)\right)$ with respect to $\psi_{i-1}, \psi_{i}$ and $\psi_{i+1}$ as follows.

$$
\begin{align*}
\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial z_{i \pm 1 / 2}} & =\frac{ \pm 1}{\exp \left( \pm z_{i \pm 1 / 2}\right)-1}-\frac{z_{i \pm 1 / 2} \exp \left( \pm z_{i \pm 1 / 2}\right)}{\left(\exp \left( \pm z_{i \pm 1 / 2}\right)-1\right)^{2}}  \tag{A.30a}\\
\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial \psi_{i}} & =\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial z_{i \pm 1 / 2}} \frac{ \pm \partial z_{i \pm 1 / 2}}{\partial \psi_{i}} \\
& =\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial z_{i \pm 1 / 2}}\left(\frac{1}{V_{T}}\right)  \tag{A.30b}\\
\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial \psi_{i+1}} & =\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial z_{i \pm 1 / 2}} \frac{ \pm \partial z_{i \pm 1 / 2}}{\partial \psi_{i+1}} \\
& =\frac{\partial \mathcal{B}\left( \pm z_{i \pm 1 / 2}\right)}{\partial z_{i \pm 1 / 2}}\left(\frac{-1}{V_{T}}\right) \tag{A.30c}
\end{align*}
$$

## A.2.3.1 $f_{p, p}, f_{p, n}, f_{p, \psi}$ matrices

The entries of matrices $\mathbf{f}_{\mathbf{p}, \mathbf{p}}, \mathbf{f}_{\mathbf{p}, \mathbf{n}}$ and $\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}}$ are derived with the help of (A.27), (A.30) and previously defined index notations.

$$
\begin{align*}
\mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i-1) & =\frac{\partial f_{p}(i)}{\partial p_{i-1}}=\frac{2 D_{p}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}}\right]  \tag{A.31a}\\
\mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i) & =\frac{\partial f_{p}(i)}{\partial p_{i}} \\
& =-\frac{2 D_{p}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}}+\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}}\right]-\frac{\partial \mathcal{R}_{i}}{\partial p_{i}}  \tag{A.31b}\\
\mathbf{f}_{\mathbf{p}, \mathbf{p}}(i, i+1) & =\frac{\partial f_{p}(i)}{\partial p_{i+1}}=\frac{2 D_{p}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}}\right]  \tag{A.31c}\\
\mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i-1) & =\frac{\partial f_{p}(i)}{\partial n_{i-1}}=0  \tag{A.31d}\\
\mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i) & =\frac{\partial f_{p}(i)}{\partial n_{i}}=-\frac{\partial \mathcal{R}_{i}}{\partial n_{i}}  \tag{A.31e}\\
\mathbf{f}_{\mathbf{p}, \mathbf{n}}(i, i+1) & =\frac{\partial f_{p}(i)}{\partial n_{i+1}}=0  \tag{A.31f}\\
\mathbf{f}_{\mathbf{p}, \mathbf{\psi}}(i, i-1) & =\frac{\partial f_{p}(i)}{\partial \psi_{i-1}} \\
& \frac{2 D_{p}}{\left(h_{i-1}+h_{i}\right)}\left[-\frac{p_{i-1}}{h_{i-1}} \frac{\partial \mathcal{B}\left(z_{i-1 / 2}\right)}{\partial \psi_{i-1}}+\frac{p_{i-1}}{h_{i-1}} \frac{\partial \mathcal{B}\left(-z_{i-1 / 2}\right)}{\partial \psi_{i-1}}\right]  \tag{A.31g}\\
\mathbf{f}_{\mathbf{p}, \psi}(i, i) & =\frac{\partial f_{p}(i)}{\partial \psi_{i}}=-\left[\frac{\partial f_{p}(i)}{\partial \psi_{i-1}}+\frac{\partial f_{p}(i)}{\partial \psi_{i+1}}\right]  \tag{A.31h}\\
\mathbf{f}_{\mathbf{p}, \mathbf{\psi}}(i, i+1) & =\frac{\partial f_{p}(i)}{\partial \psi_{i+1}} \\
& =\frac{2 D_{p}}{\left(h_{i-1}+h_{i}\right)}\left[-\frac{p_{i}}{h_{i}} \frac{\partial \mathcal{B}\left(-z_{i+1 / 2}\right)}{\partial \psi_{i+1}}+\frac{p_{i+1}}{h_{i}} \frac{\partial \mathcal{B}\left(z_{i+1 / 2}\right)}{\partial \psi_{i+1}}\right] \tag{A.31i}
\end{align*}
$$

## A.2.3.2 $f_{n, p}, f_{n, n}, f_{n, \psi}$ matrices

The entries of matrices $\mathbf{f}_{\mathbf{n}, \mathbf{p}}, \mathbf{f}_{\mathbf{n}, \mathbf{n}}$ and $\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}}$ are derived using (A.27), (A.30) and usual matrix index notation.

$$
\begin{equation*}
\mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i-1)=\frac{\partial f_{n}(i)}{\partial p_{i-1}}=0 \tag{A.32a}
\end{equation*}
$$

$$
\begin{align*}
\mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i) & =\frac{\partial f_{n}(i)}{\partial p_{i}}=-\frac{\partial \mathcal{R}_{i}}{\partial p_{i}}  \tag{A.32b}\\
\mathbf{f}_{\mathbf{n}, \mathbf{p}}(i, i+1) & =\frac{\partial f_{n}(i)}{\partial p_{i+1}}=0  \tag{A.32c}\\
\mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i-1) & =\frac{\partial f_{n}(i)}{\partial n_{i-1}}=\frac{2 D_{n}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(-z_{i-1 / 2}\right)}{h_{i-1}}\right]  \tag{A.32d}\\
\mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i) & =\frac{\partial f_{n}(i)}{\partial n_{i}} \\
& =-\frac{2 D_{n}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(z_{i-1 / 2}\right)}{h_{i-1}}+\frac{\mathcal{B}\left(-z_{i+1 / 2}\right)}{h_{i}}\right]-\frac{\partial \mathcal{R}_{i}}{\partial n_{i}}  \tag{A.32e}\\
& =\frac{2 D_{n}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{n_{i-1}}{h_{i-1}} \frac{\partial \mathcal{B}\left(-z_{i-1 / 2}\right)}{\partial \psi_{i-1}}-\frac{n_{i}}{h_{i-1}} \frac{\partial \mathcal{B}\left(z_{i-1 / 2}\right)}{\partial \psi_{i-1}}\right]  \tag{A.32f}\\
\mathbf{f}_{\mathbf{n}, \mathbf{n}}(i, i+1) & =\frac{\partial f_{n}(i)}{\partial n_{i+1}}=\frac{2 D_{n}}{\left(h_{i-1}+h_{i}\right)}\left[\frac{\mathcal{B}\left(z_{i+1 / 2}\right)}{h_{i}}\right] \\
\mathbf{f}_{\mathbf{n}, \mathbf{\psi}}(i, i-1) & =\frac{\partial f_{n}(i)}{\partial \psi_{i-1}}  \tag{A.32g}\\
\mathbf{f}_{\mathbf{n}, \mathbf{\psi}}(i, i) & =\frac{\partial f_{n}(i)}{\partial \psi_{i}}=-\left[\frac{\partial f_{n}(i)}{\partial \psi_{i-1}}+\frac{\partial f_{n}(i)}{\partial \psi_{i+1}}\right]  \tag{A.32h}\\
\mathbf{f}_{\mathbf{n}, \mathbf{\psi}}(i, i+1) & =\frac{\partial f_{n}(i)}{\partial \psi_{i+1}} \\
& =\frac{2 D_{n}}{\left(h_{i-1}+h_{i}\right)}\left[-\frac{n_{i}}{h_{i}} \frac{\partial \mathcal{B}\left(-z_{i+1 / 2}\right)}{\partial \psi_{i+1}}+\frac{n_{i+1}}{h_{i}} \frac{\partial \mathcal{B}\left(-z_{i+1 / 2}\right)}{\partial \psi_{i+1}}\right] \tag{A.32i}
\end{align*}
$$

## A.2.4 Newton-Raphson iterator

The Newton-Raphson iterator in (A.16) is modified with (A.22) to

$$
\begin{align*}
& {\left[\begin{array}{ccc}
\mathbf{I}-\Delta t^{m} \mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\Delta t^{m} \mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\Delta t^{m} \mathbf{f}_{\mathbf{p}, \psi} \\
-\Delta t^{m} \mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I}-\Delta t^{m} \mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\Delta t^{m} \mathbf{f}_{\mathbf{n}, \psi} \\
-\mathbf{D} & \mathbf{D} & \mathbf{S}
\end{array}\right]\left[\begin{array}{l}
\mathbf{p}^{m, k+1} \\
\mathbf{n}^{m, k+1} \\
\boldsymbol{\psi}^{m, k+1}
\end{array}\right]}  \tag{A.33}\\
& =\left[\begin{array}{c}
\mathbf{p}^{m, k}+\Delta t^{m}\left(\mathbf{f}_{\mathbf{p}}{ }^{m, k}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} \mathbf{p}^{m, k}-\mathbf{f}_{\mathbf{p}, \mathbf{n}} \mathbf{n}^{m, k}-\mathbf{f}_{\mathbf{p}, \psi} \mathbf{\psi}^{m, k}\right) \\
\mathbf{n}^{m, k}+\Delta t^{m}\left(\mathbf{f}_{\mathbf{n}}{ }^{m, k}-\mathbf{f}_{\mathbf{n}, \mathbf{p}} \mathbf{p}^{m, k}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} \mathbf{n}^{m, k}-\mathbf{f}_{\mathbf{n}, \psi} \mathbf{\psi}^{m, k}\right) \\
\mathbf{D}\left(\mathbf{N}_{D}^{+}-\mathbf{N}_{A}^{-}\right)+\mathbf{b}\left(v_{a}, v_{c}\right)
\end{array}\right]
\end{align*}
$$

Either (3.31) or (A.33) can be used in code implementation. In former the small change in solution between time steps are computed while in the latter the next step solution is computed directly.

## A.2.5 $f_{i, \chi}$ matrices

The Jacobian partitions considered here for derivation involves the coupling of terminal currents and the semiconductor carrier-charge densities at its boundaries (ohmic contacts). The diode terminal current equations are augmented to CCT equations for device model encapsulation. Jacobian contributions derived in this subsection is due to this augmentation of diode terminal current equations. The partitions are different for central difference and Scharfetter-Gummel methods. These contributions are derived here from the terminal current equations for these two discretization methods.

## A.2.5.1 Central difference method

The terminal current equations (3.34) is used in the following derivations.

$$
\begin{align*}
\mathbf{f}_{\mathbf{i}, \mathbf{p}}(1,1) & =\frac{\partial i_{a}}{\partial p_{1}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{1}-\psi_{2}}{h_{1}}-\frac{D_{p}}{h_{1}}\right]  \tag{A.34a}\\
\mathbf{f}_{\mathbf{i}, \mathbf{p}}(1,2) & =\frac{\partial i_{a}}{\partial p_{2}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{1}-\psi_{2}}{h_{1}}+\frac{D_{p}}{h_{1}}\right]  \tag{A.34b}\\
\mathbf{f}_{\mathbf{i}, \mathbf{p}}(2, N-1) & =\frac{\partial i_{c}}{\partial p_{N-1}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{N-1}-\psi_{N}}{h_{N-1}}+\frac{D_{p}}{h_{N-1}}\right]  \tag{A.34c}\\
\mathbf{f}_{\mathbf{i}, \mathbf{p}}(2, N) & =\frac{\partial i_{c}}{\partial p_{N}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{N-1}-\psi_{N}}{h_{N-1}}-\frac{D_{p}}{h_{N-1}}\right]  \tag{A.34d}\\
\mathbf{f}_{\mathbf{i}, \mathbf{n}}(1,1) & =\frac{\partial i_{a}}{\partial n_{1}}
\end{align*}
$$

$$
\left.\begin{array}{rl} 
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{1}-\psi_{2}}{h_{1}}-\frac{D_{n}}{h_{1}}\right] \\
\mathbf{f}_{\mathbf{i}, \mathbf{n}}(1,2) & =\frac{\partial i_{a}}{\partial n_{2}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{1}-\psi_{2}}{h_{1}}+\frac{D_{n}}{h_{1}}\right] \\
\mathbf{f}_{\mathbf{i}, \mathbf{n}}(2, N-1) & =\frac{\partial i_{c}}{\partial n_{N-1}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{N-1}-\psi_{N}}{h_{N-1}}-\frac{D_{n}}{h_{N-1}}\right] \\
& =q_{e} N_{M} a_{c}\left[\frac{\psi_{N-1}-\psi_{N}}{h_{N-1}}+\frac{D_{n}}{h_{N-1}}\right] \\
\mathbf{f}_{\mathbf{i}, \mathbf{n}}(2, N) & =\frac{\partial i_{c}}{\partial n_{N}} \\
\mathbf{f}_{\mathbf{i}, \psi}(1,1) & =\frac{\partial i_{a}}{\partial \psi_{1}} \\
& =-q_{e} N_{M} a_{c}\left[\frac{\mu_{p}\left(p_{1}+p_{2}\right)}{2 h_{1}}+\frac{\mu_{n}\left(n_{1}+n_{2}\right)}{2 h_{1}}\right] \\
\mathbf{f}_{\mathbf{i}, \psi}(1,2) & =\frac{\partial i_{a}}{\partial \psi_{2}} \\
& =q_{e} N_{M} a_{c}\left[\frac{\mu_{p}\left(p_{1}+p_{2}\right)}{2 h_{1}}+\frac{\mu_{n}\left(n_{1}+n_{2}\right)}{2 h_{1}}\right] \\
& =-q_{e} N_{M} a_{c}\left[\frac{\mu_{p}\left(p_{N-1}+p_{N}\right)}{2 h_{N-1}}+\frac{\mu_{n}\left(n_{N-1}+n_{N}\right)}{2 h_{N-1}}\right] \\
\mathbf{f}_{\mathbf{i}, \psi}(2, N-1) & \frac{\partial i_{c}}{\partial \psi_{N-1}} \\
\mathbf{f}_{\mathbf{i}, \psi}(2, N) & =\frac{\partial i_{c}}{\partial \psi_{N}} \\
2 h_{N-1} \tag{A.34l}
\end{array}\right]
$$

## A.2.5.2 Scharfetter-Gummel method

The terminal current equations (3.35) is used in the following derivations along with the Bernoulli function derivatives (A.30).

$$
\begin{align*}
\mathbf{f}_{\mathbf{i}, \mathbf{p}}(1,1) & =\frac{\partial i_{a}}{\partial p_{1}} \\
& =\frac{q_{e} N_{M} a_{c} D_{p}}{h_{1}} \mathcal{B}\left(z_{3 / 2}\right) \tag{A.35a}
\end{align*}
$$

$$
\begin{align*}
& \mathbf{f}_{\mathbf{i}, \mathbf{p}}(1,2)=\frac{\partial i_{a}}{\partial p_{2}} \\
& =-\frac{q_{e} N_{M} a_{c} D_{p}}{h_{1}} \mathcal{B}\left(-z_{3 / 2}\right)  \tag{A.35b}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{p}}(2, N-1)=\frac{\partial i_{c}}{\partial p_{N-1}} \\
& =\frac{q_{e} N_{M} a_{c} D_{p}}{h_{N-1}} \mathcal{B}\left(z_{N-1 / 2}\right)  \tag{A.35c}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{p}}(2, N)=\frac{\partial i_{c}}{\partial p_{N}} \\
& =-\frac{q_{e} N_{M} a_{c} D_{p}}{h_{N-1}} \mathcal{B}\left(-z_{N-1 / 2}\right)  \tag{A.35d}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{n}}(1,1)=\frac{\partial i_{a}}{\partial n_{1}} \\
& =-\frac{q_{e} N_{M} a_{c} D_{n}}{h_{1}} \mathcal{B}\left(-z_{3 / 2}\right)  \tag{A.35e}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{n}}(1,2)=\frac{\partial i_{a}}{\partial n_{2}} \\
& =\frac{q_{e} N_{M} a_{c} D_{n}}{h_{1}} \mathcal{B}\left(z_{3 / 2}\right)  \tag{A.35f}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{n}}(2, N-1)=\frac{\partial i_{c}}{\partial n_{N-1}} \\
& =-\frac{q_{e} N_{M} a_{c} D_{p}}{h_{N-1}} \mathcal{B}\left(-z_{N-1 / 2}\right)  \tag{A.35~g}\\
& \mathbf{f}_{\mathbf{i}, \mathbf{n}}(2, N)=\frac{\partial i_{c}}{\partial n_{N}} \\
& =\frac{q_{e} N_{M} a_{c} D_{p}}{h_{N-1}} \mathcal{B}\left(z_{N-1 / 2}\right)  \tag{A.35h}\\
& \mathbf{f}_{\mathbf{i}, \psi}(1,1)=\frac{\partial i_{a}}{\partial \psi_{1}} \\
& =\frac{q_{e} N_{M} a_{c}}{h_{1}}\left[\left(D_{p} p_{1}+D_{n} n_{2}\right) \frac{\partial \mathcal{B}\left(z_{3 / 2}\right)}{\partial \psi_{1}}\right. \\
& \left.-\left(D_{p} p_{2}+D_{n} n_{1}\right) \frac{\partial \mathcal{B}\left(-z_{3 / 2}\right)}{\partial \psi_{1}}\right]  \tag{A.35i}\\
& \mathbf{f}_{\mathbf{i}, \psi}(1,2)=\frac{\partial i_{a}}{\partial \psi_{2}} \\
& =\frac{q_{e} N_{M} a_{c}}{h_{1}}\left[\left(D_{p} p_{1}+D_{n} n_{2}\right) \frac{\partial \mathcal{B}\left(z_{3 / 2}\right)}{\partial \psi_{2}}\right. \\
& \left.-\left(D_{p} p_{2}+D_{n} n_{1}\right) \frac{\partial \mathcal{B}\left(-z_{3 / 2}\right)}{\partial \psi_{2}}\right] \tag{A.35j}
\end{align*}
$$

$$
\begin{align*}
\mathbf{f}_{\mathbf{i}, \boldsymbol{\psi}}(2, N-1)= & \frac{\partial i_{c}}{\partial \psi_{N-1}} \\
= & \frac{q_{e} N_{M} a_{c}}{h_{N-1}}\left[\left(D_{p} p_{N-1}+D_{n} n_{N}\right) \frac{\partial \mathcal{B}\left(z_{N-1 / 2}\right)}{\partial \psi_{N-1}}\right. \\
& \left.-\left(D_{p} p_{N}+D_{n} n_{N-1}\right) \frac{\partial \mathcal{B}\left(-z_{N-1 / 2}\right)}{\partial \psi_{N-1}}\right]  \tag{A.35k}\\
\mathbf{f}_{\mathbf{i}, \psi}(2, N)= & \frac{\partial i_{c}}{\partial \psi_{N}} \\
= & \frac{q_{e} N_{M} a_{c}}{h_{N-1}}\left[\left(D_{p} p_{N-1}+D_{n} n_{N}\right) \frac{\partial \mathcal{B}\left(z_{N-1 / 2}\right)}{\partial \psi_{N}}\right. \\
& \left.-\left(D_{p} p_{N}+D_{n} n_{N-1}\right) \frac{\partial \mathcal{B}\left(-z_{N-1 / 2}\right)}{\partial \psi_{N}}\right] \tag{A.351}
\end{align*}
$$

## B. MANUAL TABLEAU DERIVATION



Fig. B.1.: A simple $R L$ and diode circuit.

The state-space model of the circuit shown in Figure B. 1 is derived by writing KCL and KVL equations for the switch-on condition. The voltage across the capacitors $C_{a}$ and $C_{c}$ are $v_{a}$ and $v_{c}$, respectively. The single KVL, two KCL, and diode voltage equations are written in state-space form as follows,

$$
\begin{align*}
\frac{d}{d t}\left[\begin{array}{l}
i_{L} \\
v_{a} \\
v_{c}
\end{array}\right] & =\left[\begin{array}{ccc}
-R / L & -1 / L & 1 / L \\
1 / C_{a} & 0 & 0 \\
-1 / C_{c} & 0 & 0
\end{array}\right]\left[\begin{array}{l}
i_{L} \\
v_{a} \\
v_{c}
\end{array}\right]+\left[\begin{array}{cc}
0 & 0 \\
-1 / C_{a} & 0 \\
0 & 1 / C_{c}
\end{array}\right]\left[\begin{array}{c}
i_{a} \\
i_{c}
\end{array}\right]+\left[\begin{array}{c}
1 / L \\
0 \\
0
\end{array}\right] v_{s}  \tag{B.1a}\\
v_{d} & =\left[\begin{array}{lll}
0 & 1 & -1
\end{array}\right]\left[\begin{array}{l}
i_{L} \\
v_{a} \\
v_{c}
\end{array}\right] . \tag{B.1b}
\end{align*}
$$

A compact representation of the equations above is given in the following set of equations.

$$
\begin{align*}
\dot{\mathbf{y}} & =\mathbf{A} \mathbf{y}+\mathbf{B}_{\mathbf{1}} \mathbf{i}_{d}+\mathbf{B}_{\mathbf{2}} v_{s}=\mathbf{f}_{\mathbf{y}}  \tag{B.2a}\\
v_{d} & =\mathbf{C y} \tag{B.2b}
\end{align*}
$$

where,

$$
\begin{align*}
\mathbf{y} & =\left[\begin{array}{lll}
i_{L} & v_{a} & v_{c}
\end{array}\right]^{T}  \tag{B.3a}\\
\mathbf{i}_{d} & =\left[\begin{array}{ll}
i_{a} & i_{c}
\end{array}\right]^{T}  \tag{B.3b}\\
\mathbf{A} & =\left[\begin{array}{ccc}
-R / L & -1 / L & 1 / L \\
1 / C_{a} & 0 & 0 \\
-1 / C_{c} & 0 & 0
\end{array}\right]  \tag{B.3c}\\
\mathbf{B}_{\mathbf{1}} & =\left[\begin{array}{cc}
0 & 0 \\
-1 / C_{a} & 0 \\
0 & 1 / C_{c}
\end{array}\right]  \tag{B.3d}\\
\mathbf{B}_{\mathbf{2}} & =\left[\begin{array}{lll}
1 / L & 0 & 0
\end{array}\right]^{T}  \tag{B.3e}\\
\mathbf{C} & =\left[\begin{array}{lll}
0 & 1 & -1
\end{array}\right] \tag{B.3f}
\end{align*}
$$

The diode terminal currents are given below.

$$
\begin{align*}
\mathbf{i}_{d} & =\left[\begin{array}{c}
i_{a} \\
i_{c}
\end{array}\right]=\left[\begin{array}{c}
f\left(p_{1}, p_{2}, n_{1}, n_{2}, \psi_{1}, \psi_{2}\right) \\
f\left(p_{N-1}, p_{N}, n_{N-1}, n_{N-1}, \psi_{N-1}, \psi_{N}\right)
\end{array}\right] \\
& =\mathbf{f}_{\mathbf{i}_{d}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) \tag{B.4}
\end{align*}
$$

Equations (B.1) and (B.4) are augmented to the system of equations in (3.29). The augmented system of equations below is further supplemented with boundary conditions.

$$
\begin{align*}
\frac{d \mathbf{p}}{d t} & =\mathbf{f}_{\mathbf{p}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{B.5a}\\
\frac{d \mathbf{n}}{d t} & =\mathbf{f}_{\mathbf{n}}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}) \tag{B.5b}
\end{align*}
$$

$$
\begin{align*}
\mathbf{0} & =\mathbf{S} \boldsymbol{\psi}-\mathbf{D}\left(\mathbf{p}-\mathbf{n}+\mathbf{N}_{D}^{+}-\boldsymbol{N}_{A}^{-}\right)-\mathbf{b}\left(v_{d}\right)=\mathbf{f}_{\psi}(\mathbf{p}, \mathbf{n}, \boldsymbol{\psi})  \tag{B.5c}\\
\frac{d \mathbf{y}}{d t} & =\mathbf{f}_{\mathbf{y}}\left(\mathbf{y}, \mathbf{i}_{d}, v_{s}\right)  \tag{B.5d}\\
\mathbf{0} & =\mathbf{i}_{d}-\mathbf{f}_{\mathbf{i}_{d}}  \tag{B.5e}\\
0 & =v_{d}-\mathbf{C} \mathbf{y} \tag{B.5f}
\end{align*}
$$

with boundary conditions

$$
\begin{align*}
& n\left(x_{c}\right)=\frac{N_{D}^{+}}{2}+\sqrt{\left(\frac{N_{D}^{+}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{D}^{+}  \tag{B.6a}\\
& p\left(x_{c}\right)=\frac{n_{i e}^{2}}{N_{D}^{+}}  \tag{B.6b}\\
& \psi\left(x_{c}\right)=\frac{k T}{q} \ln \left(\frac{n\left(x_{c}\right)}{n_{i e}}\right)  \tag{B.6c}\\
& n\left(x_{a}\right)=\frac{n_{i e}^{2}}{N_{A}^{-}}  \tag{B.6d}\\
& p\left(x_{a}\right)=\frac{N_{A}^{-}}{2}+\sqrt{\left(\frac{N_{A}^{-}}{2}\right)^{2}+n_{i e}^{2}} \approx N_{A}^{-}  \tag{B.6e}\\
& \psi\left(x_{a}\right)=v_{d}-\frac{k T}{q} \ln \left(\frac{p\left(x_{a}\right)}{n_{i e}}\right) \tag{B.6f}
\end{align*}
$$

The boundary condition $\psi\left(x_{a}\right)$ depends on the voltage across the diode $v_{d}$, which is, in turn, a function of the circuit states. The external circuit dynamics gives voltage as input to the diode model. The terminal currents $i_{a}$ and $i_{c}$ are algebraic functions of $\{\mathbf{p}, \mathbf{n}, \boldsymbol{\psi}\}$, which are "current like" inputs to the external circuit model. It can be inferred that the diode model is a voltage-in current-out model.

Applying backward Euler algorithm to (B.5) gives a fully discretized system of equations. Due to the nonlinearities, an iterative approach such as the NewtonRaphson method is needed to solve these equations. Jacobian partitions related to the charge densities and electric potential are as derived in Appendix A.2. These partitions do not change when coupled with external circuits. The Jacobian partitions involving the variables $\mathbf{y}, \mathbf{i}_{d}$ and $v_{d}$ are denoted $\mathbf{f}_{\mathbf{y}, \mathbf{y}}, \mathbf{f}_{\mathbf{i}_{d}, \mathbf{i}_{d}}$ and $\mathbf{f}_{v_{d}, v_{d}}$, respectively. The last two of these are identity matrices of appropriate dimensions.

The circuit-device coupling is between the variable pairs $\left(\boldsymbol{\psi}, v_{d}\right),\left(\mathbf{i}_{d}, \mathbf{p}\right),\left(\mathbf{i}_{d}, \mathbf{n}\right)$, $\left(\mathbf{i}_{d}, \boldsymbol{\psi}\right)$ and $\left(\mathbf{i}_{d}, \mathbf{y}\right)$. The circuit-device coupling Jacobian partitions composed of these variable pairs are $\mathbf{f}_{\psi, v_{d}}, \mathbf{f}_{\mathbf{i}_{d}, \mathbf{p}}, \mathbf{f}_{\mathbf{i}_{d}, \mathbf{n}}, \mathbf{f}_{\mathbf{i}_{d}, \boldsymbol{\psi}}$ and $\mathbf{f}_{\mathbf{i}_{d}, \mathbf{y}}$ respectively. A Taylor's series expansion of the equations involving these variables about $\chi^{m}$ up to first-order terms, extending the notations (3.32) to circuit variables, and equating the system zero gives the following system of equations.

$$
\begin{align*}
\left(\mathbf{p}^{m, k}-\mathbf{p}^{m-1, \infty}\right) / \Delta t^{m}-\mathbf{f}_{\mathbf{p}}+\left(\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{p}, \mathbf{p}}\right) \Delta \mathbf{p}-\mathbf{f}_{\mathbf{p}, \mathbf{n}} \Delta \mathbf{n}-\mathbf{f}_{\mathbf{p}, \psi} \Delta \boldsymbol{\psi} & =\mathbf{0}  \tag{B.7a}\\
\left(\mathbf{n}^{m, k}-\mathbf{n}^{m-1, \infty}\right) / \Delta t^{m}-\mathbf{f}_{\mathbf{n}}+\left(\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{n}, \mathbf{n}}\right) \Delta \mathbf{n}-\mathbf{f}_{\mathbf{n}, \mathbf{p}} \Delta \mathbf{p}-\mathbf{f}_{\mathbf{n}, \psi} \Delta \boldsymbol{\psi} & =\mathbf{0}  \tag{B.7b}\\
\mathbf{f}_{\psi}+\mathbf{f}_{\psi, \mathbf{p}} \Delta \mathbf{p}+\mathbf{f}_{\psi, \mathbf{n}} \Delta \mathbf{n}+\mathbf{f}_{\psi, \psi} \Delta \boldsymbol{\psi}+\mathbf{f}_{\psi, v_{d}} \Delta v_{d} & =\mathbf{0}  \tag{B.7c}\\
\mathbf{i}^{m, k}-\mathbf{f}_{\mathbf{i}_{d}}+\mathbf{I} \Delta \mathbf{i}_{d}+\mathbf{f}_{\mathbf{i}_{d}, \mathbf{p}} \Delta \mathbf{p}+\mathbf{f}_{\mathbf{i}_{d, \mathbf{n}}} \Delta \mathbf{n}+\mathbf{f}_{\mathbf{i}_{d}, \psi} \Delta \boldsymbol{\psi} & =\mathbf{0}  \tag{B.7d}\\
\left(\mathbf{y}^{m, k}-\mathbf{y}^{m-1, \infty}\right) / \Delta t^{m}-\mathbf{f}_{\mathbf{y}}+\left(\mathbf{I}-\mathbf{f}_{\mathbf{y}, \mathbf{y}}\right) \Delta \mathbf{y}-\mathbf{f}_{\mathbf{y}, \mathbf{i}_{d}} \Delta \mathbf{i}_{d} & =\mathbf{0}  \tag{B.7e}\\
v_{d}^{m, k}-\mathbf{C y} \mathbf{y}^{m, k}+\Delta v_{d}-\mathbf{C} \Delta \mathbf{y} & =0 \tag{B.7f}
\end{align*}
$$

The additional Jacobian partitions that needs to be derived are $\mathbf{f}_{\mathbf{i}_{d}, \mathbf{p}}, \mathbf{f}_{\mathbf{i}_{d}, \mathbf{n}}, \mathbf{f}_{\mathbf{i}_{d}, \psi}$, $\mathbf{f}_{\mathbf{y}, \mathbf{y}}, \mathbf{f}_{\mathbf{y}, \mathbf{i}_{d}}$ and $\mathbf{f}_{\psi, v_{d}}$. The last three partitions are the same for both central difference and Scharfetter-Gummel discretization methods. They are derived in the following equations.

$$
\begin{align*}
\mathbf{f}_{\mathbf{y}, \mathbf{y}} & =\frac{\partial}{\partial \mathbf{y}}\left(\mathbf{A} \mathbf{y}+\mathbf{B}_{1} \mathbf{i}_{d}+\mathbf{B}_{\mathbf{2}} v_{s}\right) \\
& =\mathbf{A}  \tag{B.8a}\\
\mathbf{f}_{\mathbf{y}, \mathbf{i}_{d}} & =\frac{\partial}{\partial \mathbf{i}}\left(\mathbf{A} \mathbf{y}+\mathbf{B}_{1} \mathbf{i}_{d}+\mathbf{B}_{\mathbf{2}} v_{s}\right) \\
& =\mathbf{B}_{1}  \tag{B.8b}\\
\mathbf{f}_{\psi, v_{d}} & =\frac{\partial}{\partial v_{d}} \mathbf{f}_{\psi} \\
& =\left[\begin{array}{llll}
1 & 0 & \ldots & 0
\end{array}\right]^{T} \tag{B.8c}
\end{align*}
$$

The remaining Jacobian partitions rely on the total terminal current formulae (B.4). These formulae differ depending upon the spatial discretization method used. The terminal current formulae due to central difference and Scharfetter-Gummel methods are given in (3.34) and (3.35) respectively.

The entries of each partition is derived in Appendix A.2. After some manipulation, all the partitions derived for both the semiconductor and circuit is used to assemble the Newton-Raphson iterator (B.9). The horizontal and vertical lines in the matrix and vectors show a partitioning of this iterator to form a block partitioned linear system. The first diagonal block corresponds to the Jacobian partitions of the diode alone while the second diagonal block corresponds to that of the external circuit. The super-diagonal block describes the device-circuit coupling and the sub-diagonal block the circuit-device coupling. This partitioning of the iterator is useful to efficiently implement the variable-structure strategy proposed in Chapter 7.

$$
\begin{align*}
& {\left[\begin{array}{ccc|ccc}
\mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{p}, \mathbf{p}} & -\mathbf{f}_{\mathbf{p}, \mathbf{n}} & -\mathbf{f}_{\mathbf{p}, \boldsymbol{\psi}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
-\mathbf{f}_{\mathbf{n}, \mathbf{p}} & \mathbf{I} / \Delta t^{m}-\mathbf{f}_{\mathbf{n}, \mathbf{n}} & -\mathbf{f}_{\mathbf{n}, \boldsymbol{\psi}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{f}_{\psi, \mathbf{p}} & \mathbf{f}_{\psi, \mathbf{n}} & \mathbf{f}_{\psi, \boldsymbol{\psi}} & \mathbf{0} & \mathbf{0} & \mathbf{f}_{\psi, v_{d}} \\
\hline \mathbf{f}_{\mathbf{i}_{d, \mathbf{p}}} & \mathbf{f}_{\mathbf{i}_{d, \mathbf{n}}} & \mathbf{f}_{\mathbf{i}_{d, \psi}} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{B}_{1} & \mathbf{I} / \Delta t^{m}-\mathbf{A} & \mathbf{0} \\
0 & 0 & 0 & 0 & -\mathbf{C} & 1
\end{array}\right]\left[\begin{array}{c}
\Delta \mathbf{p} \\
\Delta \mathbf{n} \\
\Delta \boldsymbol{\psi} \\
\hline \Delta \mathbf{i}_{d} \\
\Delta \mathbf{y} \\
\Delta v_{d}
\end{array}\right]}  \tag{B.9}\\
& =\left[\begin{array}{c}
-\left(\mathbf{p}^{m, k}-\mathbf{p}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{p}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\Psi}^{m, k}\right) \\
-\left(\mathbf{n}^{m, k}-\mathbf{n}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{n}}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\Psi}^{m, k}\right) \\
-\mathbf{f}_{\psi}\left(\mathbf{p}^{m, k}, \mathbf{n}^{m, k}, \boldsymbol{\Psi}^{m, k}, \boldsymbol{N}_{A}^{-}, \boldsymbol{N}_{D}^{+}\right) \\
\mathbf{0} \\
-\left(\mathbf{y}^{m, k}-\mathbf{y}^{m-1, \infty}\right) / \Delta t^{m}+\mathbf{f}_{\mathbf{y}}\left(\mathbf{y}^{m, k}, \mathbf{i}_{d}^{m, k}, v_{s}\right) \\
0
\end{array}\right]
\end{align*}
$$

## B. 1 Appendix Summary

It can be seen that manual derivation of tableau is tedious, needs to be repeated for each circuit considered and is error prone. Only a handful of inductor currents and capacitor voltages can be considered as states while others can be expressed algebraically using these chosen states. The book keeping effort to derive the state equations and border matrices in the iterator increases as the number of semiconduc-
tor devices in the circuit increases. An algorithmic way of building system equations set forth in Chapter 4.2 is an alternative to circumvent the aforementioned issues.

## C. ADAPTIVE SPATIAL MESH

The central difference and Scharfetter-Gummel methods used to discretize the CCT equations on a fixed grid can also be applied to spatial grids that evolve as time progresses. The results of a preliminary study using an adaptive spatial mesh is described in this section. In this study, the central difference method is used for spatial discretization. A strategy to measure or estimate the spatial error in the solution is described and used to add or delete node as the simulation progresses. The error estimator relies on the difference in the solutions for two meshes with different spatial resolutions at each time step.


Fig. C.1.: Charge densities computed on adaptive mesh.

A hierarchical mesh can be generated by uniformly bisecting the mesh throughout the spatial domain at time $t^{m-1}$. A bisected mesh in one dimension has a node inserted midway between two nodes of the coarser mesh. On the bisected mesh, the charge
densities are computed using linear interpolation. The electric potential is computed by solving the Poisson's equation on the finer mesh at $t^{m-1}$. All of the information needed to compute the solution at the next time step $t^{m}$ using the bisected mesh is available at this point. The spatial solutions using the coarse and finer meshes at $t^{m}$ are computed using the iterators previously derived. The solutions based upon the coarse mesh and the uniformly bisected mesh are used for error estimation. At the midpoint of each interval in the coarse mesh, there exist a node in the bisected mesh. A triangle is formed by the charge densities at these three nodes, two from coarse mesh and one from fine mesh. The error introduced by removing an inserted node is computed by scaling the area of this triangle appropriately. A small area implies that the charge density of the inserted node can be approximated as a linear interpolate of its neighboring nodes. Therefore, it is not needed. Thus, if the error exceeds a spatial error tolerance, that node is retained and, if not, the node is not needed and is therefore eliminated. The original coarse mesh is then replaced with the bisected mesh after eliminating unnecessary nodes. This process is repeated until the spatial mesh and corresponding solution converge at time $t^{m}$. Upon convergence, the solution at the given time step is accepted, and the simulation proceeds to the next time step.

For illustration purposes, a diode with a coarse mesh is discretized using the central difference method. The parameters of the diode are as in Table 3.1. The charge densities computed using an adaptive mesh for a -50 V bias is shown in Figure C.1. This strategy results in an adaptive mesh that tends to remove the nodes where the charge densities do not vary significantly, that is in the neutral and depletion regions. The electric potential, although an algebraic function of the charge densities, varies significantly within the depletion region. Consequently, the drift component of the current density dominates the diffusion component, especially when the diode is reverse biased. In finite volume methods, the discretizing schemes should satisfy transportiveness under such conditions.


Fig. C.2.: Number of nodes (\# Nodes) versus time-step index for adaptive mesh.

From (3.19) and the Péclet constraint $P_{e} \leq 2$ discussed in Section 3.3.1, it is clear that $\left|\psi_{i}-\psi_{i+1}\right| \leq 2 V_{T}$. Thus, the spatial discretization should also ensure that the difference in electrical potentials between the neighboring nodes is less than $2 V_{T}$ throughout the simulation. In order to satisfy this criterion, the depletion region should have more nodes even though they are not required to resolve the charge densities. The evolution of the number of nodes as the simulation progresses is shown in Figure C.2. The number of steady-state nodes is 112 and the simulation required approximately 258.28 s to simulate $2.5 \mu \mathrm{~s}$ of a diode-only model.

The adaptive feature is attractive but the computational overhead is significant, even though it provides guidance on where the nodes should be placed. The computational overhead of using adaptive meshes involves the repeated solutions computed on finer meshes. Moreover, as the mesh changes, the matrix-vector dimension of the iterative difference equations changes. As a result, the sparse tableau matrices have to be completely rebuilt and refactored prior to the solution of the difference equations, both of which are computationally expensive steps. Finally, since the adaptive mesh
targets the node placement in the regions of significant change in the charge densities, the time step control mechanism chooses smaller time steps to limit the temporal error in these nodes. All of these factors when put together results in a simulation that is much slower than a fixed-mesh simulation. A fast simulation using adaptive mesh needs a more effective linear-equation solver and an error estimator that does not rely on multiple solutions involving finer meshes. While such an approach may be possible, it was concluded that an appropriately selected fixed mesh will provide a computationally efficient and accurate solution for individual devices.

## C. 1 Appendix Summary

Computing accurate solutions on different meshes before converging to a solution on spatially adapted mesh is computationally wasteful. There is a class of a-posteriori error estimation techniques available that rely on a solution computed on the current mesh to indicate where the mesh needs to refined or coarsened. Formulating such a reliable error estimator for system of PDEs is an active area of research. Instead, a carefully chosen fixed mesh that is capable of resolving the solution with sufficient accuracy is recommended for this type of problem.

## D. COMPUTER SPECIFICATIONS

Engineering Computer Network at Purdue University hosts riptide server for computationintensive applications. This computer with the following hardware specification is capable of performing both shared-memory and distributed-memory computations.

```
Architecture: x86_64
CPU op-mode(s): 32-bit, 64-bit
Byte Order: Little Endian
CPU(s): 32
```

On-line CPU(s) list: 0-31
Thread(s) per core: 1
Core(s) per socket: 4
Socket(s): 8
NUMA node (s): 8
Vendor ID: AuthenticAMD
CPU family: 16
Model: 2
Stepping: 3
CPU MHz: 2293.808
BogoMIPS: 4589.07
Virtualization: AMD-V
L1d cache: 64K
L1i cache: 64K
L2 cache: 512K
L3 cache: 2048K
NUMA nodeO CPU(s): 0-3

| NUMA node1 CPU(s): | $4-7$ |
| :--- | :--- |
| NUMA node2 CPU(s): | $8-11$ |
| NUMA node3 CPU(s): | $12-15$ |
| NUMA node4 CPU(s): | $16-19$ |
| NUMA node5 CPU(s): | $20-23$ |
| NUMA node6 CPU(s): | $24-27$ |
| NUMA node7 CPU(s): | $28-31$ |

VITA

## VITA

Anandakumar Subbiah received B.E. in electrical and electronics engineering from PSG College of Technology, Coimbatore, India in 2006. He worked for Automation division of Larsen \& Toubro Ltd, India till 2008. He pursued graduate studies at Purdue University, West Lafayette where he received M.S. and Ph.D. degrees in 2010 and 2015, all in electrical engineering. His general interests are electrical machines, power electronics, controls and scientific computing. His current interests are modeling and simulation of power electronic circuits/machines using numerical methods for partial differential equations to model relevant physical phenomena.

