Hybrid Numerical Models for Fast Design of Terahertz Plasmonic Devices

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Shubhendu Bhardwaj, M.S.

Graduate Program in Electrical and Computer Engineering

The Ohio State University

2017

Dissertation Committee:

John L. Volakis, Advisor
Siddharth Rajan, Co-Advisor
Fernando Teixeira
Kubilay Sertel
Niru K. Nahar
Karin Musier-Forsyth
Electron-plasmonic devices are of strong interest for terahertz applications. In this work, we develop rigorous computational tools using finite difference time domain (FDTD) methods for accurate modeling of these devices. Existing full-wave-hydrodynamic models already combine Maxwell’s and hydrodynamic electron-transport equations for multiphysical hybrid modeling. However, these multilevel methods are time-consuming as dense mesh is required for plasmonic modeling. Therefore, they are not suited for design and optimization. To address this issue, we propose new iterative ADI-FDTD-hydrodynamic hybrid coupled model. The new implementations provide time-efficient, yet accurate, modeling of these devices. It is demonstrated that for a typical simulation, up to 50% reduction in simulation-time is achieved with a nominal 3% error in calculations.

Using the new tool-set, we investigate several devices that operate using the properties of 2D electron gas (2DEG). We provide one of the first multiphysical numerical analyses of these devices, giving accurate estimates of their terahertz performance. The developed tool allows simulation of arbitrary 2DEG based terahertz devices, providing useful and intuitive 2D field information. This has allowed understanding of the operation and radiation principles of these devices. Specifically, we examine the known plasma-wave instability in short-channel high electron mobility transistors
(HEMTs) that leads to terahertz emissions at cryogenic temperatures. We also examine terahertz emitters that exploit resonant tunneling induced negative differential resistance (NDR) in HEMTs. Finally, using this tool we numerically demonstrate the existence of acoustic and optical-plasmonic modes within 2DEG bilayer systems in HEMTs. Methods for exciting and controlling these modes are also discussed enabling new physics among bilayer devices.
To all humanity and its prosperity...

To Mamma and Papa for their motivation and inspirations...
Acknowledgments

Foremost, I want to extend my heartiest thanks for Prof. John Volakis for his vision and mentorship through the years of Ph.D. I want to thank him for the opportunity he provided me to conduct my research with freedom and independence. Working under him has also allowed me to experience and learn from it an entrepreneurial synergy that is rarely witnessed in academic arena.

I would like to thank Prof. Siddharth Rajan for his continued advise and support. I would like to thank Sanyam and other students from his group for teaching me sample processing. Next, I would like to thank Prof. Fernando Teixeira for his help with Computational Electromagnetic. Insightful discussion with him have allowed me to develop time-efficient FDTD models (discussed in Chapter-3). Prof. Niru Nahar for her initial support with measurements and developing phaseless CP measurement methods. Prof. Kubilay Sertel for his help with interesting discussions on conducting measurement of terahertz plasmonics using non-contact probes. Special thanks to his post-doc Cosan Caglayan for conducting these measurements. My lab-mates Satheesh, Ushe, Cedric, Asif, Brock, Jingni, Dimitrios (D1) Papantonis, Dimitrios (D2) Siafarikas, for making the workplace a friendly environment with great stories to share. Michelle Diefenbach, Jennifer McCord, Yesim Anter and Sherry Vogt for
taking care of administrative needs and travel related paper work, that allowed me
to present my work nationally and internationally.

Further, I would like to thank our collaborators Prof. Berardi Sensale-Rodriguez, Prof. 
Huili Grace Xing and Prof. Patrick Fay. I would also like to acknowledge the sup-
port of the Office of Naval Research under the project Devices and Architectures for 
Terahertz Electronics Multidisciplinary University Research Initiative (DATE MURI) 
under Grant N00014-11-1-0721. Finally, I want to thank graduate school at The Ohio 
State University for providing funding for the final year of my Ph.D. in the form of 
Presidential Fellowship.
Vita

2008 ................................................. B.Tech, Indian Institute of Technology (ISM-U)- Dhanbad, India

2012 ................................................. M.S., University of California, Los Angeles

2012-present ..................................... Graduate Research Associate,
                                           The Ohio State University

Publications


**Fields of Study**

Major Field: Electrical and Computer Engineering
# Table of Contents

<table>
<thead>
<tr>
<th>Abstract</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td>1. Introduction to Terahertz Plasmonic Devices</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Need for Research in THz Electronics</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Brief Overview of Traditional and Modern Terahertz Sources</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Terahertz Plasma-wave Devices</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Terahertz Detection using Plasma-Wave Concept</td>
<td>7</td>
</tr>
<tr>
<td>1.4.1 Room-temperature Broadband Terahertz Detectors</td>
<td>8</td>
</tr>
<tr>
<td>1.4.2 Resonant Terahertz Detection using Short-gate Length Plasma-wave Devices</td>
<td>9</td>
</tr>
<tr>
<td>1.5 Terahertz Sources using Plasma-Wave Concept</td>
<td>11</td>
</tr>
<tr>
<td>1.6 Modeling of Plasmonic Phenomenon in 2D electron gas systems</td>
<td>13</td>
</tr>
<tr>
<td>1.6.1 Analytical Expressions for Electron Gas</td>
<td>14</td>
</tr>
<tr>
<td>1.6.2 Numerical Investigations for Plasma-wave Phenomenon</td>
<td>16</td>
</tr>
<tr>
<td>1.6.3 Choices for Modeling Approaches of Electron Dynamics</td>
<td>18</td>
</tr>
<tr>
<td>1.7 Summary of the Present Work</td>
<td>26</td>
</tr>
<tr>
<td>2. Full-wave-Hydrodynamic FDTD Solver</td>
<td>29</td>
</tr>
<tr>
<td>2.1 Motivations for Full-Wave-Hydrodynamic Model</td>
<td>29</td>
</tr>
</tbody>
</table>
5. Plasma-modes in a Double 2DEG Channel Systems .......................... 100
   5.1 Introduction ........................................................................... 100
   5.2 Dispersion of Plasma-waves in a Gated Double Channel 2DEG: Small Signal Analysis .................................................... 102
   5.3 Study of Excitation of Modes using Full-wave Hydrodynamic Solver 108
   5.4 Comparison of Single and Double channel systems for Terahertz Devices ................................................................. 114
   5.5 Conclusion ............................................................................ 116

6. Ongoing and Future Experimental Work ................................. 117
   6.1 Capacitively Coupled HEMT Device ........................................ 118
   6.2 Fabrication and Measurements ................................................ 120
   6.3 Conclusions and Further Steps ................................................ 122

7. Conclusions ............................................................................. 123
   7.1 Summary of the Work Conducted and Important Findings ....... 123
   7.2 Future work ......................................................................... 126

Appendices

A. Time-Difference Equations for ADI-FDTD-HD and iterative-ADI-FDTD-HD methods ................................................................. 130
   A.1 ADI-FDTD-HD Difference Equations .................................... 130
   A.2 Iterative ADI-FDTD-HD Difference Equations ....................... 135

Bibliography ................................................................................ 141
List of Tables

Table  

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 CPU-times for various $CN$ value cases for ADI-FDTD-HD method</td>
<td>59</td>
</tr>
<tr>
<td>3.2 CPU-times for various $CN$-$it$ combinations for it-ADI-FDTD-HD method</td>
<td>61</td>
</tr>
<tr>
<td>4.1 Radiated power with changing carrier velocity $v_o$, ($L_c = 100$ nm, $n_{sh} = 10^{12}$ cm$^{-2}$)</td>
<td>78</td>
</tr>
<tr>
<td>4.2 Radiated power with changing 2DEG confinement $t_{2DEG}$, ($L_c = 100$ nm, $v_o = 10^5$ m/s, $n_{sh} = 10^{12}$ cm$^{-2}$)</td>
<td>78</td>
</tr>
<tr>
<td>5.1 2DEG parameters used for the comparison of plasma-wave properties in gated single and double channel HEMTs. Results are shown in Fig. 5.5</td>
<td>113</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Traditional and emerging applications of THz technology</td>
</tr>
<tr>
<td>1.2</td>
<td>(a) A typical HEMT or FET device with gate and drain bias (b) Plasmonic oscillations in the channel shown as electron bunching effect (c) A simulation model showing propagating plasma-wave fields within a gated-HEMT device</td>
</tr>
<tr>
<td>1.3</td>
<td>Qualitative comparison of plasmons in ungated and gated 2DEG (equations (1.2) and (1.4)) with 3D plasmon frequency (given by (1.5)). Smallest wavelengths for a given frequency are achieved for gated plasmons in 2DEG channel</td>
</tr>
<tr>
<td>1.4</td>
<td>Accuracy versus ease of simulation among various simulation methods (table taken from [55])</td>
</tr>
<tr>
<td>1.5</td>
<td>Approximations used in the development of various models</td>
</tr>
<tr>
<td>1.6</td>
<td>Particle approach versus hydrodynamic model for transistor simulations [55]. ( L_g ) refers to gate length</td>
</tr>
<tr>
<td>1.7</td>
<td>(Research work presented in this dissertation in relation to some of the past research in this area)</td>
</tr>
<tr>
<td>1.8</td>
<td>Flow and organization of chapters in the dissertation</td>
</tr>
<tr>
<td>2.1</td>
<td>Typical HEMT geometry with gate discontinuity for exciting the plasmonic oscillation in channel. The red arrows show the electric field lines due to plasmonic-excitation in the channel</td>
</tr>
</tbody>
</table>
2.2 Device cross-section and corresponding modeling equations for predicting the $\vec{E}$-fields in the media and the 2DEG AC currents within the channel. EM and transport equations are solved in time iterations. $E_x$ fields as calculated from Maxwell’s equations are applied to hydrodynamic equations and channel AC current $J_e$ as calculated from the hydrodynamic equations is applied to the Maxwell’s equations within each time iteration.

2.3 Modeling of the HEMT’s cross-section using the developed hybrid numerical algorithm. Non-uniform meshing allows faster simulation times. The 2DEG channel is modeled by transport equations by employing a 1D mesh while EM fields are solved using a 2D mesh in the device cross-section. Typical cell sizes and time step values are $\Delta x = 5 \text{ nm}$, $\Delta y = 1$ to 5 nm, $\Delta t = 10^{-17} \text{ s}$.

2.4 The schematic used for the simulation for comparison with analytical formulae (schematic not to scale).

2.5 Validation of the full-wave hydrodynamic model. The parameters used for electron gas modeling are $\epsilon_r=13.9$, $m_e=0.042m_o$, $n_{sh}=1\times10^{13}\text{ cm}^{-2}$, for gated case $d_{barr}=38 \text{ nm}$, $\mu=10,000 \text{ cm}^2/(\text{V s})$ (a) Electric field distribution (in V/m) due to plasma propagation in a ungated 2DEG layer (b) Electric field distribution (in V/m) due to plasma propagation in a gated 2DEG layer. (c) Dispersion diagrams calculated from the developed numerical model versus theoretical predictions from [46], [47]

2.6 Geometry and resonances associated with the grating-gated HEMT. Top: Modulation of THz spectrum using plasma-resonances within the HEMT. Bottom: Computed and measured plasma-wave resonances. Computations are based on the proposed hybrid algorithms. The measurements are from [74] and serve to validate the developed numerical model.

3.1 Flowchart for ADI-FDTD algorithm coupled with hydrodynamic equations. As noted, the evaluation of electrodynamic fields is split in two sub-time-steps. The hydrodynamic equations may be evaluated at a smaller time-step $\Delta t/R$ for the cases of large $\Delta t$ (or $CN >500$)
3.2 Flowchart for iterative ADI-FDTD algorithm coupled with hydrodynamic equations. Fields are evaluated using ADI-FDTD with iterations in which error term is updated based on fields obtained from last iteration. The HD equations are evaluated at a smaller time-step (if need be for larger CN) to maintain stability of the algorithm. 

3.3 The device along with various dimension (left) along with the simulation domain (right) used for the modeling and verification and comparison of the ADI-FDTD based full-wave-hydrodynamic solvers.

3.4 E-field profiles (calculated as $E = \sqrt{E_x^2 + E_y^2}$ obtained at $t=3$ ps for the various cases of ADI-FDTD-HD method compared with the reference solution. Here, the reference case corresponds to explicit FDTD-HD method.

3.5 Current profile in the channel at $t=3$ ps using the reference FDTD-HD method and ADI-FDTD-HD methods.

3.6 E-field profiles obtained at $t=3$ ps for the various cases of it-ADI-FDTD-HD method compared with the reference solution. Here, the reference case corresponds to explicit FDTD-HD method. Increasing values of $it$ are needed to correct the splitting-errors as $CN$ is increased.

3.7 Channel current-profile using iterative-ADI-FDTD-HD method, compared to the reference data. Three plots refer to values of $CN=100$, 200 and 300 respectively.

3.8 Variation of error and time-cost with increasing iteration-count for $CN=100$, 200 and 300. The error terms are marked in red and time variation is plotted in blue.

3.9 Exponential fit modeling of the error in it-ADI-FDTD-HD method with increasing iteration counts.
4.1 Left: Short channel HEMT geometry, and biasing conditions and calculations showing its emission properties. Right: Mechanism of reflective gain as demonstrated in a 2DEG channel using the developed numerical model (Black curves: forward/backward waves. Red curves: reflected waves. Dashed-dotted curves: envelop of fluctuations. Employed parameters for numerical simulations were $n_{sh} = 1 \times 10^{12} \text{cm}^{-2}$, $\tau = 1 \text{ ps}$, $\varepsilon_r = 13.9 \text{ (InGaAs)}$, and $m_e = 0.042 m_o$. Also drift velocity $v_o$ is toward the right [38]. Details are discussed in future sections.

4.2 Calculated emitted spectra for the GaN/AlGaN RTD-gated HEMT. Parameters used: $\varepsilon_r = 9.5$, $d_{barr} = 30 \text{ nm}$, $L = 100 \text{ nm}$, $W = 70 \text{ nm}$, $\mu = 1200 \text{ cm}^2/\text{V.s}$, $g = -6.5 \times 10^{12} \text{ S/m}^2$. Details can be referred from [83].

4.3 Schematic of the asymmetrically biased HEMT as proposed in [21]. Here $\lambda_g$ is the plasmon wavelength at the resonance frequency.

4.4 Schematic of the asymmetrically biased HEMT as proposed in [21]. Here $\lambda_g$ is the plasmon wavelength at the resonance frequency.

4.5 Simulation domain used for the modeling of HEMT (not to scale), $d_{barr} = 22 \text{ nm}$, $d_s = 1.5 \mu\text{m}$, $p = 50 \text{ nm}$, $d_{PML} = 5 \mu\text{m}$.

4.6 Modeling of DS-instability in InGaAs HEMT (a) Schematic used to model the HEMT (not to scale), $d_{barr} = 22 \text{ nm}$, $d_s = 1.5 \mu\text{m}$, $p = 50 \text{ nm}$, $d_{PML} = 5 \mu\text{m}$ (b) Current density with time as recorded at $D$ for varying $L_c$ ($v_o = 1 \times 10^5 \text{ m/s}$) (c) Current density with time as recorded at $D$ for varying $L_c$ ($v_o = 2 \times 10^5 \text{ m/s}$) (d) AC-Current and AC-field variation in the channel near $S$ terminal during one frequency cycle. (e) Corresponding radiation-spectra as recorded at the measurement boundary.

4.7 2D electric-fields ($E_x$ component) associated to phase points $A$, $B$, $C$ and $D$ in Fig. 4.5(d).

4.8 Frequency and power dependence on channel length, $L_c$. ($n_{sh} = 1 \times 10^{12} \text{ cm}^{-2}$, $v_o = 2 \times 10^5 \text{ m/s}$). Theoretical calculations as per [35].

4.9 Frequency and power dependence on channel carrier density, $n_{sh}$, in the 2DEG layer. ($L_c = 100 \text{ nm}$, $v_o = 2 \times 10^5 \text{ m/s}$). Theoretical calculations as per [35].
4.10 Effect of velocity saturation and decreased mobility on the emitted power from the device. The considered case is for $L_c = 100$ nm, $n_{sh} = 1 \times 10^{12}$ cm$^{-2}$.  

4.11 Cross-sectional schematic of a HEMT with a vertical RTD between the gate and the channel. The dashed line represents the 2DEG and the S/D stand for source and drain.  

4.12 Propagation constant and attenuation of the plasma waves as a function of changing differential conductivity $g$ of the RTD. For calculations, we considered $\tau = 0.1366$ ps, $v_o = 10^7$ cm/s $\epsilon_r = 9.5$ for GaN, $n_o = 5 \times 10^{12}$ cm$^{-2}$, and $d = 25$ nm. (Negative values signify wave amplification).  

4.13 Attenuation constant (in Np/$\mu$m) as a function of relaxation time (\(\tau\)) and RTD-NDR (\(g\)). The marked line refers to the zero attenuation contour. For the calculations, we considered $v_o = 10^7$ cm/s, $\epsilon_r = 9.5$ for GaN, $n_o = 5 \times 10^{12}$ cm$^{-2}$, and $d = 25$ nm. (a) Plots at frequency $f = 1$ THz (b) Plots at frequency $f = 5$ THz (Negative values signify wave amplification).  

4.14 Model used for the numerical calculations. The size of the discontinuity is chosen to be $p = 40$ nm.  

4.15 Wavenumbers and the attenuation constants obtained using the full-wave method and the analytical method. The "symbol" lines refer to numerical simulations and the solid lines refer to the analytical model. Choice of parameters: $\epsilon_r = 1$, $d = 25$ nm, $n_o = 5 \times 10^{12}$ cm$^{-2}$, $\tau = 0.137$ ps, $v_o = 10^7$ cm/s and $m_e = 0.2 m_o$.  

4.16 Wavenumbers and the attenuation constants for GaN based RTD-gated HEMT obtained using the full-wave method and the analytical method. The "symbol" lines refer to numerical simulations and the solid lines refer to the analytical model. Choice of parameters: $\epsilon_r = 9.5$, $d = 25$ nm, $n_o = 5 \times 10^{12}$ cm$^{-2}$, $\tau = 0.137$ ps, $v_o = 10^7$ cm/s and $m_e = 0.2 m_o$.  

xvii
4.17 Left: Schematic of the RTD-gated HEMT with RTD typical characteristics. Right: $E_x$ field snapshots taken at $t=2\text{ps}$. Detail along the partial section of the channel. Plasma-wave propagation for varying values of the RTD-NDC $g$ is shown. Reduced plasma-wave losses and amplification can be obtained with increasingly $-g$. (Note: $x$ and $y$ axis are drawn with respect to the corner of the 2D-simulation domain and figure shows the zoomed in view.)

5.1 (a) Schematic of the double channel HEMT device based on GaN/AlGaN lattice (b) Field profiles used for the analytical solution in the regions of the gated double channel system.

5.2 Dispersion relations for the gated bilayer obtained from (5.14). The curves correspond to varying inter-channel distance with a fixed $d_1=40\text{ nm}$. The curves show that the acoustic plasmons are similar to those in the ungated bilayer case. However, the optical plasmons have smaller propagation and higher attenuation constants.

5.3 Schematic of the device used for simulations and corresponding dispersion plots and field plots. AP: acoustic plasmon, OP: optical plasmon.

5.4 Excitation and propagation of optical/acoustic modes in gated/ungated bilayer, (a) Simulation set-up used for the results shown in parts (b) and (c) (b) Fields in an ungated bilayer with a gap discontinuity (c) Fields in a gated bilayer with a gate discontinuity. We observe that spectral content of the oscillations can be controlled by choosing the appropriate gate or gap dimensions. Parameters used: $n_{o1} = n_{o2} = 5 \times 10^{12} \text{ cm}^{-2}$, $\mu = 10,000 \text{ cm}^2/\text{V.s}$, $d_1 = 28 \text{ nm}$, $\Delta d = 20 \text{ nm}$.

5.5 Comparison of the propagation constant and attenuation constants in single and double channel HEMTs. Used parameters are defined in Table-5.1.

5.6 Illustrative field line plots for the single and double channel systems for gated and ungated bilayer cases.
6.1 Device schematic of the proposed device for measurement of terahertz resonances in HEMT. (a) Top and cross-sectional views of the device. The measured electron density was $n_{sh}=1.39 \times 10^{13}$ cm$^{-2}$ and mobility was $\mu=1290$ cm$^2$/V.s (b) HFSS simulation model. (c) Field distribution at resonance frequency.

6.2 Results of HFSS simulation modeling. Left: Variations in resonance frequency, seen as the peak of $\Re(Z_{in})$ as observed at the sample plane. Dotted-black curves are at 0.2 µm steps of $L_r$. Center: Phase($Z_{in}$) at the sample plane. Right: Normalized phase of $Z_{in}$ (phase is normalized by subtracting uniformly varying phase of -80° to -40° across the band).

6.3 (a) Fabricated samples along with antenna structure for non-contact probing (b) Fabricated devices with varying sample lengths (c) Measured $\Re(Z_{in})$ for different samples (d) Measured phase and variation due to sample length variation.

7.1 Work presented in this dissertation as categorized in three blocks: 1) Physical understanding of plasmonic phenomenology, 2) New Numerical techniques and 3) Device modeling. The research in this dissertation, in relation to some other works in this area, is shown.

7.2 Summary of presented work and important findings reported in this dissertation.

A.1 Mesh-grid scheme and stencil illustration for Equation A.10. Iterative terms only apply for Equations A.27.

Chapter 1: Introduction to Terahertz Plasmonic Devices

1.1 Need for Research in THz Electronics

Efficient utilization of millimeter-wave and terahertz bands is one of the most challenging areas of the modern electronics. Innovations in these frequency bands will impact wide array of technologies including high-data rate wireless-communication, security and medical imaging, sensing and vehicular radar. For imaging applications, these frequencies are especially significant because they allow resolutions higher than microwave and can penetrate clothes and skin and yet are non-ionizing (unlike X-rays). They also exhibit unique spectrum to various gases and materials. Therefore, they are used for the identification of poisonous gases, explosives and in astronomical explorations.

Further, modern high-capacity wireless links (or 5G communication links) are expected to rely on lower end of terahertz regime [1, 2]. High attenuation bands in this spectrum, such as bands at frequencies 183 GHz, 325 GHz and 380 GHz (refer Fig. 1.1), are ideal for densely packed electronic media, where their dramatic attenuation allows ‘whisper’ radio - a communication, where weak signals do not propagate more
than a few meters before dropping below the thermal noise level. At lower attenuation bands of 77 and 240 GHz, cellular, backhaul, fiber-replacement, sensing, and vehicular radar will be viable.

In spite of above, electronics in this regime has progressed rather slowly due to lack of sources, detectors, mixers and other needed components. Ultimately, traditional high speed electronics could not be scaled for this band due to limitations in fabrication technologies. Specifically, transistors with small gate lengths have not yet been realized for operation beyond few hundreds of GHz [3–6]. When we consider opto-electronics as a possible solution, we find that material systems with sufficiently small band-gap that will act as natural terahertz sources are yet to be identified.

Lack of terahertz devices has led to search for alternative methods for terahertz generation, detection and other non-linear functions. Plasma-wave electronics, which utilizes the plasmonic-oscillations within 2D electron-gas systems, is one such method. In this dissertation work, we develop computational methods for plasma-wave modeling of terahertz devices that have potential for terahertz generation and detection. We will also apply the developed models for terahertz source applications.

Before going into the content of this dissertation, we will briefly review the approaches used for terahertz generation in past. This is considered in the following section.
1.2 Brief Overview of Traditional and Modern Terahertz Sources

1. *Up-conversion from W-band* frequencies using Amplifier Multiplier Chains is one of the most common method for THz generation. Frequencies beyond 500 GHz and upto 2.7 THz have been demonstrated using this method [7–9].

3. *Vacuum based devices*, such as backward wave oscillators (BWOs) offer the most power and frequency tunability in sub-millimeter band [2]. Another method of THz generation in the category of vacuum based devices is IR-pumped lasers. [14] provides a review of these sources.

4. *Compact Solid State Sources*: A compact THz source based on semiconductor technology is needed for many current and future applications of this technology. Some of the important development in this direction are listed as follows.

(a) *Ultrascaled HEMTs*: HEMT devices are widely used for designing oscillators and amplifiers in the microwave regime. However, in the terahertz band, these are limited by reactive parasitics or the transit time delays that cause the high frequency roll-off. Many times they have simple resistive losses that dominate the impedance at these wavelengths. Nevertheless, this technology has achieved cut-off frequency ($f_T$) higher than 550 GHz in [6]. Although, we note that such transistors must be operated well below the $f_T$ for exciting sustained oscillations as a source.

(b) *Quantum Cascade Lasers* Quantum cascade lasers (QCLs) have seen an impressive development in past several years [15–18]. They were invented in a joint collaboration by NESTPisa and Cavendish Laboratory, Cambridge [15]. QCL operation comprises of electrons being injected into a periodic structure of a superlattice under electrical bias. They undergo inter-subband transitions with THz photon emission excited by resonant tunneling through the multiple wells. As a result, a cascade process occurs,
which gives name to the source. The operating temperatures for these devices are still low (137 K for pulsed and 93 K for CW) [18]. Nevertheless, they can provide power upto several mW.

(c) Recently, advancements in antenna coupled resonant tunneling diodes have been able to provide power beyond 1THz [19, 20] An electron tunneling through a double barrier structure, experiences a negative differential resistance (NDR) due to discrete energy levels arising in the quantum-well. This NDR, when coupled to an external passive RF circuit (or an antenna), causes oscillations leading to terahertz emission.

1.3 Terahertz Plasma-wave Devices

As we search for alternative methods for making terahertz devices, there has been significant interest in terahertz plasmonic devices in recent years [21–30]. Electronic plasma-waves originate due to variations of electron-density in a electrons-gas system (or a gas of freely moving electrons). As an example, these oscillations are present in 2D electron-gas channel of the HEMTs as illustrated in Fig. 1.2. Note that the density oscillations are ‘acoustic’- much like ‘mechanical’ waves or sound waves. But they also have an associated electromagnetic field due to charged nature of the particles. Thus, such waves can carry electromagnetic power. For example, a 2D confined electron gas (or a 2DEG) channel can support propagating plasmonic modes carrying electromagnetic power in the form of electronic density variations and associated electromagnetic fields. The electric and magnetic field profiles for these can be seen
Figure 1.2: (a) A typical HEMT or FET device with gate and drain bias (b) Plasmonic oscillations in the channel shown as electron bunching effect (c) A simulation model showing propagating plasma-wave fields within a gated-HEMT device.
on both sides of the 2D channel (Fig. 1.2). Thus, the 2DEG can act as a terahertz waveguide. This observation is the basis for plasma-wave devices and related plasmonic phenomenon at terahertz.

Electronic plasma-wave concept enables confinement of electrically long terahertz waves (wavelengths of 10s and 100s of microns) within micro/nano-dimensional semiconductor devices. This is due to small wavelengths associated with the plasmons in 2DEGs of modern transistors. A key factor enabling this is typical electron-density in such 2DEGs being between $1 \times 10^{11}$ and $1 \times 10^{13}$ cm$^{-2}$. At terahertz frequencies, such 2DEG systems would allow plasmons with wavelengths ranging from few tens of nanometers to several microns. Note that these dimensions are consistent with channel lengths and gate-dimensions of the modern HEMT and FET devices. For example, if a quarter wavelength resonant cavity was to be realized, it could be achieved using HEMT/FET based 2DEGs. Following this argument, we will consider several specific applications of the plasma-wave phenomenon in following section.

1.4 Terahertz Detection using Plasma-Wave Concept

Semiconductor devices that contain 2D electron gas (2DEG) sheets can be used as efficient terahertz detectors [21–30]. Indeed, such terahertz detection has been the most practical application for the plasma-wave concept [31,32]. A terahertz wave incident upon a FET or HEMT device is rectified to DC voltage, which is then observed between the drain and the source terminals of the device. This phenomenon can be used for terahertz detection. The detection may occur via two mechanisms, a choice
decided by the conditions in the channel and the dimensions of the gate and the channel. If the channel dimensions are large and it has relatively smaller channel mobility, broadband terahertz detection is exhibited by the device [22–25]. On the other hand, in short channels with large electron-mobility, the detection mechanism is resonant and frequency dependent. This means that the detection is maximized for certain discrete resonant frequencies. This mode is often observed in short channel FETs at cryogenic temperatures [26–30]. Detailed mechanisms and example references are discussed in the following sections.

1.4.1 Room-temperature Broadband Terahertz Detectors:

Broadband terahertz detection has been the most successful application of plasma-wave phenomenon. Such detection has been demonstrated for Si-MOSFETs for sub-terahertz and terahertz frequencies upto 2.5 THz [29, 30]. Remarkably, the noise equivalent power (NEP) in these has been found to be one of the lowest for such high frequencies [30]. These have also been integrated for terahertz imaging applications as reported in [31, 32]. Apart from Si-MOSFETs, GaAs/AlGaAs and InGaAs/GaAs HEMTs have also been shown to demonstrate broadband terahertz detection at room-temperature [22–25].

For broadband detection, the plasmonic-waves in the 2DEG are over-damped owing to small relaxation times ($\tau$). This regime of operation is characterized as ‘low frequency regime’ and is expressed with condition $\omega \tau \ll 1$. Rectification in such 2DEG systems occurs due to simultaneous modulation of the electron-velocity and electron-density. The incident terahertz radiation is applied between the gate and source using an
antenna. This causes the modulation of the electron-density \((n_{sh})\) in the channel due to active gate control. Next, a plasma-wave is excited under the gate due to coupling of ambient field into the opening gap between gate and source terminals. This causes modulation of electron velocity \((v)\). These two stimuli lead to a squared sinusoidal modulation of current \((i = \langle -qn_{sh}v \rangle)\) leading to a DC rectification. Thus obtained DC current (or equivalent voltage) can be observed between the drain and source terminals of the device.

### 1.4.2 Resonant Terahertz Detection using Short-gate Length Plasma-wave Devices

Another mode of terahertz detection is possible for high frequency regime (with \(\omega \tau > 1\)), where the plasmonic modes are not overdamped [22,26–30,33]. For this case, the wave amplitude is not reduced significantly as the wave-propagation occurs along the channel. Thus, the wave reaching the end of the channel is reflected back leading to resonant cavity like operation. Of course, for such operation the channel needs to be approximately quarter wavelength long. Along with resonance, the rectification phenomenon also occurs as discussed for the broadband case, but now dominantly for specific resonant frequencies.

Thus, a resonant mode detection is achieved. Advantage is, of course, high detection potential obtained at the resonant frequencies and frequency-tunability by tuning channel electron-density (since plasmonic wavelength is a function of channel electron density) using gate biasing. But such resonant detection is only possible at low temperatures for typical HEMTs, since higher electronic-relaxation times \((\tau)\) are
needed to realize under-damped propagation modes. Another issue is related to the fabrication. For this mode of operation, resonant frequency should be high ($\omega \tau > 1$). Therefore, state-of-the-art fabrication techniques, allowing submicron length gate fabrication, are necessary for this mode of operation.

We make further note of the asymmetric boundary conditions needed to implement above mode of detection. Note that asymmetric boundary conditions [22, 33], are needed to implement the said propagation and reflection of plasmonic modes in the 2DEG channel. Specifically, low-impedance boundary conditions (at terahertz frequencies) are needed near the source-gate terminal for coupling of the incident waves. At the gate-drain terminal, high impedance boundary conditions are required to enable wave-reflection. This asymmetry can be achieved from various sources. Authors in [22] note that

“There may be various reasons of such an asymmetry. One of them is the difference in the source and drain boundary conditions due to some external (parasitic) capacitance. Another one is the asymmetry in feeding the incoming radiation, which can be achieved either by using a special antenna, or by an asymmetric design of the source and drain contact pads. Thus, the radiation may predominantly create an AC voltage between the source and the gate (or between the drain and the gate) pair of contacts. Finally, the asymmetry can naturally arise if a dc current is passed between source and drain, creating a depletion of the electron density on the drain side of the channel [3]. In most of the experiments carried out so far, the THz radiation was applied to the transistor channel, together with contact pads and bonding wires”.
1.5 Terahertz Sources using Plasma-Wave Concept

Even with a variety of methods available for terahertz sources, existing technologies fall short in providing a compact, solid state terahertz source at room-temperature. Therefore, there has been extensive research towards alternative sources. In last few decades, there has been some interest in using plasma-wave concept for terahertz sources [33,34]. A 2D confined electron gas system, such as one found in HEMTs, can be used to create plasmonic oscillations. Many different concepts are being pursued even within the broader category of plasma-wave electronics. Terahertz emissions using optically pumped grating gated devices, emissions from nano-gate length HEMTs and generation using graphene are to name a few. These will be briefly discussed in following subsections.

Emissions from Plasma Wave Oscillations in Nano-Gate Length Transistors

Like resonant detection as discussed in Section- 1.3.2, short gate-length HEMTs or FETs are also capable of emitting terahertz radiations at cryogenic temperatures. It is noted in [33,35] that under strong drain to source bias and asymmetric boundary conditions (as discussed in Sections 1.2 and 1.3), current instability can be observed in the channel. This instability creates oscillations at frequencies corresponding to channel length, electron density etc. As originally proposed, the channel should be asymmetrically biased, i.e. AC short and AC open must exist at the two ends of the channel to cause such oscillations. However, past experimental reporting do not always mention the need for such boundary conditions, as in the case of [36]
using InGaAs heterojunctions. These terahertz oscillations, primarily found at lower temperatures, persists even till room temperatures as shown for InAlAs/InGaAs and AlGaN/GaN heterojunction devices in [37].

In this dissertation work, we will further examine the mechanism of such plasma-wave instability using full-wave hydrodynamic model in Chapter-4. Use of full-wave simulations allows us to calculate the extracted power levels in such scenarios, which was not possible using analytical methods prior to this work. Details of this work can also be referred from [38]

**Plasma-wave Emission using Grating Gated Devices**

In past, grating gated plasma-wave HEMTs are used for confirmation of plasma resonances [39,40]. However, these are also shown to emit terahertz radiations. In [41], a grating gated device using In-GaP/InGaAs/GaAs material system is proposed for THz emissions. Device consists of a 2DEG channel covered by a doubly interdigitated grating gate. The grating gates are needed to modulate the electron density in the channel, while also acting as antenna for emission. The device is reported to emit broadband terahertz radiation even at room-temperature. As reported, optical pumping could be used to enhance the radiation from the device.

**Graphene based Terahertz Oscillators**

Graphene has shown potential for variety of THz devices [42]. Due to much smaller electron-effective mass (than similar 2DEG counterparts), plasma resonances are stronger, making it a good candidate for terahertz sources. Optically pumped graphene
nano-ribbons (strips of graphene) have been predicted to show terahertz emissions in [43]. Such graphene ribbons at micron dimensions have also demonstrated metamaterial properties tunable over a wide band. Another mechanism of terahertz generation using graphene is by population inversion, leading to plasmon amplification through stimulated emission as discussed in [43, 44].

Graphene technology is new and is exciting because of its unique electrical and mechanical properties. Recent advancements have already demonstrated its application to terahertz modulators and detectors. As far as terahertz sources are concerned, various concepts have been proposed, but experimental validations are still awaited.

So far, we have discussed some important devices that use the plasma-wave concept and have been demonstrated for various applications. The objective of this dissertation is to demonstrate new time-domain numerical algorithms for full-wave modeling of plasma-wave devices. Specifically, we develop efficient time-domain models for full-wave-hydrodynamic modeling of plasma-wave devices. As a background to that, we will first consider past efforts for the modeling of plasma-wave devices.

1.6 Modeling of Plasmonic Phenomenon in 2D electron gas systems

We will consider methods traditionally employed for modeling of plasma-wave phenomenon and associated device applications. Broadly, the modeling can be classified into analytical approaches, circuit modeling and full-wave numerical approaches. In the following sections, we will consider the analytical and numerical approaches used thus far.
1.6.1 Analytical Expressions for Electron Gas

Behavior of plasma-oscillations in a 2D electron gas of a HEMT or FET is quite relevant for plasma-wave devices. These transistors can have ungated regions. Modeling of plasma-wave oscillations in these regions is done by modeling a 2DEG channel embedded in a dielectric media. The dispersion relation \( k - \omega \) relation for such cases was given by Stern [45] in 1967 as

\[
\begin{align*}
k &= \frac{1}{2a} \sqrt{\omega^4 + \frac{4a^2 \omega^2 \epsilon_c}{c^2}} \\
a &= \frac{n_{sh} q^2}{4m_e \epsilon_c}
\end{align*}
\]

Here, \( n_{sh} \) is the sheet electron density, \( q \) is the charge on one electron, \( m_e \) is the effective mass of electron and \( \epsilon_c \) is the permittivity of the dielectric media. \( \omega = 2\pi f \) is the angular frequency and \( k \) is the propagation constant of the plasmonic-wave. Note that the relation 1.1 is derived using the assumption that the 2DEG sheet is infinitely large and infinitesimally thin. In practice, finite 2DEG size and thickness will alter the dispersion relation. For high frequency operation, the dispersion relations are further reduced to quadratic dependence. The expression then becomes

\[
k = \frac{\omega^2}{2a}
\]

Another important formulation commonly used is for plasmonic oscillations under the gate of a FET or HEMT device. Equivalently, dispersion relation for metal-insulator-semiconductor (or gated 2DEG) systems was developed by Quinn et. al. [46],
\[ \omega = \sqrt{\frac{n_{sh}q^2}{m_e}k} \frac{1}{\epsilon_c + \epsilon_{ox} \coth kd} \]  

(1.3)

Here \( \epsilon_{ox} \) is the permittivity of the oxide (insulator) layer between the 2DEG and metal gate. \( d \) is the thickness of the insulator layer. Above relation is approximated commonly for the case when \( d \gg \lambda_p \). \( \lambda_p \) is the plasmonic wavelength. In such cases, we use approximation \( \coth kd = \frac{1}{kd} \) to get

\[ \omega = k \sqrt{\frac{n_{sh}q^2}{m_e \epsilon_{ox}}} \]  

(1.4)

Above assumption is equivalent to assuming that the fields are confined between the 2DEG and the metal (i.e. in the insulator layer) and there are no fields below 2DEG layer. Note that, opposed to the ungated case, now the dispersion relation is linear.

Finally, it is worthwhile to mention that a 3D bulk electron gas system will have a specific resonance frequency also called plasma-frequency for the system. It is given by

\[ \omega_p = \sqrt{\frac{nq^2}{m_e \epsilon_o}} \]  

(1.5)

Here \( n \) is the electron density in 3D bulk. We do not have a \( k \) in above expression, signifying that this system does not allow a propagation mode. That is, a 3D bulk of electron gas can resonate at certain frequency, but does not support propagating modes commonly found in 2DEG and gated-2DEG systems.
A recently published work [47] is noteworthy as authors have also considered effect of finite drift velocity and momentum relaxation time in the expression for dispersion relations. These effects are relevant for 2DEG based plasmonic devices with drain to source bias. The finite relaxation time is also critical to be modeled as it dictates the attenuation of the propagating wave in the channel. Note that in equations 1.2 and 1.4 there is no imaginary part, which means that wave is non-attenuating. But having included the effect of finite relaxation time, we obtain expressions containing imaginary part of the propagation constant signifying wave attenuation.

1.6.2 Numerical Investigations for Plasma-wave Phenomenon

As noted, analytical methods are primarily applicable for canonical problems, where simplifying assumptions are made in order to solve a certain problem. For our cases,
this usually means that the plasma-wavelengths are much smaller than the dimensions of the geometry under considerations (or 2DEG is infinitely large). In HEMTs, the channel is terminated by ohmic contacts, also containing drain and source metallic regions. Thus, the plasmons frequently interact with these heterogeneous regions. These effects are not accounted by the analytical methods and therefore full-wave-numerical solutions must be developed. Note that here ‘full-wave-solution’ refers to solution of Maxwell’s electrodynamic equations, needed wave-propagation modeling outside the channel. As usual, the electron transport in the channel can be modeled using transport equations, such as Boltzmann equation, hydrodynamic equations or drift-diffusion equations. At mm-wave and terahertz frequencies, coupled solution of electrodynamics equations is critical as noted in reference [48] as following.

“The interactions between carriers and fields in semiconductors at low frequencies (<100 GHz) can be adequately described by numerical solution of the Boltzmann transport equation coupled with Poisson’s equation. As the frequency approaches the THz regime, the quasi-static approximation fails and full-wave dynamics must be considered”

Therefore, for complete modeling, the electron transport equations and full-wave equations must be solved together. This provides the most rigorous form of solution for such modeling problems. These so called ‘global-solutions’ have been pursued in past [48–54]. Among these, reference [49] presents early work in this direction, where authors developed a transient solutions for solid state devices. In [50–53], authors have used finite difference time domain (FDTD) algorithm for self-consistent solution of the hydrodynamic equations and Maxwell’s equations. For such cases, the initial field values for the electrodynamic solution are obtained by solving Poisson equation for steady state conditions. We will be using a version of this method and some
improvements over next several chapters in this dissertation work. Further efforts for full-wave modeling are reported in [53], where the non-linear circuit model (or a large signal model) is solved in conjugation with full-wave model. This unique approach allows analysis of nonlinear properties such as harmonic generation and intermodulation. At device level, the most rigorous form of solution comprises of using Boltzmann transport equations (BTEs) for electron transport along with the full-wave solution. In [48], such solution is shown using Ensemble Monte Carlo (EMC)-FDTD simulations. EMC allows particle level modeling of electron transport which is especially accurate for small channel dimensions.

Clearly, the differences among above methods emerge from the differences in modeling of the electron transport in the channel. We will be using hydrodynamic equations for this modeling in this dissertation work. Therefore, it is of interest to understand the approximations involved for such modeling along with accuracy of other models and regimes of their applicability. We will consider these in next sections.

### 1.6.3 Choices for Modeling Approaches of Electron Dynamics

We will be discussing three well known approaches for modeling electron dynamics in the 2DEG channel of a semiconductor device- 1) Particle model 2) Boltzmann Transport Equations (BTEs) 3) Hydrodynamic Model. Among these, particle model is the most rigorous approach to analyze the plasma dynamics, when compared to kinetic model or Hydrodynamic model. This model involves calculation of trajectories of the individual particles under the influence of external force and internal Coulombic
interactions. Boltzmann transport equations (specific case of kinetic theory of gases) models such system as a priory known distribution of particles, whereas hydrodynamic model is a macroscopic model which considers local bulk parameters to describe the same. In following sections, we would briefly describe each. A summary of underlying assumption and regimes is depicted in Fig. 1.4 and Fig. 1.5.

**Particle Approach**

The most rigorous solution of electron transport in a semiconductor material can be obtained by considering the trajectory of each carrier under the influence of external
and internal forces. Such particle based simulations are actually a subset of a broader realm of ‘molecular dynamics’ simulation commonly used for investigation of the thermal properties of gases. This is usually done by using Vlasov equations [55],

\[ \frac{\partial f}{\partial t} + \vec{v}.\nabla_x f + qm_e(\vec{E} + \vec{v} \times \vec{B}).\nabla_v f = 0 \] (1.6)

Figure 1.5: Approximations used in the development of various models

Here, $f$ is the distribution function which can be used to obtain initial particle distributions. Although equation uses a distribution function, we remark that simulations using particle approaches apply on each particle. A numerical implementation of this is known as particle in cell (PIC) algorithm [56, 57]. Note that collision between the particles is ignored, which is a common assumption in gases. In solids (such as metals and semiconductors), transport is known to be dominated by scattering phenomenon which is accounted for by introducing a collision term in Vlasov equations, given by [55]
Figure 1.6: Particle approach versus hydrodynamic model for transistor simulations [55]. $L_g$ refers to gate length.

\[ \frac{\partial f}{\partial t} + \vec{v}.\nabla_x f + qm_e (\vec{E} + \vec{v} \times \vec{B}).\nabla_v f = \left( \frac{\partial f}{\partial t} \right)_{\text{coll}} \]  

(1.7)

Above is Boltzmann Transport Equation (BTE). A practical implementation of solving above using particle approach is by Monte Carlo (MC) simulations. The MC algorithm consists of generating random flights for each particle under the influence of external or internal forces, choosing a type of scattering mechanism to terminate the flight, then calculating the final momentum and energy of the particle. This is then repeated for the next particle. The fields due to each particle (in case of charged particles) can be calculated using Maxwell’s equations. This particle-based picture, in which the particle motion is decomposed into free flights terminated by instantaneous collisions, is basically the same approximate picture underlying the derivation of the semiclassical Boltzmann Transport Equation (BTE). In fact, it may be shown that the one-particle distribution function obtained from the random walk Monte Carlo technique satisfies the BTE for a homogeneous system in the long-time limit.
Boltzmann Transport Equations

Kinetic theory, more commonly known as Boltzmann Kinetic equations, is intermediate level description of particle dynamics lying somewhere between the macroscopic level hydrodynamic equations and microscopic level atomistic view. Note that BTE cannot be solved analytically, since they are 7 dimensional (3-phase, 3-space, and time) equations governing the plasma dynamics. Therefore, to solve them we must rely on either particle methods (Monte Carlo simulations) or approximations such as hydrodynamic model. More recently BTE has been proposed to be solved using spheroidal harmonic expansion of the distribution function.

Nevertheless, particle methods are more rigorous, since BTE fundamentally assumes following:

1. The number of particles should be very large. Mathematically, it should be infinite. This is so because now the distribution can be described by a probability function in 6-dimensional phase-space.

2. To account for the collisions between the particles, collision integral is used, which assumes that 1) at a time only two particles collide with each other and 2) colliding particles are not correlated (or have any history). Here, we quickly present the BTE and significance of its terms. For an electron distribution \( f \), the BTE states that distribution is always conversed, i.e.,

\[
\frac{\partial f}{\partial t} + \vec{v}.\nabla_x f + qm_e(\vec{E} + \vec{v} \times \vec{B}).\nabla_v f = \left( \frac{\partial f}{\partial t} \right)_{coll}
\]  

(1.8)
Distribution \( f \) can be any generic distribution however commonly used distribution is Maxwellian distribution. Also note that distribution is only a initial choice, i.e., depending on the conditions non-equilibrium state may be achieved which can change the shape of the distribution. Thus BTE allows the possibility of general distribution.

In the equation, \( \frac{\partial f}{\partial t} \) represents time rate of change in the electron distribution which is accounted for by either a by a diffusion process \((\vec{v}.\nabla_x f)\), which could be result of concentration or temperature gradient or by a force term \((\vec{v} \times \vec{B}).\nabla_v f\) which accounts for change in momentum due to electric field force and Lorentz force. Another mechanism by which the distribution can change is by inter-particle collision which changes the momentum of the electron. This is accounted for by the term \( \left( \frac{\partial f}{\partial t} \right)_{coll} \) in the equation.

The collision term can be explained as following. Assuming that the probability of transition of electrons from a state \( k \) to \( k' \) during the collision is given by \( S(k, k') \) and noting that probability of the electron state \( k \) being full and \( k' \) being empty are given by \( f(k) \) and \( 1 - f(k') \), we write [55],

\[
\left( \frac{\partial f}{\partial t} \right)_{coll} = \sum_k \left[ -S(k, k')f(k)(1 - f(k')) + S(k', k)f(k')(1 - f(k)) \right] \quad (1.9)
\]

As noted before, collision integral accounts for only two-particle collisions.
Hydrodynamic Equations

Boltzmann equation is complex to solve because of its higher dimensionality. In return it also gives us a lot of information. Although in most practical cases, we do not require so much information. Usually, we need to know the variations in \( n \) and \( v \) with time and space. For this, we can make certain assumptions and reduce the equation. General advantages and disadvantages of particle approach as compared to hydrodynamic model are given in Fig. 1.6.

Under the assumptions that variations from the initially assumed distribution are slowly varying in time and are in small in amplitude, we can safely apply conservation equations on local bulk parameters such as \( v \) and \( n_{sh} \). Specifically, if the macroscopic observation length (\( L \)) and macroscopic observation time (\( t \)) follow

\[
L >> \lambda_{mfp} \quad \text{and} \quad t >> \tau_{mft},
\]

then macroscopic approximations, namely hydrodynamic equations, can be applied [55]. Here \( \lambda_{mfp} \) is the distance traveled by electron between two collision and \( \tau_{mft} \) is the time between two collisions (also called momentum relaxation time \( \tau \)).

A formal derivation of the hydrodynamic equations can be obtained by considering moments of BTE. To that end, we multiply BTE by 1, \( mv \) and \( mv^2/2 \) and then integrate over momentum space (or velocity space). It can be shown that first two moments of the equation reduce to
\[
\frac{\partial \rho}{\partial t} + \vec{\nabla} . (\rho \vec{v}) = 0 \quad \text{and} \quad (1.11)
\]

\[
\frac{\partial \vec{p}}{\partial t} + (\vec{p} \cdot \vec{\nabla} ) (\vec{v}) = -q(\vec{E} + \vec{v} \times \vec{B}) - \frac{KT}{n} \frac{\partial n}{\partial x} - \frac{\vec{p}}{\tau} \quad (1.12)
\]

Here \(\tau\) is the phenomenological relaxation time constant, which is either obtained from experiments or from bulk Monte Carlo simulations. The term \(\frac{KT}{n} \frac{\partial n}{\partial x}\) is pressure term accounting for the force due to gradient in concentration. \(\rho = (m_e n)\) is local bulk density, \(n\) is local bulk electron concentration and likewise \(\vec{p}\) is the local bulk momentum of electrons. We note that in case of charged particles we must also account for field variation due to moving electrons, which can either be done using Poisson equation or using Maxwell’s equations as shown before in previous sections.

Note that it will not be possible to solve HD equations unless a closure equation is assumed. Each new HD equation (due to new moment of BTE), will introduce a new variable. E.g. first HD equation has \(n\) and \(v\) as variables, and second would have \(p\) (momentum), \(n\) and \(v\) as variable. So we have 2 equations and 3 variables. Therefore, we use following closure relation

\[
p = m_e n v, \quad (1.13)
\]

which comes from assumptions of parabolic bands in the conduction band of the semiconductor. Also, such parabolic band structure points to the fact that particle distribution is Maxwellian, which can be written as

\[
f \propto \exp \left[ -\frac{|\vec{p} - m_e \vec{v}|^2}{2m_e k_B T_c} \right]. \quad (1.14)
\]
Therefore, we inadvertently assume that above HD equations are applicable for Maxwellian distribution only.

1.7 Summary of the Present Work

In this dissertation work, we will develop and demonstrate efficient time-domain models for full-wave-hydrodynamic modeling for plasmonic devices and phenomenology. The models account for the electromagnetic wave-propagation and the electron transport and therefore are multiphysical in nature. Our approach is use of electrodynamic equations for the wave-propagation and hydrodynamic equations for the modeling of the 2DEG channel. Hydrodynamic equations provide computationally inexpensive solution- especially advantageous since Maxwell’s equations already require computational resources for this modeling. Even with its macroscopic nature, the hydrodynamic model is quite accurate, as long as relaxation time ($\tau$) is modeled accurately. To that end, $\tau$ can either be extracted from measurements or from a separate particle simulations. Overall, use of hydrodynamic equations allows fairly accurate modeling, while providing computational efficiency to plasma-wave modeling.

We note that similar modeling method is also proposed recently in [58] for plasmonic modeling of FET/HEMT devices. Even before this work, different researchers have developed solution of same equations, targeting different applications [50,51]. However, simulation-times in these proposed methods could be very high due to small FDTD(finite difference time domain) time-steps that are needed for fine mesh-sizes.
Figure 1.7: (Research work presented in this dissertation in relation to some of the past research in this area)

In this work, we offset the time-cost issues by developing unconditionally stable FDTD algorithms coupled with hydrodynamic equations (Chapter-3). Specifically, we use Alternate-Directional-Implicit FDTD method which achieves fast simulation times at the cost of some accuracy [59]. We further employ iterative corrective methods to improving the solution-accuracy leading to time-efficient and accurate solutions. We also use the developed numerical models to study several plasma-wave device-concepts and phenomenon.

In Chapters 5 and 6, we employ the developed models for examination and optimization of plasma-wave devices. Specifically, we numerically examine terahertz emitters based on plasma-wave instability in short channel HEMTs (also published in [38]) and RTD-gated HEMTs (published in [60]). We also apply the developed numerical model to investigate the plasmonic modes in a double channel system (also referred to
These device-models are also verified by analytical and measurement results. For RTD-gated HEMT and double channel HEMT cases, we have developed our own analytical models to verify the device operation and numerical model.

A graphical illustration of the research accomplishments of this work in relation to work published in past is shown in Fig. 1.7. Organization of the chapters and contents is shown in Fig. 1.8.
Chapter 2: Full-wave-Hydrodynamic FDTD Solver

2.1 Motivations for Full-Wave-Hydrodynamic Model

Modeling of plasmonic phenomenon comprises of modeling of electromagnetic field propagation, along with the modeling of electron transport within the channel. Various approaches for such modeling can be grouped in three categories - 1) analytical-methods [39,45–47,60,62], 2) circuit based methods [63–65] and 3) numerical-methods [48,49,51,52,58,66].

Analytical approaches can be applied for canonical problems where dispersion relations can be derived by making simplifying assumptions. For example, wave-propagation constant can be calculated for infinitely large 2DEG sheet embedded in a dielectric media or metal backed infinitely large 2DEG sheet in dielectric media. In general, they cannot be used for a specific device geometry. Although, approximate estimates and qualitative trends can be evaluated [21,33,34]. Likewise, circuit approaches allow for the understanding of already known phenomenon using lumped elements. Although easy to use, circuit approaches can only be applied in specific scenarios and ultimately use the insights and information obtained from either full-wave/analytical modeling or measurements.
On the other hand, numerical approaches can be used to model specific device geometries more accurately. In past, these models have often employed Drude model for modeling the conductivity [67, 68] in doped and 2DEG regions of the devices. For active device modeling, where the electrons are drifting under the influence of external DC bias, Drude model cannot be applied. This is so because the terahertz conductivity of electrons under drift is fundamentally different from those of stationary electrons [47, 69]. Secondly, nonlinear effects such as velocity-saturation due to electron heating under high field bias are not considered by Drude model. Usually they can be accounted for by considering third hydrodynamic (or energy conservation) equation in the model. It is worth noting that several attempts of modeling the wave propagation using Poisson’s equation have also been made in past [66], but these do not provide a complete picture and insights of wave propagation within the devices.

Hydrodynamic (HD) equations coupled with Maxwell’s equation provides a method for active modeling, while also accounting for EM wave-propagation within the devices. As opposed to Drude model, the HD equations are solved numerically and are concurrently coupled to the Maxwell’s electrodynamic equations. Therefore, they can account for drifting electrons within the channel. In this chapter, we develop FDTD-hydrodynamic solver (referred to as FDTD-HD solver in this work) using traditional explicit-FDTD scheme (also known as Yee-FDTD scheme) coupled with upwind solution of hydrodynamic equations. Note that the algorithm used for this modeling is similar to that used in [58], which uses so-called Global Modeling method as previously developed by researchers in [48, 49, 51, 52].
Figure 2.1: Typical HEMT geometry with gate discontinuity for exciting the plasmonic oscillation in channel. The red arrows show the electric field lines due to plasmonic-excitation in the channel.

A typical device geometry modeled using the simulation model is shown in Fig. 2.1. As shown, a terahertz plane-wave is incident on a HEMT device with a small gate-discontinuity. The incident wave is diffracted due to the discontinuity, creating number of wave-vectors. Among these, the wave-vector that matches the plasma-wave modes in 2DEG-sheet couples with the 2DEG layer and propagates along the 2DEG channel. Using this model schematic, we will discuss the full-wave-hydrodynamic model using FDTD-HD algorithm in following sections.
2.2 Governing Equations

2.2.1 Hydrodynamic Modeling of the 2DEG Channel

We assume steady state conditions as initial condition for our solution. E.g. we begin our solution by assumption of 2DEG layer present in the dielectric media. In our analysis, we are focused on electrodynamic solution only, therefore we simply assume the presence of 2DEG in the dielectric.\(^1\) This is sufficient for our purposes, since electrodynamic field variations are decoupled from the DC bias conditions such as electron-drift due to gate-to-drain and drain-to-source biases. They only depend on outcomes of bias conditions such as, concentration \(n_{sh}\), electron drift velocity \(v_o\) etc. These are assumed as initial conditions in our solution.

For channel modeling, Boltzmann Transport equations (BTE) provides a complete description of the electron-transport phenomenon in a 2DEG channel. Since this solution is computer-memory and time expensive, we consider its first two moments for description of electron-transport, i.e. particle conservation (or continuity equation) and momentum conservation equations [51]. These equations collectively are called hydrodynamic (HD) equations. We also referred to these equations in Chapter-1, but they were written for a 3D electron gas system. We re-write them for a 2D confined electron gas as

\(^1\)A more complete solution can be pursued by solving Poisson equation in the 2D region to obtain the initial conditions. E.g. Non-linear Poisson equation (Poisson equation with band gap, electron-hole concentration considerations) in vertical direction can be used to derive the 2DEG sheet charge density [70]. Likewise, Poisson equation in horizontal direction provides the density distribution in lateral direction.
\begin{equation}
\frac{\partial n_{sh}}{\partial t} + \frac{\partial j}{\partial x} = 0, \quad \text{and} \quad \frac{\partial j}{\partial t} + v \frac{\partial j}{\partial x} + j \frac{\partial v}{\partial x} = -\frac{q n_{sh} E_x}{m_e} - \frac{j}{\tau} - \frac{K T}{m_e} \frac{\partial n_{sh}}{\partial x}.
\end{equation}

Here $n_{sh}$ is the sheet carrier density, $j(= n_{sh} v)$ is the sheet current and $v$ is the electron velocity within the 2DEG channel. $E_x$ is the $x$-directed electric field along the channel (see Fig. 2.1), $\tau$ refers to the momentum relaxation time and $m_e$ is the effective electron mass. $q = 1.6 \times 10^{-19}$ C is the charge-magnitude of a single electron. $T$ is the electron temperature in the channel and $K$ is the Boltzmann constant.

Equation 2.1 is the particle conservation equation, signifying the equality between rate-change of electrons and the gradient of the sheet-current at any given point in space. Equation 2.2 represent conservation of momentum under an externally applied force (given by the term $q n_{sh} E_x / m_e$). The term $j/\tau$ represents loss of momentum due to scattering within the channel. For our work, we use $\tau$ as phenomenological parameter, obtained either from experimental results or from separate particle simulations.

Third hydrodynamic equation, i.e. energy conservation equation, can also be considered for more complete solution [55]. But, we avoid its use for plasmonic applications to reduce solution’s complexity and its time-cost. Ignoring energy conservation automatically means that a thermal equilibrium is assumed between the electron gas and the lattice of the semiconductor material (or $T = T_l$ where $T_l$ is the lattice temperature). This is a valid assumption for small bias voltages. For high bias scenarios, electrons can gain energy under influence of strong electric fields and the temperature
$T$ of the electron-gas could rise. This introduces additional driving force proportional to the pressure gradient term ($nKT$), which would require the use of energy conservation equation for its modeling.

Finally, we also note that the solution of hydrodynamic equations can be pursued for both types of carriers, electrons and holes, existing in a FET or HEMT device. However, we will only consider majority carriers, i.e. electrons, and develop appropriate transport equations. The analysis can be extended for minority-carriers, if need be.

2.2.2 Electrodynamic Equations (or Maxwell’s Equation):

Field variations lead to variations in electron-density within the channel. Since electrons are charged particles, a plasma-mode can be excited leading to propagation of AC power within the channel. Since the channel is electrically thin ($\approx 5$-$15$ nm), these propagating fields interact with the media above and below the channel. For example, in Fig. 2.1, propagating plasma-wave has $E_x$ field components within the channel, but also exhibits electrodynamic fields outside the channel, attenuating in $\hat{y}$ direction away from the channel. This field is modeled using Maxwell’s electrodynamic equations solved self-consistently with HD equations. This AC modeling of fields can be done using

$$\nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J} + \sigma \vec{E} \quad \text{and} \quad (2.3)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} - \sigma^{*} \vec{H} \quad (2.4)$$

For 2D solution, we consider electric field intensity as $\vec{E} = \hat{x}E_x + \hat{y}E_y$. Likewise, $\vec{H}$ is the magnetic field intensity, $\vec{D} = \epsilon \vec{E}$ and $\vec{B} = \mu \vec{H}$ are electric and magnetic
flux densities—all defined as 2D vectors. These symbols carry same meaning through this work. Permittivity $\epsilon$ and permeability $\mu$ are assumed to be scalar, since we are only interested in isotropic media for our problems. Electrical and magnetic conductivity are represented by $\sigma$ and $\sigma^*$ (also referred to as $\sigma_m$) respectively. We use the conductivity terms for implementing perfectly matched layer (PML) boundary conditions [71] around the device for absorptive and reflection-less termination of the simulation domain. Source term $\vec{J}$ is due to the current generated by the oscillations caused by the plasma waves as per 2.1 and 2.2. This term is related to current in equation 2.3 by

$$\vec{J} = \frac{-q(j - j_o)}{t_{2\text{DEG}}} \hat{x}, \quad (2.5)$$

where, $t_{2\text{DEG}}$ is the thickness of the 2DEG layer. Also, $j - j_o$ represents the AC current responsible for radiation from the channel, with $j_o$ being the initial DC current in the channel that does not contribute to electromagnetic radiation.

### 2.3 Finite Difference Time Domain (FDTD) Solution:

For self-consistent solution of the two sets of equations, we use finite difference time domain (FDTD) approach [71]. Although the algorithm is modified to add the simultaneous solution of HD equations using upwind scheme [47]. For each time iteration, we calculate the values of currents and fields for next time steps using discretized time and space difference equations (as also further detailed in next section). For coupling between the two solvers, the HD and full-wave equation solvers are run in tandem with careful exchange of variables within each time-step cycle. Specifically, within each time-step, current value ($j$) updated from the HD solver is used in the full-wave
Figure 2.2: Device cross-section and corresponding modeling equations for predicting the $\vec{E}$-fields in the media and the 2DEG AC currents within the channel. EM and transport equations are solved in time iterations. $E_x$ fields as calculated from Maxwell’s equations are applied to hydrodynamic equations and channel AC current $J_e$ as calculated from the hydrodynamic equations is applied to the Maxwell’s equations within each time iteration.

solver as source term. In return, field component $E_x$, updated by the full-wave solver, is fed back into the HD solver. Thus, oscillating current within the 2DEG and associated electric field values outside are consistent with each other. This process is repeated until desired simulation time is reached. The scheme is illustrated in Fig. 2.2.

The HD equations are solved in time domain using upwind scheme. Therefore, both solutions being in time domain, FDTD scheme is easily coupled with HD equation. We only need to model the cross-sections of the device. Therefore, a 2D FDTD solver would be sufficient. We used Yee’s time-space 2-D staggered grid [71] for discretization
of the device cross section. 2DEG layer was modeled using a 1-D grid of the electrons sandwiched within the dielectric (Fig. 2.3).

![2D Meshing Scheme](image)

Figure 2.3: Modeling of the HEMT’s cross-section using the developed hybrid numerical algorithm. Non-uniform meshing allows faster simulation times. The 2DEG channel is modeled by transport equations by employing a 1D mesh while EM fields are solved using a 2D mesh in the device cross-section. Typical cell sizes and time step values are $\Delta x = 5$ nm, $\Delta y = 1$ to 5 nm, $\Delta t = 10^{-17}$ s.

### 2.3.1 Time-Space Discretization of HD Equations

For the numerical solution, the partial derivative Equations 2.1-2.4 are converted into their corresponding time and space difference equations. We first consider the discretization of HD equations. It is known that HD equations are nonlinear and are highly coupled advection equations [55]. Their numerical solutions could be unstable when center difference approximation is used. This is especially true if we consider cases with strong DC bias across the channel. Therefore, upwind difference scheme was applied to maintain the stability of the algorithm [55,72]. Thus, we write the difference equations for 2.1 and 2.2 as
\[ n_{sh}^{i+1} = n_{sh}^{i} - \frac{\Delta t}{\Delta x} [j_{i+1/2}^n - j_{i-1/2}^n], \quad (2.6) \]

and
\[
j_{i+1/2}^{n+1} = j_{i+1/2}^n - \Delta t q n_{sh}^{i+1/2} \frac{E_{ch}^{i+1/2}}{m_e} - \Delta t \frac{j_{i+1/2}^n}{\tau_n} - \Delta t \frac{K_T}{m_e} \left[ n_{sh}^{i+1} - n_{sh}^i \right] \Delta x - Conv_{i+1/2}^n \quad (2.7)\]

where, for upwind scheme we define the convective term, \( Conv \), based on the direction of flow of electrons at that point [72].

For \( v_{i+1/2}^n > 0 \), we define
\[
Conv_{i+1/2}^n = \frac{\Delta t}{\Delta x} j_{i+1/2}^n \left[ v_{i+1/2}^n - v_{i-1/2}^n \right] + \frac{\Delta t}{\Delta x} v_{i+1/2}^n \left[ j_{i+1/2}^n - j_{i-1/2}^n \right] \quad (2.8)
\]

and for \( v_{i+1/2}^n < 0 \), we define
\[
Conv_{i+1/2}^n = \frac{\Delta t}{\Delta x} j_{i+1/2}^n \left[ v_{i+3/2}^n - v_{i+1/2}^n \right] + \frac{\Delta t}{\Delta x} v_{i+3/2}^n \left[ j_{i+3/2}^n - j_{i+1/2}^n \right] \quad (2.9)
\]

In above equations, \( i \) refers to the grid-point index along \( \hat{x} \)-direction and \( n \) represents the time iteration index. Time-step is \( \Delta t \) and mesh step-size along \( \hat{x} \)-direction is \( \Delta x \). Note \( \Delta x \) could be non-uniform outside the channel but is chosen to be uniform across the channel. More details on upwind modeling of the hydrodynamic equations can be referred from [72].

### 2.3.2 Time-Space Discretization of Maxwell’s Equation

Next, we consider the discretized version of electrodynamic equations. Since the \( E \)-field perturbations in the channel is along the \( \hat{x} \)-direction only, we only expect field
variations in $E_x$, $E_y$ and $H_z$ field components and therefore modeling a $TE_z$ (TE to the $\hat{z}$-direction) field profile is sufficient. Removing other field components and by using first order space and time differences approximations \cite{71}, we get

\begin{equation}
E_y|_{i+1/2,j}^{n+1} = E_y|_{i+1/2,j}^n \left( \frac{1 - \sigma_x \Delta t/2 \epsilon}{1 + \sigma_x \Delta t/2 \epsilon} \right) - \frac{\Delta t}{\Delta x \epsilon} \left[ H_z|_{i+1,j}^{n+1/2} - H_z|_{i,j}^{n+1/2} \right] \tag{2.10}
\end{equation}

\begin{equation}
E_x|_{i,j+1/2}^{n+1} = E_x|_{i,j+1/2}^n \left( \frac{1 - \sigma_y \Delta t/2 \epsilon}{1 + \sigma_y \Delta t/2 \epsilon} \right) + \frac{\Delta t}{\Delta y \epsilon} \left[ H_z|_{i,j+1}^{n+1/2} - H_z|_{i,j}^{n+1/2} \right] - \frac{\Delta t J_z}{\epsilon} \tag{2.11}
\end{equation}

\begin{equation}
H_{zx}|_{i,j}^{n+1/2} = H_{zx}|_{i,j}^{n-1/2} \left( \frac{1 - \sigma_x \Delta t/2 \mu}{1 + \sigma_x \Delta t/2 \mu} \right) - \frac{\Delta t}{\Delta x \mu} \left[ E_y|_{i+1/2,j}^n - E_y|_{i-1/2,j}^n \right] \tag{2.12}
\end{equation}

\begin{equation}
H_{zy}|_{i,j}^{n+1/2} = H_{zy}|_{i,j}^{n-1/2} \left( \frac{1 - \sigma_y \Delta t/2 \mu}{1 + \sigma_y \Delta t/2 \mu} \right) + \frac{\Delta t}{\Delta y \mu} \left[ E_x|_{i,j+1/2}^n - E_x|_{i,j-1/2}^n \right] \tag{2.13}
\end{equation}

As before, indices $i$ and $j$ refer to cell numbers along $\hat{x}$ and $\hat{y}$ directions, and $n$ is the time iteration index. \( \epsilon = \epsilon_r \epsilon_o \) and \( \mu \) are respectively the permittivity and permeability of the media having a dielectric constant \( \epsilon_r \). \( \sigma_x, \sigma_y \) and \( \sigma_x^*, \sigma_y^* \) are the electric and magnetic conductivity of the media for the waves traveling in $x, y$-directions, respectively. For the implementation of the PML boundary conditions, we model the PML layer as a hypothetical anisotropic media, which requires such
separation of the conductivity for $x$ and $y$-propagating waves [71]. This is also the reason we have two equations for updating the magnetic fields, 2.12 and 2.13.

We start with an assumed channel sheet electron density and electron drift velocity given by $v_o = -\frac{\tau q V_{ds}}{L_c m_e}$, where $V_{ds}$ is the externally applied drain to source voltage, $L_c$ is the length of the channel. Using the obtained initial DC conditions in the channel, i.e. $n_{sh}$ and $v$, update-equations 2.6 to 2.13 are sequentially executed. Therefore, sheet current obtained from HD equations is the source for radiation and used in Maxwell’s equations. Then Maxwell’s equations are used to update the field values. The updated $E_x$ component from the Maxwell’s solver is used in the HD solver for next iteration. This process is repeated till the desired simulation time is reached.

The solver’s time step is restricted by the Courant-Friedrich-Levy (CFL) condition [71]. For 2D solution, the condition states that

$$
\Delta t \leq \frac{1}{c \sqrt{(\Delta x_{min})^2 + (\Delta y_{min})^2}},
$$

where, $\Delta x_{min}$ and $\Delta y_{min}$ are the minimum step sizes in the grid. Since the plasma wavelengths are much smaller that the free-space wavelengths ($<100$ times), $\Delta x$ in the 2DEG region must be approximately 1 to 5 nm. This makes the time step small ($10^{-17}$ s), increasing the total simulation times. Nevertheless, simulation-times can be reduced to some extent by reducing the number of cells by the use of non-uniform meshing, as shown in Fig. 2.3. We will further solve the time-cost issues using unconditionally stable FDTD algorithms as discussed in Chapter-3.
Figure 2.4: The schematic used for the simulation for comparison with analytical formulae (schematic not to scale).

2.4 Verification using Analytical-Models and Experimental Data

2.4.1 Model Validation via Analytical method

For validation, we use the developed numerical model to find the plasmonic wavelength and wave velocity for two canonical problems. We examine the plasmonic propagation phenomenon for 1) a cases when a 2D electron gas layer is embedded in dielectric media and 2) a case when the 2DEG layer is placed next to a metal gate and is embedded in dielectric media The schematic for these cases are shown in Fig. 2.5(a) and (b). Here, we do not model an actual device (e.g. a HEMT), so that
the model truly represents the analytical formula and approximations made in their
derivations. Thus, a comparison can be made for model validation.

The schematic used for the modeling of the described cases is shown in Fig. 2.4. As
shown, a dielectric media with dielectric constant $\epsilon_r = 13.9$ (for GaAs semiconductor)
is assumed. The channel carrier density was set to $n_{sh} = 10^{13}$ cm$^{-2}$ and the channel-
length was $L_c = 6 \mu m$, large as compared to the expected plasma wavelength in the
frequency band under consideration. We use a 2DEG channel thickness of $t_{2DEG} =
10$ nm for these simulations. The mesh size near the channel was chosen to be $\Delta x
= 5$ nm and $\Delta y = 10$ nm. In vertical direction, we increased the mesh size in steps
to $\Delta y = 10$ nm, 50 nm, 100 nm and 600 nm as we moved away from the channel
in either direction. 600 nm was the mesh size of the PML region which was 50 $\mu m$
thick. In horizontal direction, the mesh within the 2DEG region was chosen to be
$\Delta x = 5$ nm. Away from the 2DEG, on either side, the mesh size was increased to 10
nm, 100 nm, 200 nm and then to 700 nm in the PML region.

By introducing a time harmonic $E_x$ field, of frequency $f$, at the point $x = 4 \mu m$, a
2D plasma-wave is excited within the 2DEG channel. The same excitation process
was again used for the case when the 2DEG layer is covered with a metal gate with a
barrier of thickness $d_{barr} = 38$ nm. As seen in Fig. 2.5(a) and (b), a plasma wave prop-
agating away from the point of excitation was observed. The resulting propagation
constant ($k = 2\pi/\lambda$) and phase-velocity ($v_p = \omega/k$), calculated for varying frequency
excitation, is plotted in Fig. 2.5(c). For ungated 2DEG layer, the dispersion relation
from [45, 47, 73], given by,
\[ k = \frac{1}{2a} \sqrt{\omega^4 + \frac{4a^2 \omega^2 \epsilon_r}{c^2}}, \quad (2.15) \]

is also plotted. Here, \( a = \frac{n_{sh} q^2}{4 \alpha_0 \epsilon_r m_e}, \quad \epsilon_o = 8.85 \times 10^{-12} \text{F/m} \) and \( \omega \) is the angular frequency. It is seen in Fig. 2.5(c) that the numerical data is in agreement with analytical model in (2.15). Next, for gated case, the dispersion relation is given by [46]

\[ k = \frac{\omega^2 m_e}{n_{sh} q^2} (\epsilon_o \epsilon_r + \epsilon_o \epsilon_b \coth kd_{barr}) \quad (2.16) \]

As known, \( \epsilon_b (= \epsilon_r) \), is the dielectric constant of the barrier layer. Again, the agreement between (2.16) and the numerical data is quite good.

We observe that for gated case, there is strong \( E_y \) field and much weaker \( E_x \) field between the gate and the 2DEG layer (refer field plots in Fig. 2.5(b)). This is especially true for small \( d_{barr} \), i.e. \( kd_{barr} \ll 1 \). In such cases, the 2DEG and gate system acts like a parallel plate waveguide, supporting a TEM mode. Therefore, due to this field profile, we expect a linear relation between the frequency \( \omega \) and the propagation constant \( k \), a characteristic of the TEM mode propagation. Indeed, for \( kd_{barr} \ll 1 \), (2.16) reduces to \( k = \omega \sqrt{\frac{m_e \epsilon_o}{n_{sh} q^2 d_{barr}}}, \) verifying this linear relation.

2.4.2 Model Validation via Prior Measurements

Another validation of the developed model was conducted by comparing the results from a prior measurement. We consider a grating-gated HEMT device (shown in Fig. 2.6) with experimentally observed resonances [74]. These resonance modes occur due
to periodic boundary conditions introduced by the grating-gates and roughly occur at frequencies given by

\[ f = \frac{1}{L} \sqrt{\frac{n_{sh} q^2 d}{m_e \epsilon_o \epsilon_r}} \] (2.17)

In the above, \( d \) is the thickness of the barrier (distance from the metal gate to the channel) and \( L \) is the periodicity of the grating-gate. The formula (2.17) assumes that: 1) channel is completely covered with metal gate, 2) the 2DEG thickness is infinitely small, and 3) \( d \) is much smaller than the plasma-mode wavelength. Due to these approximations in the derivation of (2.17), it is only an approximate representation of the resonance frequencies for a grating-gated HEMT. Therefore, we used our developed full-wave-hydrodynamic model to predict the transmission spectra. In Fig. 2.6, we compare the simulation results with that of measurements conducted in [74].

The experiment in [74] was conducted using an AlGaN/GaN HEMT sample of size 1.7 \( \times \) 1.6 mm\(^2\), with \( L = 1.5 \) \( \mu \)m and \( W = 1.2 \) \( \mu \)m (as shown in Fig. 2.6). That is, more than 1000 fingers (gratings) were used. Also, the electron density in the channel was \( n_{sh} = 7.5 \times 10^{12}/\text{cm}^2 \) and barrier thickness was \( d = 30 \) nm. Notably, transmission spectra in the terahertz frequency range are recorded for varying temperature. As expected, at lower temperatures, the plasma-resonances are sharp due to high electron mobility \( \mu \) (and high relaxation time \( \tau \)). At room temperature, these resonances are broadened due to the reduced channel mobility. These affects are also modeled in our simulations as shown in the plots of Fig. 2.6.
As shown, there is an excellent agreement between measured and simulated data. Specifically, the position of resonance frequencies and shape of transmission spectra are well predicted by the developed numerical algorithm. The simulations are able to predict higher order resonances which are present in the measurements as well. We observe some differences between the simulation and measured data towards the higher frequency band, but these differences can arise from the fabrication tolerances in the sample. Another reason contributing to this variation is the finite size of the sample which is comparable to incident THz wavelength. In simulations, the sample is assumed to infinitely large due to use of periodic boundary conditions employed for efficient device modeling.
Figure 2.5: Validation of the full-wave hydrodynamic model. The parameters used for electron gas modeling are $\epsilon_r=13.9$, $m_e=0.042 m_o$, $n_{sh}=1 \times 10^{13} \text{cm}^{-2}$, for gated case $d_{barr}=38 \text{ nm}$, $\mu=10,000 \text{ cm}^2/(\text{V s})$ (a) Electric field distribution (in V/m) due to plasma propagation in a ungated 2DEG layer (b) Electric field distribution (in V/m) due to plasma propagation in a gated 2DEG layer. (c) Dispersion diagrams calculated from the developed numerical model versus theoretical predictions from [46], [47].
Figure 2.6: Geometry and resonances associated with the grating-gated HEMT. Top: Modulation of THz spectrum using plasma-resonances within the HEMT. Bottom: Computed and measured plasma-wave resonances. Computations are based on the proposed hybrid algorithms. The measurements are from [74] and serve to validate the developed numerical model.
Chapter 3: Efficient Time-Domain Modeling of Plasmonic Devices

3.1 Motivations for ADI and Iterative-ADI-FDTD Based Full-Wave-Hydrodynamic Solvers

In the previous chapter, we developed full-wave-hydrodynamic method using explicit FDTD scheme for the modeling of plasma-wave devices. Despite its accuracy and rigor, this approach requires from long simulation times because of small time-steps dictated by Courant-Friedrich-Levy (CFL) condition. These small-time steps are ultimately dictated by the fine mesh needed for accurate plasmonic modeling and therefore are unavoidable.

Plasmonic-oscillations exhibit strong spatial field-variations near the 2DEG channel and evanescent-field variations vertically-away from the channel. Specifically, the plasma-wave wavelengths are up to two orders smaller than the free-space wavelength, requiring around 5000-10000 cells per free-space-wavelength. At terahertz frequencies, this translates to a time-step $\Delta t$ of around $10^{-17}$ s. To conduct a simulation for up to several picoseconds, tens of thousands of iterations are required.
To avoid these time-cost issues, so-called unconditionally-stable algorithms are adopted. Alternate-Directional-Implicit FDTD (ADI-FDTD) method [59, 75–81], is one such method. This approach uses a splitting-operator on electrodynamic equations to yield an implicit form of difference equations [59]. Notably, these equations are tridiagonal system of linear equations that can be solved at a small computational cost. In the meantime, the time-steps can be arbitrarily large reducing the simulation-times. The method has been applied for several traditional EM problems, but its applications for electron-plasma wave related application is yet to be demonstrated.

In this chapter, we will use ADI-FDTD algorithm for electrodynamic equations coupled with upwind algorithm for hydrodynamic equations for time-efficient full-wave-hydrodynamic modeling for plasmonic applications. We call this newly developed approach as ADI-FDTD-HD method. Note that similar ADI-FDTD based algorithms have been used for fast modeling of microwave transistors in [82], but they have not been considered for plasmonic applications prior to this work. The model uses the unconditional stability to attain time-efficiency even for refined mesh scenarios in our simulation model. Although, it is limited in its accuracy due to truncation of second order terms (which is required to achieve a tridiagonal system of linear implicit equations). Notably, these errors (so-called splitting errors or truncation errors) could become dominant in the regions with larger spatial derivatives [75, 81], i.e. regions near the 2DEG channel.

Therefore, to address the accuracy of ADI-FDTD-HD method, we will use iterative-ADI-FDTD method [80] coupled with upwind scheme for the solution of the HD equations. This developed approach is abbreviated as it-ADI-FDTD-HD method.
We will show that the method improves the accuracy of ADI-FDTD-HD model while maintaining time-efficiency over the traditional FDTD-HD method that was discussed in previous chapter.

We will consider ADI-FDTD-HD and it-ADI-FDTD-HD models by first discussing the underlying equations and their corresponding FDTD implementations. The performances of these methods will be demonstrated by modeling the plasmonic oscillations in a GaN/AlGaN HEMT device (shown in Fig. 3.3). Finally, the results for these implementations will be compared with a reference model for validation, accuracy evaluation and time-cost analysis.

3.2 Governing Equations for Unconditionally Stable FDTD-HD Methods

Underlying principle of the ADI-FDTD and iterative-ADI-FDTD methods involves splitting of single time-step into two sub time-steps and evaluation of implicit equations in each of the sub time-steps. In ADI-FDTD method, a second order term is ignored for fast solution. Whereas, in iterative-ADI-FDTD method, we also use an added iterative term to account for this approximation made in the ADI-FDTD. We will consider the governing equations for these methods in following subsections.

3.2.1 ADI-FDTD-HD Method

We start with Maxwell’s electrodynamic equations for a TE$_z$ solution. We account for the existence of PMLs in the simulation-domain by modeling the anisotropic
conductivity in the media. Following the analysis shown in [80], we write
\[
\frac{\partial \vec{u}}{\partial t} = [A] \vec{u} + [B] \vec{u}. \tag{3.1}
\]
Here, operators \([A]\) and \([B]\) are
\[
[A] = \begin{bmatrix}
-\frac{\sigma_y}{2\epsilon} & 0 & \frac{1}{\epsilon} \frac{\partial}{\partial y} & \frac{1}{\epsilon} \frac{\partial}{\partial y} \\
0 & -\frac{\sigma_x}{2\epsilon} & 0 & 0 \\
0 & 0 & -\frac{\sigma_x^*}{2\mu} & 0 \\
\frac{1}{\mu} \frac{\partial}{\partial y} & 0 & 0 & -\frac{\sigma_x^*}{2\mu}
\end{bmatrix}, \quad
[B] = \begin{bmatrix}
-\frac{\sigma_y}{2\epsilon} & 0 & 0 & 0 \\
0 & -\frac{\sigma_x}{2\epsilon} & -\frac{1}{\mu} \frac{\partial}{\partial x} & -\frac{1}{\mu} \frac{\partial}{\partial x} \\
0 & 0 & -\frac{\sigma_x^*}{2\mu} & 0 \\
0 & 0 & 0 & -\frac{\sigma_x^*}{2\mu}
\end{bmatrix}, \tag{3.2}
\]
with fields given by matrix \(\vec{u}\) (note that here the vector sign represents a matrix, not a space-vector) as
\[
\vec{u} = \begin{bmatrix}
E_x \\
E_y \\
H_{xx} \\
H_{xy}
\end{bmatrix}. \tag{3.3}
\]
Here H-field splitting implies \(H_z = H_{xx} + H_{xy}\). For time discretization, we apply Crank Nicholson scheme at time step \(n + 1/2\) to obtain
\[
\left( I - \frac{\Delta t}{2} [A] - \frac{\Delta t}{2} [B] \right) \vec{u}^{n+1} = \left( I + \frac{\Delta t}{2} [A] + \frac{\Delta t}{2} [B] \right) \vec{u}^n. \tag{3.4}
\]
Here, \(\Delta t\) is the time-step used for difference equations. Above can be re-written in factorized form as
\[
\left( I - \frac{\Delta t}{2} [A] \right) \left( I - \frac{\Delta t}{2} [B] \right) \vec{u}^{n+1} = \left( I + \frac{\Delta t}{2} [A] \right) \left( I + \frac{\Delta t}{2} [B] \right) \vec{u}^n + \frac{\Delta t^2}{4} [A][B] \left( \vec{u}^{n+1} - \vec{u}^n \right). \tag{3.5}
\]
In the above a second-order term is present in addition to the factorized terms. Ignoring this term provides us
\[
\left( I - \frac{\Delta t}{2} [A] \right) \left( I - \frac{\Delta t}{2} [B] \right) \vec{u}^{n+1} = \left( I + \frac{\Delta t}{2} [A] \right) \left( I + \frac{\Delta t}{2} [B] \right) \vec{u}^n, \tag{3.6}
\]
which can be solved in exactly two steps, i.e.,
\[
\left( I - \frac{\Delta t}{2} [A] \right) \vec{u}^{mp} = \left( I + \frac{\Delta t}{2} [B] \right) \vec{u}^n, \quad \text{and} \tag{3.7}
\]
\[
\left( I - \frac{\Delta t}{2} [B] \right) \vec{u}^{n+1} = \left( I + \frac{\Delta t}{2} [A] \right) \vec{u}^{tmp}.
\] (3.8)

Expanded difference equations derived from the above equations are provided in Appendix-A. Here \(\vec{u}^{tmp}\) is intermediate solution, denoting field values at sub-time-step. Equations 3.7 and 3.8 are tridiagonal system of linear equations that can be solved using Gauss-Seidel method. Thus, no matrix inversion is needed for the solution of these equations.

Figure 3.1: Flowchart for ADI-FDTD algorithm coupled with hydrodynamic equations. As noted, the evaluation of electrodynamic fields is split in two sub-time-steps. The hydrodynamic equations may be evaluated at a smaller time-step \(\Delta t/R\) for the cases of large \(\Delta t\) (or \(CN > 500\))
Further, these equations are unconditionally stable as shown in the analysis conducted by Namiki in [59]. Therefore, the time-step can be arbitrarily increased to any \( \Delta t = CN \times \Delta t_{FDTD} \), where \( CN \) denotes so-called Courant number and \( \Delta t_{FDTD} \) is the time-step for traditional explicit FDTD algorithm. We will use \( CN \) in the following sections to represent the increase in time-step for our analysis.

Although the solution of electrodynamic equations is modified using implicit method, there are no changes in the solution of HD equations and their coupling to the electrodynamic solver. The only change experienced by the upwind equation solver is increase in the time-step due to the use of unconditionally stable method for electrodynamic equations. With regards to that, solution of hydrodynamic equations was found to be robust for increased time-steps. But for very large time-steps (\( CN > 500 \)), the upwind scheme becomes unstable. For such cases, we can choose smaller time-step as defined by \( \Delta t_{HD} = \Delta t/R \), where \( R > 1 \) is specified to bring the upwind solver in stable regime \(^2\). Thus, the upwind equation solver runs \( R \)-times for each \( \Delta t \) time step of the FDTD solver. Of course, the exchange of variables \( J_e \) and \( E_x \) between the two solvers occurs at every \( \Delta t \) (not every \( \Delta t_{HD} \)). Even so, the upwind solution remains stable due to this smaller choice of \( \Delta t_{HD} \). Fig. 3.1 describes the flowchart for the ADI-FDTD-HD approach.

\(^2\)Further studies and/or literature-survey is needed to explore the stability conditions for the upwind scheme with increasing time-steps. Since our simulations did not use \( CN > 500 \) cases, this study was not pursued in this work.
Figure 3.2: Flowchart for iterative ADI-FDTD algorithm coupled with hydrodynamic equations. Fields are evaluated using ADI-FDTD with iterations in which error term is updated based on fields obtained from last iteration. The HD equations are evaluated at a smaller time-step (if need be for larger CN) to maintain stability of the algorithm.

### 3.2.2 Iterative-ADI-FDTD-HD Method

As noted in previous sections, to reach the tridiagonal system of equations, we ignored the second order term

\[
\frac{\Delta t^2}{4} [A][B]\left(\bar{u}^{n+1} - \bar{u}^n\right)
\]

in 3.5. If we expand this matrix term, we find that it contains second order space-derivatives of the field, which become dominant near the 2DEG channel where the plasma-wave fields vary along the channel. Therefore, there are significant errors in the solution due to this ignored term. In iterative-ADI-FDTD method, we reconstruct
this term by using an iterative process within each time step [80]. Specifically, the splitting equations 3.7 and 3.8 are re-written to read

\[
\begin{align*}
(I - \Delta t \frac{1}{2}[A]) \vec{u}_{k+1}^{tmp} &= \left( I + \Delta t \frac{1}{2}[B] \right) \vec{u}^{n} + \frac{\Delta t^{2}}{8}[A][B]\left( \vec{u}_{k+1}^{n+1} - \vec{u}^{n} \right) \quad \text{and} \quad (3.10) \\
(I - \Delta t \frac{1}{2}[B]) \vec{u}_{k+1}^{n+1} &= \left( I + \Delta t \frac{1}{2}[A] \right) \vec{u}_{k+1}^{tmp} + \frac{\Delta t^{2}}{8}[A][B]\left( \vec{u}_{k}^{n+1} - \vec{u}^{n} \right). \quad (3.11)
\end{align*}
\]

In the above, we have introduced another variable \( k \) that denotes the iteration number within a single-time step calculation. We denote the number of iteration by parameter \( it \). Thus, the term \( \vec{u}_{k+1}^{tmp} \) for \( k+1 \)st iteration is calculated \( it \) times, using the field values from the previous iteration \( \left( \vec{u}_{k}^{n+1} - \vec{u}^{n} \right) \). By choosing a different initial guess and/or a few ADI-FDTD iterations for the solution, the splitting-error can be controlled. In our simulations, we chose the initial guess to be \( \left( \vec{u}_{0}^{n+1} = \vec{u}^{n} \right) \). As before, the derived difference equations corresponding to 3.10 and 3.11 are provided in Appendix A.

Iterative-ADI-FDTD method seeks to reduce the splitting-error in the ADI-FDTD method using a chosen number of iterations within each time-step. Of course, the simulation-time increases linearly with the number of iterations. But each time the error-term reduces and ADI-FDTD solution converges to explicit-FDTD solution. Overall, the method provides a way to reach a compromise between the time-cost and the accuracy by adjusting the number of iterations. In the following sections, we will also examine this tuning by varying the number of iterations and Courant number \( (CN) \) for iterative ADI-FDTD method coupled with HD equations.

The method for the solution of HD equations and coupling to the Maxwell’s equations remains same as explained in previous section. Flowchart, illustrating the complete algorithm, is shown in Fig. 3.2.
3.3 Model Implementation and Performance Evaluation

Next, we will examine the performance of the developed time-efficient algorithms for terahertz plasmonic applications. As noted, the chosen time-step $\Delta t = CN \times \Delta t_{FDTD}$ can be increased by varying the Courant number $CN$. For ADI-FDTD-HD case, with increasing $CN$, the simulation times are proportionately decreased. However, accuracy of solution is compromised due to dominant splitting-errors that are proportional to $\Delta t^2$ (refer equation 3.9).

These splitting-errors can be reduced by using it-ADI-FDTD-HD method by increasing the number of iterations (denoted by $it$). This again increases the simulation time by a factor of $it$. Therefore, for a given simulation domain, a compromise between the $CN$ and $it$ exists, i.e. time and accuracy can be optimized by tuning these parameters.

To understand these effects in detail, we model the plasmonic oscillations within a HEMT using these methods. Here, we consider an ungated AlGaN/GaN HEMT with a small gate-discontinuity for the coupling of an incoming terahertz wave. The geometrical dimensions of the device along with the schematic of the simulation domain are shown in the Fig. 3.3. As shown, a plane-wave of chosen frequency of 5 THz is incident on the device. The parameters associated with the 2DEG channel were chosen as $n_{sh} = 5 \times 10^{12} \text{ cm}^{-2}$, $\epsilon_r = 9.5$, $m_e = 0.2 m_o$ and $\tau = 1.14$ ps. The thickness of the 2DEG channel which was also the smallest cell size in vertical direction ($\Delta y$) was chosen to be 1 nm. Note that, this chosen dimension is smaller than the usually found 2DEGs in practical heterojunctions (which is around 4-10 nm), but thin channel is
so-chosen to demonstrate the algorithm for the case of refined mesh sizes. Furthermore, for applications with graphene, the 2DEG thicknesses are smaller than 1 nm. The horizontal cell size was chosen as $\Delta x = 4$ nm along the 2DEG channel. This led to an overall number of grids to be 1460 cells in horizontal direction and 740 in the vertical direction in a rectangular grid. Based on these chosen $\Delta x$ and $\Delta y$, values, using the CFL condition, the maximum allowable time-step $\Delta t_{FDTD}$ was $3.2 \times 10^{-18}$ s. The simulation was conducted from $t = 0$ to 3 ps.

![Diagram](image)

Figure 3.3: The device along with various dimension (left) along with the simulation domain (right) used for the modeling and verification and comparison of the ADI-FDTD based full-wave-hydrodynamic solvers.

### 3.3.1 Performance of the ADI-FDTD-HD Method

Using the device model described in previous sections, simulation was conducted using traditional FDTD-HD method and ADI-FDTD-HD method. For ADI method, $CN$ values of 100, 200 and 300 were chosen for the simulations. The cell size, time-step size and total simulation times for 3 ps of simulation for these cases are tabulated.
Figure 3.4: E-field profiles (calculated as $E = \sqrt{E_x^2 + E_y^2}$ obtained at $t=3$ ps for the various cases of ADI-FDTD-HD method compared with the reference solution. Here, the reference case corresponds to explicit FDTD-HD method.

in the Table 3.1. As observed in the table, owing to 100 to 300 times increase in the time-step size, the number of time-iterations are decreased. Therefore, even with each time-step taking more CPU-time for ADI-FDTD-HD method, overall CPU times are considerably reduced for the simulation. Note that the simulation models were implemented and executed in MATLAB.

However, the reduction in CPU times comes at a price of reduced accuracy. For $CN=300$ case, the accuracy of solution is unacceptably low. This is shown in Fig. 3.4 using the 2D E-field plots for these cases, as recorded at the end of the simulation of 3 ps. As shown, the $CN=100$ case shows almost no difference when compared to reference case. But as we move to $CN=200$ and 300 cases, there are appreciable differences in the field profiles, specially under the metal-gate region of the device. This is expected, since the inaccuracies are dominant in the regions with large spatial
Table 3.1: CPU-times for various $CN$ value cases for ADI-FDTD-HD method

<table>
<thead>
<tr>
<th>Method</th>
<th>$\Delta x$, $\Delta y$</th>
<th>$\Delta T$ (s)</th>
<th>CPU-time per time-step (s.)</th>
<th>CPU-TIME total (Hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDTD-HD</td>
<td>4 nm, 1nm</td>
<td>$3.24 \times 10^{-18}$</td>
<td>0.35</td>
<td>90</td>
</tr>
<tr>
<td>ADI-FDTD-HD, $CN=100$</td>
<td>4 nm, 1nm</td>
<td>$3.24 \times 10^{-16}$</td>
<td>7.59</td>
<td>19.57</td>
</tr>
<tr>
<td>ADI-FDTD-HD, $CN=200$</td>
<td>4 nm, 1nm</td>
<td>$6.48 \times 10^{-16}$</td>
<td>7.59</td>
<td>9.79</td>
</tr>
<tr>
<td>ADI-FDTD-HD, $CN=300$</td>
<td>4 nm, 1nm</td>
<td>$9.72 \times 10^{-16}$</td>
<td>7.59</td>
<td>6.52</td>
</tr>
</tbody>
</table>

Figure 3.5: Current profile in the channel at $t=3$ ps using the reference FDTD-HD method and ADI-FDTD-HD methods.

field derivative. For a closer look, we also examine the current in the channel for varying $CN$ values using 1D plots. The comparative current plots are shown in Fig. 3.5. As shown, the calculated current spuriously exhibits decreased plasma-wave wavelength for larger $CN$ values. Note that even for $CN = 100$ case, where the fields seemed to be accurate by visual inspection, there are significant errors in the calculated current. Furthermore, spurious oscillations are present near the gate region at $x = 1.2 \ \mu m$ for $CN = 200$ case.
Figure 3.6: E-field profiles obtained at $t=3$ ps for the various cases of it-ADI-FDTD-HD method compared with the reference solution. Here, the reference case corresponds to explicit FDTD-HD method. Increasing values of $it$ are needed to correct the splitting-errors as $CN$ is increased.

We conclude that care should be taken while using ADI-FDTD-HD method for the plasmonic applications where spatial field derivative can affect the accuracy of the solution. It at all used, smaller $CN$ should be used to maintain the integrity of the solution. Furthermore, possibility also exists for using higher $CN$ values, with iterative correction using the it-ADI-FDTD-HD for these problems. This is examined in next section.

### 3.3.2 Performance of it-ADI-FDTD-HD Method

Next, we use the same simulation cases as examined in previous sections, but this time several iterations are used using the algorithm outlined in Section 3.2.2. For each $CN$ case, the number of iterations were increased from 2 to up till we obtained
Table 3.2: CPU-times for various $CN$-$it$ combinations for it-ADI-FDTD-HD method

<table>
<thead>
<tr>
<th></th>
<th>Min. $\Delta x$, $\Delta y$</th>
<th>$\Delta T$ (s)</th>
<th>CPU-time per time-step (s.)</th>
<th>CPU-TIME total (Hrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDTD-HD</td>
<td>4 nm, 1 nm</td>
<td>$3.24 \times 10^{-18}$</td>
<td>0.35</td>
<td>90</td>
</tr>
<tr>
<td>it-ADI-FDTD-HD, $CN=100$, $it=2$</td>
<td>4 nm, 1 nm</td>
<td>$3.24 \times 10^{-16}$</td>
<td>15.17</td>
<td>39.14</td>
</tr>
<tr>
<td>it-ADI-FDTD-HD, $CN=200$, $it=4$</td>
<td>4 nm, 1 nm</td>
<td>$6.48 \times 10^{-16}$</td>
<td>30.25</td>
<td>39.16</td>
</tr>
<tr>
<td>it-ADI-FDTD-HD, $CN=300$, $it=7$</td>
<td>4 nm, 1 nm</td>
<td>$9.72 \times 10^{-16}$</td>
<td>53.11</td>
<td>43.64</td>
</tr>
</tbody>
</table>

good agreement with reference data. Chosen field-plots for specific cases are shown in Fig. 3.7. These cases are (1) Reference (2) $\langle CN$, $it \rangle = \langle 100$, 2 $\rangle$ (3) $\langle CN$, $it \rangle = \langle 200$, 4 $\rangle$ and (4) $\langle CN$, $it \rangle = \langle 300$, 7 $\rangle$. The chosen $it$ in these cases is the smallest iteration-count that provided good agreement with the reference data.

As seen, there is good agreement between the reference data and the field-profiles obtained from the proposed algorithm. When we compare the field-plots from $CN=300$ cases between Fig. 3.4 and 3.6, we find that the errors associated with the oscillations under the gate-region are corrected. As before, we further examine the current in the channel at $t = 3$ ps for several $CN$ and $it$ values in Fig. 3.7. In all the cases, the errors associated with the wavelength reduction and oscillatory behavior are removed with increasing $it$ value. As expected, the simulation-times are also increased with increasing iteration-count. Table-3.2 provides a comparison between the CPU-times for chosen three cases with best error-performance among these.

Thus, from the table we conclude that it-ADI-FDTD-HD method has allowed reduction in simulation time by a factor of 0.42 while maintaining almost same accuracy.
Figure 3.7: Channel current-profile using iterative-ADI-FDTD-HD method, compared to the reference data. Three plots refer to values of $CN=100$, 200 and 300 respectively.

Remarkably, the time performance of three cases is quite similar. To further understand the time-cost versus error compromise in these simulations, we examine these for a larger set of $it$ values. For this examination, we first define an error term as the normalized deviation of the current from reference current obtained from the simulations. We write

$$Error\% = \frac{\sum_{x=0}^{L_c} [J_{data}(x) - J_{ref}(x)]}{\sum_{x=0}^{L_c} J_{ref}(x)} \times 100$$  \hspace{1cm} (3.12)$$

where, $J_{data}(x)$ is the current calculated along the channel for the specific case under examination. Likewise, $J_{ref}(x)$ is the current for the reference case. $L_c$ is the length of channel.
The error term and time-cost for various $CN$-$it$ combinations are plotted in the Fig. 3.8. As shown, the error-term reduces with increasing $it$. However, for larger $CN$, this rate of reduction is smaller. That is, for $CN = 300$, the improvement in accuracy with increasing $it$ is not as good as for that for $CN = 100$. Previously we discussed three cases with $(CN, it) = (100, 2), (200, 4)$ and $(300, 7)$ that showed the best error-performances for corresponding $CN$ values. Using the quantitative error data, it is further clear that $(100, 2)$ case outperforms the other two cases (Refer Fig. 3.8). Therefore, using lower $CN$ and lower $it$ values outperforms using higher $CN$ and higher $it$ values, for same time-cost cases. A possible reason for this could be higher dispersion errors experienced by larger $CN$ case. Note that iterations only recover the splitting-error terms, but not the errors associated to wave dispersion due to finite cell-size and finite time-step. These errors could ultimately be the limiting factor for using larger $\Delta t$ (or larger $CN$).

Even with better error performance for low-$CN$-low-$it$ case, there are test cases when higher $CN$ values simulations will find applications. If time-cost is a priority over accuracy, especially when only a rough estimate or qualitative idea of the plasma-wave is needed, higher-$CN$-low-$it$ case can be used allowing small simulation times with some cost to accuracy of the solution. In other words, iterative ADI based method provides tunability between time-cost and error by modifying these parameters.

Finally, it will be of interest to understand the behavior of accuracy with increasing $it$. We model this using exponential-fit curves as shown in Fig. 3.9. Indeed the exponential fit shows excellent fit with the errors as a function of increasing $it$. As
Figure 3.8: Variation of error and time-cost with increasing iteration-count for CN=100, 200 and 300. The error terms are marked in red and time variation is plotted in blue.

noted before, exponent power is smaller for larger $CN$ values, signifying decreasing rate of improvement with increasing $it$.

3.4 Concluding Remarks

In this work, we have proposed new methods for modeling the plasma-wave phenomenology based on unconditional stable FDTD methods. Specifically, in this analysis, we have shown that for 5 THz excitation frequency, the total simulation time
can be reduced by a factor of 0.42 by using it-ADI-FDTD-HD method at a nominal cost of 3% error.

Due to highly refined mesh, traditional FDTD methods are quite slow for plasmonic applications. When we seek ADI-FDTD based algorithms, we encounter significant splitting-errors (or truncation errors) due to large spatial field derivative in such problems. Iterative-ADI-FDTD based methods, therefore, become effective tool to maintain the accuracy with some time-cost. Overall, we maintain the accuracy-levels with significant time-cost advantages when compared to traditional explicit-FDTD modeling.
Further, it-ADI-FDTD-HD method provides a way of tuning the time-cost and accuracy levels for different applications within plasmonic analysis. For scenarios where high accuracy is desired, a low $CN$-low-$it$ combination could be used, but for the cases when time-cost is a concern, we can use high-$CN$-low-$it$ set-up for reduced simulation times with some inaccuracy.
Chapter 4: HEMT based Terahertz Emitters Using Plasma-Waves Instability

4.1 Modeling of HEMT Based Terahertz Emitters

In this chapter, we will apply the developed numerical solvers for the prediction of performances of the terahertz emitters that use the plasma-wave instability. These are one of the first applications for such full-wave solution methods for the estimation of terahertz emissions. Specifically, we will consider two devices using our tools. We consider

1. terahertz emissions from an ungated short channel HEMT (schematic shown Fig. 4.1), and

2. a new configuration of RTD-gated HEMT device for terahertz generation (schematic shown in Fig. 4.2).

First, we will numerically examine an ungated InGaAs-HEMT with a short channel that has been known to emit terahertz radiation at cryogenic temperatures (schematic
shown in Fig. 1). Note that, the origin of these emission is known to be the plasma-wave instability in the channel, however, prior analyses only provide qualitative understanding of the phenomenon [21,34,35]. In this work, we will provide rigorous full-wave-hydrodynamic examination of the phenomenon. This examination provides us accurate emission power levels, since the developed tools solve the Maxwell’s equation and hydrodynamic equations in the entire simulation domain. Further, non-uniform HEMT surroundings around the channel, e.g. presence of source and the drain terminals, is accounted for in our models. These effects are especially significant when small channel lengths (≈100 nm) and large electron densities are used and causing smaller
Figure 4.2: Calculated emitted spectra for the GaN/AlGaN RTD-gated HEMT. Parameters used: $\epsilon_r=9.5$, $d_{bar}=30\text{nm}$, $L=100\text{nm}$, $W=70\text{nm}$, $\mu=1200\ \text{cm}^2/\text{V.s}$, $g=-6.5\times10^{12}\ \text{S/m}^2$. Details can be referred from [83].

wave-confinement and stronger interaction with non-uniform HEMT surroundings. The coupled full-wave hydrodynamic solver [48–54,58] considers these effects for accurate HEMT emission performance predictions.

Secondly, we will analyze the plasmonic-oscillations in an RTD-gated HEMT. In past, such devices have been analyzed for terahertz amplification using circuit model [65]. But such models do not take into account the multi-physical nature of the problem. By combining the electron transport and full-wave analysis, here we provide more rigorous analysis with full-wave hydrodynamic solution. In this device, a resonant-tunneling-diode (RTD) is positioned at the gate of a HEMT-device and thus interacts with the plasmonic oscillations within the channel [60]. By the virtue of negative differential resistance (NDR) of the RTD, an amplification of the plasmonic waves occurs. This amplification can be used either for generating oscillations (leading to
terahertz sources) [83] or can be used as a gain-media for terahertz amplifications [60].

An schematic for this device concept in emitter configuration is shown in Fig. 2.

In the next sections, we will conduct a numerical examination of these devices, starting with short-channel HEMT device followed by the RTD-gated HEMT device.

4.2 Full-Wave-Hydrodynamic Modeling of Terahertz Emissions from an Short Channel HEMT [38]

A mechanism for creating plasma instability, namely the Dyakonov-Shur (DS) instability, achieved by asymmetric boundary conditions at the source and drain terminals (Fig. 4.3), was proposed in [34,35]. This configuration has the potential to provide a range of terahertz devices including mixers, detectors and sources [21]. The concept itself is analyzed using linearized hydrodynamic model in [21, 34, 35]. A number of numerical studies have also been conducted to study this phenomenon, such as shown in [66,84,85], but they only consider 1D electron-transport in the channel and ignore the electromagnetic wave propagation in the media. As a result, it is not clear how much power can be derived via emission from such a configuration.

Here, we show the existence of the terahertz oscillations and power-emissions in such devices using full-wave hydrodynamic simulations. Note that the 2DEG channel in the device can be modeled using the hydrodynamic equations (as shown in previous sections). The device is operated under strong drain to source bias conditions and therefore use of hydrodynamic equations is apt for the modeling of this class of active-devices. The method models the hydrodynamic non-linearity in the channel and the field coupling to the surroundings. Therefore it accounts for mechanism of scattering
loss, reflection gain and the radiation. This enables the prediction of realistic estimates of the emitted power from the device. In the following sections, we will discuss the mechanism of wave-amplification and radiation, followed by a discussion on the results.

4.2.1 Dyakanov-Shur Instability

The device under examination is shown in Fig. 4.3. The ungated short channel HEMT consists of a 2DEG layer terminated by a perfect electrical conductor (PEC) source and drain terminal. We chose $m_e = 0.042m_o$, $\tau = 1$ ps and a uniform carrier concentration $n_{sh}$ in the channel. To understand the phenomenology, we assumed uniform velocity. This assumption is quite similar to prior analyses presented in [21, 34, 35].

The drift velocity is given by $v_o = -\tau qV_{ds}/L_c m_e$, where $L_c$ is the channel-length and $V_{ds}$ is the applied source-to-drain voltage. To implement the asymmetric boundary conditions, we fixed the AC current at $(D)$ terminal using $j_{(x=L_c)} = v_o n_{sh}$, thus emulating an AC open circuit boundary. In practice this is equivalent of having a
large inductor connected to $D$ or having a constant current source. At the source terminal $(S, x = 0)$, we used the Nuemann boundary condition $\frac{\partial V}{\partial t} = 0$ and $\frac{\partial V}{\partial x} = 0$. These enable an AC short and are equivalent to having a large capacitor or constant voltage source. To keep the terminals charge neutral, the boundary conditions $n(x=0) = n(x=L_c) = n_{sh}$ was also used. We note that for the FDTD model, we used the perfectly matched layer (PML), placed on all four sides of the device, to truncate the numerical grid. Estimation of the power radiated by the device was done by integrating the fields over the defined boundary-line between the PML and the device. Specifically, we have

$$P_{rad} = \int_C \frac{1}{2} Re(\vec{E} \times \vec{H}^*) \cdot \hat{n} dl, \quad (4.1)$$

where boundary $C$ is noted in simulation-domain shown in Fig. 4.5. Here $\hat{n}$ is unit vector normal to the boundary $C$ with $dl$ being incremental distance along $C$. As usual, $\vec{E}$ and $\vec{H}$ are the complex fields obtained from Fourier transform of recorded time domain fields.

### 4.2.2 Instability Mechanism

Before examining the instability and radiation from the device, we first consider the gain-mechanism within the channel using a 2DEG model. As shown in [34], gain is a result of the open AC boundary condition at the drain and large electron-velocities due to the applied drain to source bias. Under such conditions, when a plasmonic wave riding along the electron drift encounters an AC-open boundary condition, it experiences reflective gain. That is, the amplitude of the reflected wave becomes larger than that of the incident wave at the reflection boundary. We use numerical
model to illustrate this in Fig. 4.4, where a 4 µm-long channel is considered with open boundary conditions enforced at the end (x = 4 µm). The excitation is a time-harmonic field at x = 3.15 µm, leading to density-fluctuations to propagate in the forward (right) and the backward (left) directions (shown with black curves). The reflected wave (shown in red) is also plotted. The respective envelopes (maximum amplitude at each point) are shown with dotted curves. We observe that for non-zero drift velocity, the reflected wave at the drain shows increased amplitude as compared to the incident wave. As the electron drift velocity increases, the reflective gain continues to increase as well.

Above explains how an open-AC boundary condition can lead to wave amplification. Of course, competing loss mechanism also exists due to channel scattering. Further,
the AC-short boundary conditions (at the source) causes reflection without amplification. Thus, for a finite length 2DEG channel, a small fluctuation may travel between the source and drain and amplify in each trip. At certain frequency, such fluctuations would interfere constructively leading to sustained oscillations.

![Simulation domain used for the modeling of HEMT](image)

(a)

Figure 4.5: Simulation domain used for the modeling of HEMT (not to scale), $d_{barr} = 22 \text{ nm}$, $d_s = 1.5 \mu\text{m}$, $p = 50 \text{ nm}$, $d_{PML} = 5 \mu\text{m}$

### 4.2.3 Instability in un-gated InGaAs HEMT

Having examined the amplification-mechanism in the channel, we next model HEMT consisting of a InGaAs based heterojunction. Using our numerical model, we consider
Figure 4.6: Modeling of DS-instability in InGaAs HEMT (a) Schematic used to model the HEMT (not to scale), $d_{barr} = 22$ nm, $d_s = 1.5$ µm, $p = 50$ nm, $d_{PML} = 5$ µm (b) Current density with time as recorded at $D$ for varying $L_c$, ($v_o = 1 \times 10^5$ m/s) (c) Current density with time as recorded at $D$ for varying $L_c$ ($v_o = 2 \times 10^5$ m/s) (d) AC-Current and AC-field variation in the channel near $S$ terminal during one frequency cycle. (e) Corresponding radiation-spectra as recorded at the measurement boundary.
the case when $v_o = 10^5 \text{ m/s}$, $n_{sh} = 10^{12} \text{ cm}^{-2}$, with channel lengths of $L_c = 1 \mu\text{m}, 500 \text{ nm}$ and $100 \text{ nm}$. To start the instability, the channel was excited by introducing a square 0.1 ps long voltage pulse with amplitude of $1 \mu\text{V}$ at the drain terminal, $D$. We note that pulse magnitude itself does not effect the final oscillation amplitude. That is, the oscillation amplitude remains same no matter how the excitation is introduced. In practice, this excitation can come from the ambient noise itself (similar to any other oscillator). After excitation, we proceeded to record the current, observed at the drain terminal as a function of time, plotted in Fig. 4.6(a) and (b). We observe that due to the plasma wave propagation from $D$ to $S$ and its reflection from $S$, current oscillations build up near $D$, as shown in Fig. 4.6(a). For large $L_c (= 1 \mu\text{m}, 500 \text{ nm})$, the power loss due to the channel scattering and radiation is greater than the gain due
asymmetrical source termination, causing decaying current oscillations. In contrast, for $L_c = 100$ nm case, the scattering losses are reduced, due to shorter propagation distance. This causes a net gain in each round trip of the wave, leading to instability. Note that, in such a case, the oscillations do not amplify infinitely, rather saturate at a certain amplitude. This is due to the non-linearity of the channel-hydrodynamics, which balances the current growth with increased propagation losses as the current amplitude increases.

In other words, for $L_c = 100$ nm, the condition $2v_o/L_c > 1/\tau$ is satisfied and oscillations are supported as suggested in [21]. For $v_o = 2 \times 10^5$ m/s (Fig. 4.6(b)), we again observe oscillations. But this time, oscillations can be sustained for $L_c = 500$ nm or smaller ($2v_o/L_c > 0.8$ps). Thus, finite substrate dimensions and presence of ohmic contacts alters the condition of instability, making it more relaxed (e.g. velocity may be slightly smaller or channel length larger) as compared to that proposed in [21].

Fig. 4.6(c) shows the channel-current and channel-field variations near the terminal $S$ during one frequency-cycle (we considered $n_{sh} = 1 \times 10^{12}$ cm$^{-2}$, $L_c = 100$ nm and $v_o = 1 \times 10^5$ m/s for these plots). At steady state, 2D E-field distribution for phase points $A$, $B$, $C$ and $D$ are also shown in Fig. 4.7. The plots show that the channel radiates as a leaky-cavity. That is, EM-waves in the vicinity of the channel interact (via reflection and diffraction) with the ohmic contacts and radiates. Therefore, the structure inherently allows for the free-space coupling of plasmonic waves. Such coupling is usually achieved by using grating-gate couplers [54,86].

Thus, sustained current oscillations in the channel act as a THz emitter. This power is determined using Equation 5.1 and is shown in Fig. 4.6(d) after normalization.
Absolute power levels along with first harmonic resonance frequencies for these cases are shown as a function of varying channel length $L_c$ in Fig. 4.8. As expected, the resonance frequencies decrease with increasing $L_c$. We note that calculated resonant frequencies are smaller than that predicted from theoretical analysis in [21]. This implies that the resonance occurs at a larger 2DEG lengths than expected quarter wavelengths. Likely reason is that the fringing fields spread well outside the channel dimensions.

From Fig. 4.6(d), we observe that the most of the power is radiated at the fundamental resonance frequency, with a decrease of 20 dB or more at higher resonances.

Tens of nW of power can be expected from a mm wide device with $n_{sh} = 10^{12}$ cm$^{-2}$ and $v_o = 2 \times 10^5$ m/s (Fig. 4.8). We also observe that, although amplitude of the channel current decreases with increasing $L_c$, the total radiated power is increased.

Table 4.1: Radiated power with changing carrier velocity $v_o$, ($L_c = 100$ nm, $n_{sh} = 10^{12}$ cm$^{-2}$)

<table>
<thead>
<tr>
<th>$v_o$ (m/s)</th>
<th>$10^5$</th>
<th>$2 \times 10^5$</th>
<th>$3 \times 10^5$</th>
<th>$4 \times 10^5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{rad}$ (µW/mm)</td>
<td>0.013</td>
<td>0.027</td>
<td>0.032</td>
<td>0.033</td>
</tr>
</tbody>
</table>

Table 4.2: Radiated power with changing 2DEG confinement $t_{2DEG}$, ($L_c = 100$ nm, $v_o = 10^5$ m/s, $n_{sh} = 10^{12}$ cm$^{-2}$)

<table>
<thead>
<tr>
<th>$t_{2DEG}$ (nm)</th>
<th>5</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{rad}$ (µW/mm)</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.01236</td>
</tr>
</tbody>
</table>

78
Figure 4.8: Frequency and power dependence on channel length, \( L_c \). \((n_{sh}=1\times10^{12}\text{cm}^{-2}, v_o=2\times10^5\text{m/s})\). Theoretical calculations as per [35].

This is because longer \( L_c \) enables larger radiating area, and therefore an increase in the radiated power.

Fig. 4.9 shows the effect of changing the electron carrier concentration in the channel on the frequency and power in the radiated signal. The resonance frequency increases with increasing electron concentration and is confirmed in Fig. 4.9. Again we observe that due to the presence of ohmic contacts, finite dimensions of the channel predicted plasma resonances are significantly different from those calculated using the analytical model. Also, the power of the radiated field increases (with near quadratic dependence) with the increase in the carrier concentration. This is attributed to an
increase in the current density of the channel. Indeed, the analysis shows that, as large as 0.5 $\mu$W of power can be expected from this type device.

Our simulations also reveal that the resonant frequency is not a strong function of electron drift velocity or the thickness of the 2DEG layer. Although, the emitted power is indeed a function of these parameters. Tables 5.1 and 5.2 describe the corresponding trends. With increasing drift velocity, we also notice increase in the radiated power. This is again expected, since current amplification is larger at higher velocities [21]. Lastly, radiated power is decreasing with increasing 2DEG thickness, although this effect is rather small (Table - 5.2).
Figure 4.10: Effect of velocity saturation and decreased mobility on the emitted power from the device. The considered case is for $L_c = 100$ nm, $n_{sh} = 1 \times 10^{12}$ cm$^{-2}$.

4.2.4 Effects of velocity saturation and reduced mobility

So far, we have simplified our model by assuming a linear relationship between the channel E-field and the electron velocity. Thus, at steady state, the oscillation amplitude is limited by the non-linearity in the channel, rather than by the velocity-saturation effect. In this section, we consider this effect, by introducing a velocity saturation at high fields. We further include the effect of reduced momentum relaxation times, since $\tau = 1$ ps can only be observed at cryogenic temperatures [87, 88]. Here, we consider these effects in our simulations. In accordance with [89], we have used saturation velocity at $v_o = 1.2 \times 10^7$ cm/s.
Fig. 4.10 shows the effect of velocity saturation and decreased relaxation time on the emitted power. We consider two cases: (1) velocity saturation effect and (2) velocity saturation effect combined with decrease in the relaxation time. In both cases, the power emitted at fundamental resonance frequency is decreased. In the former case, we observe a 12 dB drop in the power due to clipping of the velocity-amplitude in the channel. In the latter case, the emitted power further decreases by 10 dB. This is due to increased scattering losses in the channel. We also note that for $\tau < 0.8$ ps, oscillations were not sustained due to excessive scattering losses. Therefore, no emitted power was observed for $\tau < 0.8$ ps. This leads to the conclusion that, this phenomenon should be expected at cryogenic temperatures only.

### 4.2.5 Summary of Findings

Using a coupled full-wave numerical model, we have shown that the asymmetrical boundary conditions in HEMTs can cause a plasma-wave instability, leading to terahertz emissions. This numerical model contained the interaction between 2DEG plasma-waves and the surrounding fields, accounting for non-uniform media and finite dimensions of the device. Our model was used to predict 1) more accurate plasma resonance threshold conditions 2) accurate resonance frequency for a given set of channel parameters and (3) maximum achievable power emissions from such devices.
Figure 4.11: Cross-sectional schematic of a HEMT with a vertical RTD between the gate and the channel. The dashed line represents the 2DEG and the S/D stand for source and drain.

Figure 4.12: Propagation constant and attenuation of the plasma waves as a function of changing differential conductivity $g$ of the RTD. For calculations, we considered $\tau=0.1366$ ps, $v_o = 10^7$ cm/s $\epsilon_r=9.5$ for GaN, $n_o = 5 \times 10^{12}$ cm$^{-2}$, and $d=25$ nm. (Negative values signify wave amplification).
4.3 RTD Assisted Amplification of Plasmons in HEMTs

As noted in previous sections, low loss electron plasma-wave propagation at room-temperature as well as self-sustained oscillations in traditional semiconductor 2DEG structures are not yet practical. This is due to the limited electron mobility and large electron plasma-wave damping at room-temperature in these materials. To develop such plasma-wave THz devices at room temperature, this low mobility limitations must be overcome by inserting additional gain mechanisms [43, 44, 90] One of such approaches can be realized with the assistance of resonant tunnel diode (RTD) introduced at the gate and biased in the negative differential conductance region of its I-V characteristics [65, 91]. RTDs operating beyond 1 THz have already been demonstrated [19, 20]. In this work, we show via analytical and numerical means that plasmons in RTD-gated HEMTs can indeed show amplification. Such a demonstration could indeed pave way to THz sources and amplifiers. Such RTD-gated devices could also serve as replacements for lossy dielectric waveguides in future terahertz integrated circuits.

In the following, we first derive the attenuation constants for propagating plasma-modes in a 2DEG channel of an RTD-gated HEMT. We find that appropriately large negative differential conductance (NDC) leads to negative attenuation constant, indicating growth of plasmonic oscillations. This analysis is next verified using full-wave simulations that model the Maxwell-Hydrodynamic equations. Finally, we discuss...
their relevance and show practical designs that can be realized for THz sources, amplifiers, and waveguide applications. This part of the dissertation is published in ref. [60, 83].

4.3.1 Plasma Propagation in RTD-gated 2DEG: Small Signal Analysis

As discussed in Chapter-1, the dispersion relations for propagating plasma waves in 2DEG channels have been derived in the literature [39, 45–47, 73]. Specifically, [47] gives the attenuation due to finite mobility for a biased gated 2DEG layer. We will expand on this analysis to include the effect of RTD placed at the gate, adjacent to the 2DEG layer.

A schematic of the device under consideration is shown in Fig. 4.11. We consider a long, thin 2DEG channel under a metal gate separated by a vertical RTD at the barrier layer. The 2DEG channel is biased by applying source to drain voltage, while the RTD is biased using the gate’s voltage with respect to the channel’s local potential. With this set-up, the current through the RTD is added to the 2DEG current. To simplify the analysis, and to be able to obtain closed form expression, we make following assumptions:

1. RTD DC current is assumed to be small and does not cause significant variations in sheet-electron density or velocity along the channel. In other words, the DC conditions in the channel are dominated by drain to source bias (and not by the gate to channel bias).
2. Variations in the RTD bias are small, allowing almost uniform NDC as we move along the channel.

3. Plasma wavelength is much larger than barrier thickness $d$, but is much smaller than the free-space wavelength.

With the above assumptions, we proceed to employ the electron transport equations in the channel. These are the first two moments of Boltzmann Transport Equation (BTE) [55], viz.

$$\frac{\partial n_{sh}}{\partial t} + \frac{\partial j}{\partial x} = \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}, \quad \text{and}$$

$$\frac{\partial j}{\partial t} + v \frac{\partial j}{\partial x} + j \frac{\partial v}{\partial x} = - \frac{q n_{sh} E_{ch}}{m_e} - \frac{j}{\tau} + v \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}. \quad (4.3)$$

Here, $n_{sh}$ is the sheet carrier density, $j = n_{sh} v$ is the sheet electron flux and $v$ is the electron velocity within the 2DEG channel. Also, $E_{ch}$ is the $x$-directed electric field along the channel (see Fig. 4.10), $\tau$ refers to momentum relaxation time, $m_e$ is the effective electron mass and $q = 1.6 \times 10^{-19}$C is the charge of a single electron. To account for the small-signal current due to RTD, the term $\left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}$ is added on the right hand side of (1). This term is much similar to generation recombination terms used in [92]. Similarly, a change in the local momentum due to the RTD can be accounted for by introducing the term $v \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}$ in (2). This addition assumes that the electrons added due to the RTD quickly attain the local-bulk velocity in the channel, altering the momentum by $v \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}$. This is a reasonable assumption since electron tunneling is vertical to the 2DEG and electrons conserve momentum in the $x$-$z$ plane. We note that Equations 4.2 and 4.3 can be simplified by replacing...
\( j = n_sh v \) in Equation 4.3, giving,

\[
\frac{\partial n_{sh}}{\partial t} + \frac{\partial j}{\partial x} = \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD}, \quad \text{and} \quad (4.4)
\]

\[
\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = -\frac{q E_{ch}}{m_e} - \frac{v}{\tau}, \quad (4.5)
\]

To derive the dispersion formulae from the above, our strategy is to first establish a relation between the channel AC current and the electric-field. This will provide an expression for the channel conductivity \( \sigma_{ac} \) in terms of \( \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD} \), which is related to the E-field profile in the barrier layer. Using this \( \sigma_{ac} \), we can then employ the E and H-field boundary conditions to obtain the desired dispersion relations.

We proceed by introducing a small signal, space-time field perturbation, given by \( E_{ch} = E_o + E_{ac} \exp(j\omega t - j \beta_x x) \). This leads to represent the electron density as \( n_{sh} = n_o + n_{ac} \exp(j\omega t - j \beta_x x) \). Further, the associated velocity is given by \( v = v_o + v_{ac} \exp(j\omega t - j \beta_x x) \). As can be seen, \( n_{ac} \) and \( v_{ac} \) are the amplitudes of the AC variations in the electron density and velocity, respectively. As expected, \( \omega (= 2\pi f) \) is the angular frequency and \( \beta_x \) is the propagation constant along the channel (i.e., in \( x \)-direction). We next introduce the known AC current representation

\[
J_{ac} = -qv_o n_{ac} - qn_o v_{ac} - qn_{ac} v_{ac} \quad (4.6)
\]

and by invoking Equations 4.4 and 4.5, we get

\[
J_{ac} = \frac{q^2 E_{ac}}{m} \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD} + j \omega n_o \frac{(\partial n_o/\partial t)_{RTD}}{j \omega - j \beta_x v_o} - \frac{q v_o (\partial n_{sh}/\partial t)_{RTD}}{j \omega - j \beta_x v_o} \quad (4.7)
\]

In the above, \( J_{ac} \) can be further simplified by introducing \( \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD} = -\frac{g V_{RTD}}{q} \), where \( V_{RTD} \) is the AC voltage across the RTD barrier and \( g \) is the RTD differential conductance (in S/m^2). As expected, \( V_{RTD} = -\int_{y=-d}^{y=0} E_y dy \) is related to the 2DEG
field, $E_{ac}$. To derive this relation, we further consider the field profiles in the barrier layer $-d < y \leq 0$ and below the 2DEG layer $y \leq -d$. Since we do not expect any $E_z$ fields, we begin by introducing the vector potentials $\vec{F} = \hat{z}F_z$, where $F_z$ is

$$F_z = \begin{cases} 
A \cos(\beta_y y) \exp(j \omega t - j \beta_x x); & -d \leq y \leq 0 \\
B \exp(\alpha_y y) \exp(j \omega t - j \beta_x x); & y < -d 
\end{cases} \quad (4.8)$$

Here, $\beta_y$ is the propagation constant in the $y$-direction for the region $-d < y \leq 0$ and $\alpha_y$ is the attenuation constant for when $y \leq -d$. Further, by invoking the wave equation, we obtain the characteristic equations, $\beta_y^2 + \beta_x^2 = \beta^2$ and $-\alpha_y^2 + \beta_x^2 = \beta^2$ where $\beta = \omega \sqrt{\epsilon \mu_o}$ is the propagation constant of the wave in the dielectric. Here, we have $\epsilon = \epsilon_r \epsilon_o$, with $\epsilon_r$ being the dielectric constant. $\epsilon_o$ and $\mu_o$ are permittivity and permeability in the vacuum, respectively. In the above, $A$ and $B$ are to be eliminated via enforcement of the boundary conditions.

The field components corresponding to $\vec{F}$ are found from $\vec{E} = \frac{-1}{\epsilon} \nabla \times \vec{F}$, giving

$$E_x = \frac{-1}{\epsilon} \frac{\partial F_z}{\partial y} = \begin{cases} 
\frac{A \beta_x}{\epsilon} \sin(\beta_y y) \exp(j \omega t - j \beta_x x); & -d \leq y \leq 0 \\
-\frac{B \alpha_y}{\epsilon} \exp(\alpha_y y) \exp(j \omega t - j \beta_x x); & y < -d 
\end{cases} \quad (4.9)$$

$$E_y = \frac{1}{\epsilon} \frac{\partial F_z}{\partial x} = \begin{cases} 
\frac{-A \beta_x}{\epsilon} \cos(\beta_y y) \exp(j \omega t - j \beta_x x); & -d \leq y \leq 0 \\
-\frac{B \beta_x}{\epsilon} \exp(\alpha_y y) \exp(j \omega t - j \beta_x x); & y < -d 
\end{cases} \quad (4.10)$$

also, $H_z = -j \frac{1}{\omega \mu_e} \left( \frac{\partial^2}{\partial z^2} + \beta^2 \right) F_z$, viz.,

$$H_z = -j \frac{\beta^2 F_z}{\omega \mu_e} = \begin{cases} 
\frac{-j \omega \mu_e \beta^2 A \cos(\beta_y y) \exp(j \omega t - j \beta_x x); & -d \leq y \leq 0 \\
\frac{-j \omega \mu_e \beta^2 B \exp(\alpha_y y) \exp(j \omega t - j \beta_x x); & y < -d 
\end{cases} \quad (4.11)$$

Recognizing that the channel-field $E_{ac}$ can be obtained by $E_{ac} = E_y^{y=-d}$, and then dividing $E_{ac}$ by 4.10 gives $E_y$ in terms of $E_{ac}$,

$$E_y = \frac{j \beta_x E_{ac}}{\beta_y \sin(\beta_y d)} \cos(\beta_y y) \quad (4.12)$$
Also, from

\[ V_{RTD} = - \int_{y=-d}^{y=0} E_y dy = -\frac{j \beta_x E_{ac}}{\beta_y^2} \]  

(4.13)

we get,

\[ \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD} = -\frac{gV_{RTD}}{q} = g \frac{j \beta_x E_{ac}}{q \beta_y^2} \]  

(4.14)

Above can is used to replace the term \( \left( \frac{\partial n_{sh}}{\partial t} \right)_{RTD} \) in (6). After further algebraic manipulations, 4.7 becomes

\[ J_{ac} = \sigma_{ac} E_{ac} + \gamma E_{ac}^2, \]  

(4.15)

where,

\[ \sigma_{ac} = \frac{q^2}{m_e (j\omega - j \beta_x v_o)(j\omega - j \beta_x v_o + 1/\tau)} - \frac{v_o \beta_x g}{\beta_y^2 (\omega - v_o \beta_x)}, \]

and

\[ \gamma = \frac{q^2}{m_e (j\omega - j \beta_x v_o)(j\omega - j \beta_x v_o + 1/\tau)} \times \left[ \frac{-jq \beta_x n_o}{m(j\omega - \beta_x v_o + 1/\tau)} - \frac{j \beta_x g}{q \beta_y^2} \right]. \]

Since \( E_{ac}^2 \) is small, an additional simplification is to ignore the \( E_{ac}^2 \) term. Doing so, we obtain a linear current-field relation for determining the channel AC conductivity \( \sigma_{ac} \). An expression for channel-conductivity was derived in [47] as well, but 4.15 above also accounts for the effect of NDC \( g \), as controlled by the second term of \( \sigma_{ac} \).

We next proceed to enforce the other boundary conditions and therefore eliminate the \( A \) and \( B \) constants. Specifically, we have \( H_z|_{y=-d^+} - H_z|_{y=-d^-} = \sigma_{ac} E_x|_{y=-d} \) long with \( E_x|_{y=-d^+} = E_x|_{y=-d^-} \). When these are substituted in (8)-(10), we get

\[ \cot(\beta_y d) \pm j = \frac{\sigma_{ac} \beta_y}{j\omega \epsilon} \]  

(4.16)
Figure 4.13: Attenuation constant (in Np/µm) as a function of relaxation time ($\tau$) and RTD-NDR ($g$). The marked line refers to the zero attenuation contour. For the calculations, we considered $v_o = 10^7$ cm/s, $\epsilon_r = 9.5$ for GaN, $n_o = 5 \times 10^{12}$ cm$^{-2}$, and $d = 25$ nm. (a) Plots at frequency $f = 1$THz (b) Plots at frequency $f = 5$THz (Negative values signify wave amplification)

with $\sigma_{ac}$ as given in 4.15. This can be simplified to a quadratic equation in $\beta_y$, using $\cot \beta_y d \approx \frac{1}{\beta_y d}$ since the barrier thickness $d$ is much smaller than the plasma-wavelength and $\beta_x \gg \beta$. Thus we have

$$
\beta_x^2 \left(-v_o^2 + \frac{q^2n_o d}{m_e \epsilon} - \frac{jgdv_o^2}{\omega \epsilon}\right) + \beta_x \left(2v_o \omega - \frac{jv_o}{\tau} + \frac{jv_o gd}{\epsilon \omega \tau} + \frac{v_o gd}{\omega \tau} \right) + \frac{j\omega}{\tau} - \omega^2 = 0
$$

(4.17)

This new relation allows us to find the plasma-wave propagation constant as a function of $\omega$, while considering the effect of NDC $g$. We note that for $g = 0$, 4.17 converges to the dispersion relation for the plasma-wave propagation in the gated 2-DEG as derived in [47]. This serves as verification for above calculations.

Dispersion curves can be obtained by solving the quadratic Equation 4.17. The two solutions to Equation 4.17 correspond to wave moving in forward and backward directions. Here, we define forward wave as that traveling in the direction of the
electron drift velocity due to drain to source bias, i.e. \(\Re(\beta_x) > 0\) for \(v_o > 0\). Hereon, we only consider the forward wave, since the backward wave is associated with much higher attenuation and is not of interest (as also reported in \([47, 58]\) for non RTD gated cases).

**Solution for RTD-gated GaN/AlGaN Heterojunction**

As a next step, we evaluated the developed solution for a GaN/AlGaN heterojunction with an RTD placed at the gate. For the given material system, parameters chosen are \(\varepsilon_r=9.5\), \(n_o=5\times10^{12}\ \text{cm}^{-2}\), \(d=25\ \text{nm}\), \(\tau = 0.1366\ \text{ps}\) and \(v_o = 10^7\ \text{cm/s}\). Fig. 4.12 shows the effect of RTD differential conductance \(g\) on the wave propagation. Specifically, the phase constant (\(\Re(\beta_x)\)) and attenuation constant \(\alpha (=-\Im(\beta_x))\) are plotted as a function of frequency. We observe that increasing the value \(-g\) gradually decreases the attenuation constant, eventually causing amplification. From these plots, we can conclude a \(g_{RTD} < -0.5 \text{S/\mu m}^2\) is needed for plasma-wave amplification. However, since plasma-wave losses are higher for smaller relaxation time, \(\tau\), the attenuation constant is expected to increase accordingly.

Fig. 4.13 shows the attenuation constants at frequencies 1 THz and 5 THz as a function of \(\tau\) and \(g\). As expected, the wave amplification is highest for large \(\tau\) and large \(-g\). We also note that NDC needed for amplification becomes smaller, when \(\tau\) is large (refer to the \(\alpha = 0\) contour).

The above analysis suggests that the RTD-gate can reinforce plasma-wave propagation, reducing losses and even causing amplification in accordance with the NDC.
value provided by the RTD. To further validate the analysis, we consider comparisons
with the full-wave hydrodynamic simulations. These comparisons are presented next.

4.4 Full-wave Modeling of RTD-Gated HEMT: Validation and
Results

In this section, we validate the aforementioned analysis using numerical finite dif-
ference time domain (FDTD) simulations. In this work, we expand on this model
to consider the presence of an RTD-NDC between the 2DEG channel and the gate.
Below, we briefly discuss the developed numerical model and then use it to calculate
dispersion relations. These will be compared with the analytical model given in the
previous sections.

4.4.1 Full-Wave-Hydrodynamic FDTD analysis for RTD-gated
HEMT

The approach adopted here is same as discussed in Chapter 2. We use the FDTD-HD
solver, with hydrodynamic equations modified to account for the RTD AC current
(Equations 4.2 and 4.3). Likewise the EM fields were accounted for by the use of
Maxwell’s electrodynamic equations. For extraction of propagation constant via nu-
merical solution, the model shown in Fig. 4.14 was used.
Figure 4.14: Model used for the numerical calculations. The size of the discontinuity is chosen to be $p=40\text{nm}$.

### 4.4.2 Comparison of Numerical and Analytical Solutions

As already noted, the derivation of the dispersion relation (16), assumes linearization of the hydrodynamic equations. Specifically, second order variations ($v_{ac}^2, n_{ac}^2$ etc.) are ignored, as noted in 4.15. Additionally, we also assume that $\beta_y d \ll 1$ and $\beta_x^2 \approx -\beta_y^2 \gg \beta^2$. These assumptions were needed to obtain a closed form solution. Therefore, to validate the analytical results using full-wave simulations, we must carefully choose the frequency band and the parameters where these approximations hold true. For the comparison, we chose parameters typical to a 2DEG channel in GaN/AlGaN heterojunctions, except for its dielectric constant. We chose $\epsilon_r = 1$, so that approximation $\beta_y d \ll 1$ is valid due to larger plasma-wavelength for smaller $\epsilon_r$. Other parameters used were $n_0 = 5 \times 10^{12} \text{ cm}^{-2}$, $\tau = 0.137 \text{ ps}$, $v_o = 10^7 \text{ cm/s}$ and $m_e = 0.2m_o$. We chose the thickness of the 2DEG layer to be $t_{2DEG} = 2.5 \text{ nm}$.
Figure 4.15: Wavenumbers and the attenuation constants obtained using the full-wave method and the analytical method. The "symbol" lines refer to numerical simulations and the solid lines refer to the analytical model. Choice of parameters: $\epsilon_r=1$, $d=25$ nm, $n_o = 5 \times 10^{12}$ cm$^{-2}$, $\tau=0.137$ ps, $v_o=10^7$ cm/s and $m_e=0.2m_o$.

To obtain the dispersion curve, we need to excite a broadband THz plasma-wave in 2DEG and then record its propagation and attenuation constant as it travels along the channel. To do so, we model a THz pulse incident on the HEMT which contains a long and RTD-gated channel ($L_c=10$ µm), having a small gap-discontinuity at the gate (Fig. 4.14). The incident wave diffracts through the gap and couples with the channel to excite the propagating plasmonic modes. We record the time domain signal at different points in the channel and extract the attenuation and phase constants by Fourier transformation across the band of interest. The results for the $\epsilon_r = 1$ case are plotted in Fig. 4.15.
Figure 4.16: Wavenumbers and the attenuation constants for GaN based RTD-gated HEMT obtained using the full-wave method and the analytical method. The "symbol" lines refer to numerical simulations and the solid lines refer to the analytical model. Choice of parameters: $\epsilon_r=9.5$, $d=25$ nm, $n_o = 5 \times 10^{12}$ cm$^{-2}$, $\tau=0.137$ ps, $v_o=10^7$ cm/s and $m_e=0.2m_o$.

Notably, both models predict only a negligible change in the wavenumber with varying $g$. These plots also demonstrate an agreement in the attenuation constants calculated via the two models. Here, we also confirm that the attenuation constant decreases with the increasing $-g$. Small variations in the simulated results were observed due to the inherent non-linearity of the numerical solver and reflection from source and drain terminals - both effects not accounted for by the analytical solution.

Next, a similar comparison is made for a practical case of GaN/AlGaN device, where dielectric constant is set $\epsilon_r = 9.5$. The comparisons are shown in Fig. 4.16. We
Figure 4.17: Left: Schematic of the RTD-gated HEMT with RTD typical characteristics. Right: $E_x$ field snapshots taken at $t=2$ps. Detail along the partial section of the channel. Plasma-wave propagation for varying values of the RTD-NDC $g$ is shown. Reduced plasma-wave losses and amplification can be obtained with increasingly $-g$.

(Note: $x$ and $y$ axis are drawn with respect to the corner of the 2D-simulation domain and figure shows the zoomed in view.)

again confirm the reduction of the attenuation constant as predicted by the analytical model. Notably, more deviations are observed between the two models specially at larger frequencies. This is due to the higher dielectric constant causing $\beta_y d$ to become large, violating the assumptions made in the analysis. E.g. at 3 THz $\beta_y d$ is close to 0.3.

From above, we infer that the analytical model correctly captures the phenomenon in the regime of long plasma-wavelengths. On the other hand, the full-wave-hydrodynamic
model provides rigorous and accurate solution for all frequencies. It is also noted that for $\epsilon_r > 1$, analytical solution deviates from full-wave simulations. In such cases, use of full-wave-hydrodynamic simulations is recommended for design while the analytical model can be used for initial estimates.

4.5 Plasmon Propagation in RTD-gated GaN/AlGaN Heterojunctions

In this section, we employ numerical solution to show the field-visualization for intuitive understanding of the proposed concept. The model and the excitation method are the same as in the previous sub-section, but now we chose single excitation frequency of 5 THz. Other parameters are set as $n_o = 5 \times 10^{12} \text{cm}^{-2}$, $\tau = 0.137 \text{ps}$, $v_o = 10^7 \text{cm/s}$, $\epsilon_r = 9.5$ and $m_e = 0.2m_o$.

Fig. 4.17 shows the coupling and propagation of plasma-waves for varying values of RTD-NDC. For $g=-0.5 \text{S/}\mu\text{m}^2$, plasmon attenuation is reduced allowing propagation to longer distances. This observation is very useful for active terahertz waveguides. For the case of large NDC ($g=-0.8 \text{S/}\mu\text{m}^2$ and $-1 \text{S/}\mu\text{m}^2$), a growing propagating plasma-wave was also observed. With appropriate antennas at the input (source-gate) and output (drain-gate) terminals, an antenna-coupled amplifier configuration can be achieved [58]. Clearly, this would require impedance-matching of the input and output terminals with the corresponding radiating structures. If left un-matched, the amplified plasma-wave would simply be reflected back from the drain terminal. This backward traveling wave would quickly attenuate due to opposite flow of electrons.
and large \( -g \). Another way of achieving input-output coupling in the amplifier mode could be by using grating-gates, similar to in [54].

The shown concept can also be utilized to obtain a terahertz source as well, as shown in Fig. 4.2. The RTD-gate provides the much needed gain media for the propagating plasmons. Alongside this gain-media, a resonance mechanism is also needed to complete the oscillator action. This resonance can be obtained by having a periodic grating-gates instead of a single continuous gate. Further, these grating-gates also provide the needed method for coupling of the plasmons to the free-space. This mechanism is further described in [83]. Thus obtained RTD-gated, grating-gate HEMT is expected to emit terahertz radiations. For such devices, the frequency of resonance would be decided by the grating periodicity, 2DEG electron density and barrier thickness.

### 4.6 Conclusion

We demonstrated that RTD-gated HEMTs support low-attenuation and growing plasma modes within the 2DEG channel. Depending on the value of the RTD-NDC, either a long distance plasma-wave propagation or amplification action can be sustained within the gated channel. This mode enhancement is supported by the RTD-gain mechanism, which counters the scattering losses in the channel. Our full-wave-hydrodynamic model and analytical expressions are validated for this concept.

We also concluded that, in practice such devices would depend on fabrication of the state-of-the-art RTDs, operating at THz frequencies. Such RTDs are already
being reported [19, 20]. The NDC values available in these RTDs are of the order of \(10 \, mS/\mu m^2\). Although these values are smaller than the required, development of low-loss active THz-waveguides is possible using these RTDs. For THz-sources and amplifiers, higher NDC values would be required, which could allow these applications in future.
Chapter 5: Plasma-modes in a Double 2DEG Channel Systems

5.1 Introduction

In this chapter, we present a rigorous study of the plasma-wave modes in a bilayer (or double 2DEG channel) system using analytical and numerical tools that were developed in the previous chapters. Such bilayer structures are developed by using a lattice of high bandgap and low bandgap semiconductor materials placed periodically [93]. Indeed, this process has even provided us multiple 2DEG-channel systems in past [94]. Here, we pursue study for the plasmonic-behavior of double channel systems using the developed numerical and analytical models.

Theoretical analysis of the plasma-waves in such two channel systems (also known as bilayer) has been considered previously [95,96]. These showed that small inter-channel distances ($\Delta d \ll \lambda_p$, $\lambda_p$ being plasmon wavelength) imply a strong electromagnetic coupling that leads to collective plasmonic modes in the channels. Two such modes have been identified: 1) optical mode; caused due to in-phase oscillations of the electron densities in the two channels, and 2) acoustic mode; arising due to out-of-phase oscillations. These modes have also been confirmed experimentally as
demonstrated in ref [97], among others. Inter-channel tunneling is also possible in a bilayer configurations leading to interesting physics in such systems as theoretically and experimentally demonstrated in [98,99].

In the theoretical work so far, bilayer structure was placed in an ungated environment, i.e. the bilayer is assumed to be embedded in homogeneous or heterogeneous dielectric media. Here, we instead consider a practically relevant case. Specifically, the bilayer or the dual channel is placed in the vicinity of metal gate. Such metal gates are frequently used to provide free-space coupling of plasmons and for modulating the electron density [69,100,101]. Recently, a gate was shown to modulate electron densities in multiple overlaying channels using a castellated configuration [102]. Another practical scenario of gated bilayer is tunneling coupled bilayer where gate voltage is used to tune subband energy levels [99]. We note that prior to this work, gate considerations were also made for double gated graphene bilayers in [103], but dispersion relations have yet to be derived.

In this chapter (also published in [61]), we will derive dispersion relations for gated double channels or bilayers using electromagnetic field profiles between the channel (similar to [73]). We find that the presence of the metal-gate can selectively affect the optical plasmonic modes in the bilayer. This implies a decrease in their wavelength and thus an increase in the attenuation constant \(Np/\mu m\). However, the acoustic plasmons remain unchanged in presence of the metal gate. These observations are then verified using our full-wave-hydrodynamic FDTD solver. The solver is subsequently used to examine these modes in the gated/ungated bilayers. Specifically, we study the mechanism of coupling of these modes, excited by an incident plane wave.
5.2 Dispersion of Plasma-waves in a Gated Double Channel 2DEG: Small Signal Analysis

We consider a long, thin 2DEG channel placed under a metallic gate at a vertical distance $d_1$ (see Fig. 5.1). A second 2DEG channel is positioned at a distance $d_2$ from the metallic gate. We also assume a drain-to-source bias to enable drift velocities $v_{o1}$ and $v_{o2}$ in the upper and lower channels, respectively. The corresponding initial electron density in the two channels are assumed to be $n_{o1}$ and $n_{o2}$, respectively. To calculate the dispersion relation, we proceed to first calculate the AC conductivity of
the channel using hydrodynamic equations. Next, we solve for the EM-fields within
the channel and in its vicinity, subject to the boundary conditions in the 2DEG and
gate. The derived EM-fields are then used to produce the dispersion relation.

Below, we consider the currents and fields in the upper channel to calculate its AC
conductivity. To begin with, we invoke the hydrodynamic approximations based
on Boltzmann’s Transport Equations (BTE), [43]. That is, we consider continuity
equation coupled with the momentum conservation equations, viz,

\[
\frac{\partial n_1}{\partial t} + \frac{\partial j_1}{\partial x} = 0 \tag{5.1}
\]

\[
\frac{\partial j_1}{\partial t} + v_1 \frac{\partial j_1}{\partial x} + j_1 \frac{\partial v_1}{\partial x} = -\frac{q n_1 E_{x1}}{m_e} - j_1 \tau \tag{5.2}
\]

where \( j_1(= n_1 v_1) \) is the sheet electron flux with \( n_1 \) and \( v_1 \) being the total electron
density (i.e AC and DC) and velocity in the 2DEG channel. \( E_{x1} \) is the \( x \)-directed
electric field along the channel. In addition, \( \tau \) refers to the momentum relaxation
time and \( m_e \) is the effective electron mass and \( q=1.6\times10^{-19} \) C is the charge on a
single electron. Suffix 1 in above equations refers to the first channel in the HEMT.
Similar analysis is done for second channel.

For small signal analysis, we assume that an electric field of the form \( E_{x1} = E_{o1} +
E_{ac} \exp(j\omega t - j\beta x) \) exists within the top 2DEG channel in Fig. 5.1. We also define
the corresponding electron density and velocity as \( n_1 = n_{o1} + n_{ac} \exp(j\omega t - j\beta x) \)
and \( v_1 = v_{o1} + v_{ac} \exp(j\omega t - j\beta x) \), respectively. \( E_{ac}, n_{ac} \) and \( v_{ac} \) are the amplitudes
of the AC electric-field, electron density and associated velocity, respectively. Here,
\( \omega(= 2\pi f) \) is the angular frequency and \( \beta_x \) is the propagation constant along the
channel (\( x \)-direction).
Using (5.1) and (5.2), and noting that $J_{ac} = -qv_1n_{ac} - qn_{o1}v_{ac}$, we have

$$J_{ac} = \frac{q^2E_{ac}}{m_e} \frac{j\omega n_{o1}}{(j\omega - j\beta_x v_{o1} + \frac{1}{r})(j\omega - j\beta_x v_{o1})}$$

(5.3)

$J_{ac}$ is associated with the channel conductivity $\sigma_{ac1}$ via $J_{ac} = \sigma_{ac1}E_{ac}$. Thus, we have

$$\sigma_{ac1} = \frac{q^2}{m_e} \frac{j\omega n_{o1}}{(j\omega - j\beta_x v_{o1} + \frac{1}{r})(j\omega - j\beta_x v_{o1})}$$

(5.4)

This AC conductivity is the same as that derived previously [47, 69]. Using similar steps as above, we can also write the AC conductivity of the lower 2DEG layer in Fig. 5.1(b), as

$$\sigma_{ac2} = \frac{q^2}{m_e} \frac{j\omega n_{o2}}{(j\omega - j\beta_x v_{o2} + \frac{1}{r})(j\omega - j\beta_x v_{o2})}$$

(5.5)

Having obtained the conductivity of the two channels, we proceed to find the coupled field profiles in presence of the two channels. Considering the need to enforce field continuity between the two 2DEG channels, we introduce the field components [104, 105]

$$E_x = \frac{-1}{\epsilon} \frac{\partial F_z}{\partial y} \quad \text{and} \quad H_z = \frac{-j}{\omega \mu \epsilon} \beta^2 F_z$$

(5.6)

as shown in Fig. 5.1(b). Here the potential $F_z$ is given by

$$F_z = \begin{cases} 
A \cos(\beta_y y) \exp(j\omega t - j\beta_x x) & -d_1 \leq y \leq 0 \\
B \exp(-j\beta_y y) + C \exp(j\beta_y y)] \exp(j\omega t - j\beta_x x) & -d_2 < y < -d_1 \\
D \exp(\alpha_y y) \exp(j\omega t - j\beta_x x) & y < -d_2 
\end{cases}$$

(5.7)

As usual, $\beta_y$ is the propagation constant in the $y$-direction for the region $-d_2 < y \leq 0$ and $\alpha_y$ is the attenuation constant for the region $y \leq -d_2$.

Next, we use (5.7) in the wave equation $\nabla^2 F_z + \beta^2 F_z = 0$ to giving the dispersion relation $\beta_y^2 + \beta_x^2 = \beta^2$ and $-\alpha_y^2 + \beta_x^2 = \beta^2$.

In (5.7), $A$, $B$, $C$ and $D$ are constants to be eliminated by enforcing the boundary conditions at $y = 0$, $y = -d_1$ and $y = -d_2$. Prior to applying the boundary conditions,
Figure 5.2: Dispersion relations for the gated bilayer obtained from (5.14). The curves correspond to varying inter-channel distance with a fixed $d_1=40$ nm. The curves show that the acoustic plasmons are similar to those in the ungated bilayer case. However, the optical plasmons have smaller propagation and higher attenuation constants.

we expand (5.6) to read

$$E_x = \begin{cases} \frac{A \beta_y}{c} \sin(\beta_y y) \exp(j \omega t - j \beta_x x) & -d_1 \leq y \leq 0 \\ -j \beta_y \left[-B \exp(-j \beta_y y) + C \exp(j \beta_y y)\right] \exp(j \omega t - j \beta_x x) & -d_2 < y < -d_1 \\ -D \alpha_x \exp(\alpha_y y) \exp(j \omega t - j \beta_x x) & y < -d_2 \end{cases}$$

and

$$H_z = \begin{cases} \frac{-1}{\omega \mu \epsilon} \beta^2 A \cos(\beta_y y) \exp(j \omega t - j \beta_x x) & -d_1 \leq y \leq 0 \\ \frac{-1}{\omega \mu \epsilon} \beta^2 [B \cos(-j \beta_y y) + C \cos(j \beta_y y)] \exp(j \omega t - j \beta_x x) & -d_2 \leq y \leq d_1 \\ \frac{-1}{\omega \mu \epsilon} \beta^2 D \exp(\alpha_y y) \exp(j \omega t - j \beta_x x) & y < -d_2 \end{cases}$$

To find $A$, $B$, $C$, and $D$, we enforce the boundary conditions on the tangential $E$ and $H$-fields across the two 2DEGs. From $E$-field continuity conditions: $E_x|_{y=-d_1^+} =$
\( E_x|_{y=-d_1^+} \) and \( E_x|_{y=-d_2^+} = E_x|_{y=-d_2^-} \), we get
\[
A \sin(\beta_y d_1) = C \exp(-j \beta_y d_1) - B \exp(j \beta_y d_1) \tag{5.10}
\]

and
\[
D \alpha_y \exp(-\alpha_y d_2) = j \beta_y \left[ C \exp(-j \beta_y d_2) - B \exp(j \beta_y d_2) \right] \tag{5.11}
\]

Further, from \( H_z|_{y=-d_1^+} - H_z|_{y=-d_1^-} = \sigma_{ac1} E_x|_{y=-d_1} \), we have
\[
A \left[ \frac{j \beta^2}{\omega \mu} \cos(\beta_y d_1) - \sigma_{ac1} \beta_y \sin(\beta_y d_1) \right] = \frac{j \beta^2}{\omega \mu} B \exp(j \beta_y d_1) + \frac{j \beta^2}{\omega \mu} C \exp(-j \beta_y d_1) \tag{5.12}
\]

Likewise, at \( y = -d_2 \) we have \( H_z|_{y=-d_2^+} - H_z|_{y=-d_2^-} = \sigma_{ac2} E_x|_{y=-d_2} \), giving
\[
D \left[ \frac{j \beta^2}{\omega \mu} \exp(-\alpha_y d_2) - \sigma_{ac2} \alpha_y \exp(-\alpha_y d_2) \right] = \frac{j \beta^2}{\omega \mu} \left[ B \exp(j \beta_y d_2) + C \exp(-j \beta_y d_2) \right] \tag{5.13}
\]

We can further manipulate (5.10)-(5.13), to eliminate all constants \( A, B, C \) and \( D \).

This process gives us,
\[
\frac{\sigma_{ac1} \beta_y - j \cot(\beta_y d_1) - 1}{\sigma_{ac1} \beta_y - j \cot(\beta_y d_1) + 1} = \exp(2j \beta_y \Delta d) \frac{1 + \frac{j \beta_y}{\alpha_y} + \frac{\sigma_{ac2} \beta_y}{\omega \mu}}{-1 + \frac{j \beta_y}{\alpha_y} + \frac{\sigma_{ac2} \beta_y}{\omega \mu}} \tag{5.14}
\]

where \( \Delta d = d_2 - d_1 \). In the above, \( \sigma_{ac1} \) and \( \sigma_{ac2} \) are functions of \( \beta_x \) as described by (5.4) and (5.5). Further, \( \beta_y \) and \( \alpha_y \) can be written in terms of \( \beta_x \) using the characteristic equation \( \beta_x^2 + \beta_y^2 = \beta^2 \) with \( \alpha_y = \pm j \beta_y \). That is, for a specified frequency, (5.14) defines the variable \( \beta_x \). In effect, it represents the dispersion relation of the propagating plasma-wave modes along the \( x \) direction in a double channel system. This dispersion relation can be solved numerically using functions like \textit{fzero} or \textit{newtzero} in MATLAB.

We remark that (5.14) reduces to the known dispersion relation for a single channel system, when \( n_{02} = 0 \) (leading to \( \sigma_{ac2} = 0 \)). Indeed, setting \( \sigma_{ac2} = 0 \) in (5.14), we
obtain
\[
\frac{\sigma_{ac1} \beta_y}{\omega \epsilon} - j \cot(\beta_y d_1) \pm 1 = 0. \tag{5.15}
\]
This result is in agreement with the previously derived dispersion relation [47] and serves to verify our calculations.

We note that (5.14) has four roots signifying two propagating modes in the forward and backward directions. But the two directions are equivalent since we have considered zero drift velocity in the channel. Thus, we only need to plot the solutions for the forward direction. For these plots, we have used GaN/AlGaN material system with parameters chosen as \( n_{o1} = n_{o2} = 7.5 \times 10^{12} \text{ cm}^{-2}, \epsilon_r = 9.5, m_e = 0.2 m_0 \) and the channel mobility \( \mu = 1800 \text{ cm}^2/\text{V.s} \ (\tau = 0.2 \text{ ps}) \) for both the channels. The obtained modes, referred to as optical and acoustic modes, are shown in Fig. 5.2. 2D field-profiles for these modes is further shown in Fig. 5.6.

As shown in the figure, analysis is done for varying \( \Delta d \) with fixed \( d_1 (= 40 \text{ nm}) \). We observe that the acoustic modes become increasingly lossy as \( \Delta d \) decreases, whereas, the optical modes experience little change. This observation is similar to those for ungated bilayers. This suggests that inter-channel distances play same role for gated bilayer as in ungated one.

A comparison with ungated bilayers \((d_1 \rightarrow \infty \text{ in (5.14)})\) reveals that the phase and attenuation constants of the acoustic branch remains unchanged, even for small \( \Delta d \) values. This is because the electric field is confined within the bilayer. Therefore, it does not interact with the gate. However, the optical plasmons interact strongly with
the added gate metal. Thus, they have decreased phase velocities and higher attenuation per \( \mu m \). In conclusion, gate metal selectively modifies the optical plasmonic modes in the bilayer, while maintaining the acoustic modes as such.

Next, we validate theoretical calculations using our full-wave-hydrodynamic numerical solver. This solver will also be used to provide more insights on mechanisms of excitation of these modes in the bilayer.

### 5.3 Study of Excitation of Modes using Full-wave Hydrodynamic Solver

The developed 2D solver uses a finite difference time domain (FDTD) algorithm to solve Maxwell’s and hydrodynamic equations in one or more channels. This hybrid approach provides a self consistent modeling of the electron plasma-fluctuations and the associated fields in the media. A complete description of the solver is provided in previous chapters.

We consider a GaN/AlGaN bilayer channel of length 10\( \mu m \) covered with a metal-gate. To enable excitation from an incident plane wave, we introduce a small gap of length 200 nm in the gate. This is shown in Fig. 5.3. The parameters chosen for this simulation are same as those used for analytical solution in Section 5.2, except that \( d_1 = 28 \) nm and \( \Delta d = 20 \) nm. A broadband terahertz gaussian pulse is incident on the device and is allowed to couple to the bilayer. The data of the propagating pulse is recorded over time and subsequently Fourier transform is calculated to extract the phase and attenuation constants across a band of frequencies. To better understand the plasmon coupling and propagation, we plot the field distribution within and
around the channels. Specifically, we plot the $E_y$ (vertical) component to distinctly identify optical and acoustic modes. As is known, optical mode is characterized by large $E_y$ field component outside the bilayer, whereas the acoustic mode is mostly confined within the bilayer (channel-pair).

The obtained simulation results are plotted in Fig. 5.3. We observe that the full-wave simulation results for the phase and attenuation constants are in agreement with the analytical results. Specifically, simulation results coincide with the optical branch
Figure 5.4: Excitation and propagation of optical/acoustic modes in gated/ungated bilayer, (a) Simulation set-up used for the results shown in parts (b) and (c) (b) Fields in an ungated bilayer with a gap discontinuity (c) Fields in a gated bilayer with a gate discontinuity. We observe that spectral content of the oscillations can be controlled by choosing the appropriate gate or gap dimensions. Parameters used: 

\[ n_{o1} = n_{o2} = 5 \times 10^{12} \text{ cm}^{-2}, \mu = 10,000 \text{ cm}^2/\text{V.s}, d_1 = 28 \text{ nm}, \Delta d = 20 \text{ nm}. \]

meaning that the optical modes are excited predominantly. Fig. 5.3 also includes the time snapshots of the \( E_y \) fields at \( t = 1 \text{ ps}, 2 \text{ ps} \) and \( 3 \text{ ps} \). The field plots confirm
that the incident energy couples primarily to the optical modes in the bilayer. This is evident from small $E_y$ fields between the channels. It is well known that spectra in such experiments is dominated by the optical modes, but this simulation provides further insights. It reveals that the reason for weak acoustic oscillations is its weaker excitation rather than stronger attenuation. In other words, the gate discontinuity naturally favors the excitation of the optical modes. This can be reasoned by the fact that ambient incident fields are outside the bilayer - a field profile that matches better with the optical mode field profile. Acoustic modes are difficult to excite due to their confined field-profile that exists between the channels.

In summary, we find that numerical and analytical models are in good agreement. We also verify that optical modes are predominantly excited in a gated-bilayer with a gap discontinuity. We also note that in spite of the tendency for the incident fields to couple to optical modes, it is possible to selectively excite one mode over another by modifying the gate-gap. Specifically, the gate-gap (or other forms of discontinuity) can be tailored to excite a particular mode. We consider more simulations to test this premise.

First, we consider an ungated bilayer with a small gate that acts as a discontinuity (Fig. 5.4(a)). We use an incident plane wave as the excitation but the frequency is chosen to be 3 THz. It is known from the prior simulations that opposite field profiles exist at the two ends of the discontinuity. That is, the discontinuity region acts as a half wavelength cavity. Note that this wavelength can be calculated using (5.14). Therefore, we consider two cases: 1) when the gate length is $L_g = 330$ nm which is $\lambda_p/2$ for optical mode and 2) with $L_g = 170$ nm that is $\lambda_p/2$ for acoustic mode. We
plot the time snapshots for the field profiles at $t = 12.5$ ps (> steady state time) in Fig. 5.4(a). On the right, we also show the spectral content of the fluctuations as obtained by spatial Fourier transform of the electron density.

In case-1 ($L_g = 330$ nm), we observe almost 4 times larger wave amplitude for optical modes as compared that for the acoustic modes. This strong optical mode excitation comes from optical mode resonance under the gate. The said resonance mechanism is confirmed from the field plots, where we observe a large $E_y$ field present outside the bilayer. This generates an optical-mode field profile. On the contrary, in the second case ($L_g = 170$ nm), strong $E_y$ field appears between the channels. As expected, this is due to the acoustic mode resonance at the gate. This leads to increased acoustic content in the channels. In conclusion, by setting the gate to a resonant dimension, corresponding spectral content can be enhanced in the channel.

Next, a similar study is conducted for the gated-bilayer. Here the gate discontinuity size ($L_{gap}$) was modified. The results (shown in Fig 5.4(c)) show similar trends. That is, we observe that the ratio of power in these modes can be controlled by $L_{gap}$. Although, we note that total coupled power decreases for the smaller $L_{gap}$ case. This is due to overall decrease in the exposure of the bilayer to the ambient incident field.

The illustrative summary of modes and their field-profiles for gated- and ungated-bilayer and single layer cases are also shown in Fig. 5.6.
Table 5.1: 2DEG parameters used for the comparison of plasma-wave properties in gated single and double channel HEMTs. Results are shown in Fig. 5.5

<table>
<thead>
<tr>
<th></th>
<th>( n_{o1} ) (cm(^{-2}))</th>
<th>( n_{o2} ) (cm(^{-2}))</th>
<th>( d_1 ) (nm)</th>
<th>( d_2 ) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Channel</td>
<td>7.5 ( \times ) 10(^{12})</td>
<td>-</td>
<td>37</td>
<td>-</td>
</tr>
<tr>
<td>Double Channel</td>
<td>7.5 ( \times ) 10(^{12})</td>
<td>7.5 ( \times ) 10(^{12})</td>
<td>25</td>
<td>50</td>
</tr>
</tbody>
</table>

Figure 5.5: Comparison of the propagation constant and attenuation constants in single and double channel HEMTs. Used parameters are defined in Table-5.1
5.4 Comparison of Single and Double channel systems for Terahertz Devices

In this section, we compare the dispersion properties of the gated bilayer with the that of corresponding gated single layer. Specifically, we examine the potential of these devices as terahertz detectors. To that end, we consider a case of single and double channel gated HEMTs with parameters as chosen and shown in Table 5.1.

Calculated dispersion curves using (5.14) and (5.15), for single and double channel respectively, are plotted in Fig. 5.5. Due to ease of excitation and smaller attenuation, optical branch of the bilayer would be useful for practical applications. Therefore, we show only this branch in Fig. 5.5. It is clear that plasma-waves can propagate to longer distances when a double channel HEMT is used. This property is due to the increased electron density supported by the channel-pair. This could be important for the plasma-wave devices based on non-resonant plasma-wave propagation within the channel. More details on the resonant and non-resonant plasma mechanisms for the plasma-wave devices can be found in [21,22].

In addition to the above, double channel HEMTs are associated with smaller propagation constants. Such increased wavelengths would lead to longer channel lengths for the resonant plasma-wave devices operating at the same frequency. However, we must ascertain that losses per wavelength remain same as the channel length scales. Therefore, we also calculate this parameter for single and double channel HEMTs (Fig. 5.5). We find that the double channel and the single channel HEMTs show about the same loss performance. In conclusion, the advantage lies in the increased
channel dimensions offered by the double channel HEMTs. This leads to larger gate-lengths for the same operational frequency, thus allowing for improved fabrications. Alternatively, we can also say that the performance can be scaled to larger frequencies when using double channel HEMTs with small gate lengths.

As an example, let us consider a resonant plasma-wave detector at 2 THz. Using the parameters in Table-I and assuming channel to be completely covered by a conducting gate, we estimate that the needed channel length would be 315 nm ($\approx \lambda_p/4$) using the double channel HEMT, as opposed to 200 nm using a single channel HEMT. Alternatively, a double channel design with 200 nm channel length would have a resonance at 3 THz. Thus, large gate-length designs and/or higher operational frequencies are possible using the proposed concept.
5.5 Conclusion

We presented dispersion curves and closed form propagation constant formulae for gated bilayers which could be used for the design of terahertz detectors. The formulae are also verified using physical 2D simulations using the full-wave hydrodynamic numerical model. Our analysis showed that when compared to ungated bilayer, addition of a gate selectively affects the optical modes by decreasing their wavelengths and increasing the attenuation constants. Meanwhile, the acoustic modes largely remain unchanged. Using simulations, we further examine the excitation of these modes via an incident plane wave. We find that the most incident energy couples to the optical modes in these scenarios. However, acoustic modes can also be excited, given the discontinuity dimensions are tailored accordingly. We note that the double channel systems are of importance for resonant and non-resonant plasma-wave devices. Non-resonant devices can benefit due to smaller plasmonic losses. For the resonant devices, we can have larger channel lengths (and gate-lengths), allowing for easier fabrications with improved performance.
Chapter 6: Ongoing and Future Experimental Work

A number of experiments have been conducted in past to verify the understanding of terahertz plasmonic resonances in 2DEG systems [40, 41, 74]. Commonly, a periodic grating gate is placed over a 2DEG sheet for generating plasma-wave resonances. Effect of these resonances can be observed as transmission valleys in the spectra due to an impinging terahertz wave. In other words, researchers have used large-area bulk-mode resonances in the 2DEG-grating-gate system to validate the expected resonances in such devices. Clearly, terahertz spectroscopic measurement experimental set-up is ideal for conducting such measurements.

Most practical terahertz devices, however, utilize plasma-resonances within a single device (or using a single gate). For example, a resonant terahertz detector uses a single gated-2DEG resonance, as opposed to resonances in large area grating-gated structures. With that in mind, we develop experimental devices to demonstrate the presence of plasmonic-resonances in a finite length of 2DEG channel.
Figure 6.1: Device schematic of the proposed device for measurement of terahertz resonances in HEMT. (a) Top and cross-sectional views of the device. The measured electron density was $n_{sh} = 1.39 \times 10^{13}$ cm$^{-2}$ and mobility was $\mu = 1290$ cm$^2$/V.s (b) HFSS simulation model. (c) Field distribution at resonance frequency.

### 6.1 Capacitively Coupled HEMT Device

For the proposed experiment, circuit-based measurements are needed, since spectroscopic measurements will be ineffective for such micrometer scale devices. Therefore, the device and feeding mechanisms should be carefully designed.

The developed device is shown in Fig. 6.1. As shown, we use co-planar-waveguide (CPW) that is terminated with an open circuit. At the open end of the CPW, the device mesa is etched (below the CPW-line). The mesa is passivated using 40nm thick
Figure 6.2: Results of HFSS simulation modeling. *Left:* Variations in resonance frequency, seen as the peak of $\Re(Z_{in})$ as observed at the sample plane. Dotted-black curves are at 0.2 $\mu$m steps of $L_r$. *Center:* Phase($Z_{in}$) at the sample plane. *Right:* Normalized phase of $Z_{in}$ (phase is normalized by subtracting uniformly varying phase of -80° to -40° across the band)

Al₂O₃ layer. Thus, there is not ohmic contact between the input signal/ground lines and the heterojunction mesa. Note that we employ non-contact (antenna based)-one-port network measurement approach for our measurements. Details of the measurements and calibration method can be referred from [106]. We measure the reflection coefficients and using the CPW-line characteristic impedance ($Z_0=64$ $\Omega$), we calculate the input impedance of the 2DEG sample.

Motivations of this device and the measurement set-up come from another similar experiment conducted at microwave frequencies [107], where such plasma-wave resonances were measured using a AlGaAs/GaAs FET fed via a coaxial cable. In the experiment a 2DEG channel along with the metal gate and an open boundary condition was used to create a FET resonant cavity. Note that there was no ohmic contact between the 2DEG and source/drain terminals. That is, the cavity is excited by
capacitive coupling of input power to the 2DEG. The resonances were observed as peaks in real part of input impedance (as de-embedded to the device plane).

The modeling of the device along with the CPW line feeding was conducted using High Frequency Structural Simulation (HFSS tool). The 2DEG channel was modeled using Drude model, using the measured DC mobility and electron density data. Fig. 6.2 shows the $\Re(Z_{in})$ as obtained at the sample plane. As noted, we observed broad peaks in impedance, with variation of peaking frequency for different $L_r$. The variation of resonance frequency suggests that observed resonances are indeed plasma-wave resonances in the channel cavity.

### 6.2 Fabrication and Measurements

The sample were grown and processed by our collaborators at University of Notre Dame and Cornell University. Fig. 6.3(a) and (b) shows the devices along with feeding antenna structures. Several samples of varying $L_r$ were identified on the die and measurements were conducted.

The real part of the input impedance at the sample plane along with phase of the input impedance are shown in Fig. 6.3(c) and (d). Measurement results show that, we observe the expected trend of decreasing resonant frequency with increasing $L_r$. Although, the shift is the resonant frequency is quite small. We register a frequency shift of 15 GHz due to $L_r$ variation of 1.67 $\mu m$ (see Fig. 6.3(c)). According to simulation data this shift should be close to 150 GHz. We observe similar artifacts in the phase of the measured input impedance as shown in Fig. 6.4(d).
Figure 6.3: (a) Fabricated samples along with antenna structure for non-contact probing (b) Fabricated devices with varying sample lengths (c) Measured $\Re(Z_{in})$ for different samples (d) Measured phase and variation due to sample length variation.
6.3 Conclusions and Further Steps

We consider the results only partially successful. We conclude that the shift in the resonance frequencies is small, as only a small portion of 2DEG-sample actually participates in the wave-resonances. This could be due to inferior quality of 2DEG towards the edges of the mesa. However, this reasoning needs to be further verified. In order to improve the 2DEG edges, next step is to fabricate the devices using ion-implantation method. After this initial fabrication and experimental iteration, we are currently further pursuing these leads.
Chapter 7: Conclusions

7.1 Summary of the Work Conducted and Important Findings

In this dissertation work, rigorous computational methods are developed for the modeling of terahertz plasmonic devices. The developed time-domain methods utilize 2-D finite difference time domain (FDTD) method which is coupled to the electron transport model for accurate analysis. The so-called ‘full-wave-hydrodynamic’ solver can self-consistently solve multiphysical problem of modeling electron-transport and full-wave propagation within a device. Further, time-cost issues are addressed using unconditionally stable FDTD algorithms. We have also demonstrated its application for design, optimization and physical verification of active terahertz plasmonic devices and phenomenology. In this chapter, we provide summary of important findings of this dissertation work. The contributions presented in this work, in comparison to some of the other works in this area are presented in Fig. 6.1. Also, a summary of specific results pertaining to different methods and devices are presented in Fig. 6.2. We will briefly discuss these in the following.
One major drawback of the traditional FDTD-HD (or also called Global Modeling) is its extraneous time-cost. Due to small plasmonic wavelengths, the cell-sizes are small increasing the FDTD time-step and the overall simulation time. For efficient device design and optimization, we solve this issue by developing time-efficient models using ADI-FDTD and iterative ADI-FDTD methods for plasmonic applications. Specifically, we couple ADI-FDTD and iterative ADI-FDTD algorithms with hydrodynamic equations to obtain a class of plasmonic solvers—so-called ADI-FDTD-HD and iterative ADI-FDTD-HD models (presented in Chapter-3). Note that the application of ADI-FDTD-HD method for transistor modeling is done at microwave frequencies in a prior work. But in the presented work, these are considered for plasmonic applications where the propagation wavelengths are far smaller.

We find that when we choose large $\Delta t$ (or $CN$), ADI-FDTD methods do not work well for plasmonic applications. This is so because they omit the second order field-derivative term, which is dominant in the plasma-wave field regions near the 2DEG. Therefore, this splitting error (or truncation error) term leads to inaccuracies in the model.

To improve the accuracy for large $CN$, we develop iterative ADI-FDTD-HD method. This method corrects the errors associated with the splitting-error in subsequent iterations within each time-step. Thus, the truncation errors are drastically reduced at some added time-cost. Overall time-costs are still smaller than the explicit (or traditional) FDTD-HD method. For our considered example, we showed that a simulation-time reduction by a factor of 0.42 is achieved with a nominal error of 3% by the use of iterative ADI-FDTD-HD method.
We further use the developed models for the design and optimization of the terahertz devices. Specifically, we consider three devices: 1) short-channel terahertz emitter using plasma-wave instability; 2) RTD-gated HEMT for terahertz plasmonic-wave amplification and emission and 3) HEMT with bilayer channel configuration, enabling multi-mode plasmonic propagation. These devices are analyzed using the full-wave hydrodynamic solver allowing physical verification of the phenomenon and trends associated with their geometry. These analyses have provided valuable insights for design and optimization of these devices.

Specifically, we estimate that few tens of nW of terahertz emissions are possible from short-channel HEMT devices at cryogenic temperatures. We also attain valuable understanding of mechanism of radiation from such devices by looking at the near-field plots of these field radiations. We find that emissions occur by the virtue of drain and source contacts, which diffracts the plasma-wave fields enabling free-space radiation. It is interesting to note that such devices do not require the grating gate like coupling methods for radiation.

From the analyses of the RTD-gated HEMTs, we calculated the required values of the Resonant tunneling diode NDR (negative differential resistance) needed for plasma-wave growth. We found that for GaN/AlGaN system, 0.5 S/\(\mu\)m\(^2\) of NDR is needed for amplification. Although, we note that such high NDR values maybe too optimistic for GaN/AlGaN system currently. Nevertheless, development of low-loss active THz-waveguides is possible using these RTDs. For THz-sources and amplifiers, higher NDC values would be required, which could allow these applications in future.
New interesting physical insights are reported for 2D electron gas bilayer system. We verify the presence of acoustic and optical plasmonic modes via field visualization. The application of full-wave solvers allowed us to understand the methods for preferentially exciting one mode over the other. New analytical derivations, derived using hydrodynamic conductivity expression and EM boundary conditions, for these modes are also presented. Thus, the full-wave analyses and presence of modes was confirmed using the analytical method.

7.2 Future work

Overall, the developed methods are powerful tool for understanding the physics of electronic-plasmonic phenomenon and conducting accurate device simulations for a class of electronic-plasmonic devices. For future work, several possible research areas are open. First, a more rigorous modeling can be pursued by using particle based modeling for electron transport phenomenon. Such modeling will improve the model accuracy, especially in cases when the channel dimensions are small (few tens of nm) and/or high drain to source biases are used. Secondly, a more complete model can be achieved by incorporating 2D electron transport dynamics in the device. Current work considers 1D electron dynamics, i.e. electron transport equations in the horizontal direction (along the 2DEG channel). By including the electron transport in vertical direction the effect of gate-source bias in the device can also be modeled. A broader realm of applications can then be pursued using this model, including terahertz detector.
For physical understanding among multiple channel systems, plasma-wave dynamics in higher number of channels (more than two) can also be pursued. The developed numerical tools will not only reveal dominant modes for these systems, but also a method for coupling of power to a specific mode from terahertz incident wave. Note that, for higher number of channels, the number of modes increase linearly, e.g. for three channel systems, we expect three possible plasma-propagating modes. Thus, the developed tools can be applied for better understanding and more accurate designs and analysis for such cases.
Figure 7.1: Work presented in this dissertation as categorized in three blocks: 1) Physical understanding of plasmonic phenomenology, 2) New Numerical techniques and 3) Device modeling. The research in this dissertation, in relation to some other works in this area, is shown.
Summary of presented work and important findings reported in this dissertation:

- **Time-efficient model based on unconditionally stable FDTD-HD**
  - ADI-FDTD-HD models save time, but becomes inaccurate for large Δt or CN iterative ADI-FDTD-HD models provide accurate and time-efficient modeling.
  - Using a ADI-FDTD-HD modeling, time improvement by 0.42 times was achieved as compared to explicit FDTD-HD.

- **Modeling of THz emissions from Ungated-HEMT**
  - Numerical verification of Terahertz emissions from Short channel ungated HEMT devices.
  - Various findings are as following:
    - Few tens of nW of power possible from 1mm wide device at cryogenic temperatures.
    - Power coupled to free-space via interactions with Drain and Source terminals.
    - Power and frequency can be controlled with channel electron density and channel length.

- **Numerical Analysis of RTD-gated HEMT:**
  - Negative Differential Resistance of RTDs can be used for plasma-wave growth & amplification.
  - High NDR values ~0.5s/μm² are needed for such growth.
  - Emitter configurations are possible with grating-gated RTD-gated HEMTs.

- **Numerical Analysis of Gated/Ungated Bilayer System:**
  - Acoustic and Optical plasma-wave modes identified using analytical and FDTD-HD numerical models.
  - Respective mode excitation can be done by using half wavelength wide metal gate for respective modes.

---

Figure 7.2: Summary of presented work and important findings reported in this dissertation.
Appendix A: Time-Difference Equations for ADI-FDTD-HD and iterative-ADI-FDTD-HD methods

In this appendix, we will develop the difference equations for ADI-FDTD and it-ADI-FDTD methods that were used in Chapter 3. First we will consider the ADI-FDTD method.

A.1 ADI-FDTD-HD Difference Equations

We start with the matrix Equations 3.7 and 3.8

\[
\begin{align*}
\left( I - \frac{\Delta t}{2} [A] \right) \vec{u}^{tmp} &= \left( I + \frac{\Delta t}{2} [B] \right) \vec{u}^n, \\
\left( I - \frac{\Delta t}{2} [B] \right) \vec{u}^{n+1} &= \left( I + \frac{\Delta t}{2} [A] \right) \vec{u}^{tmp}.
\end{align*}
\]  

(A.1)

(A.2)

With \( \vec{u}, [A] \) and \([B]\) defined as

\[
\vec{u} = \begin{bmatrix} E_x \\ E_y \\ H_{zx} \\ H_{zy} \end{bmatrix}, \quad [A] = \begin{bmatrix} -\sigma_y & 0 & 1/\epsilon \partial_y & 1/\epsilon \partial_y \\ 0 & -\sigma_x & 0 & 0 \\ 0 & 0 & -\sigma^*_{xy} & 0 \\ 1/\mu \partial_y & 0 & 0 & -\sigma^*_{xy}/\mu \end{bmatrix}, \quad [B] = \begin{bmatrix} -\sigma_y/2\epsilon & 0 & 0 & 0 \\ 0 & -\sigma_x/2\epsilon & -1/\epsilon \partial_x & -1/\epsilon \partial_x \\ 0 & -1/\mu \partial_y & -\sigma^*_{xy}/2\mu & 0 \\ 0 & 0 & 0 & -\sigma^*_{xy}/2\mu \end{bmatrix},
\]

with \( H_z \) being sum of \( H_{zx} \) and \( H_{zy} \) ( or \( H_z = H_{zx} + H_{zy} \)).

\[
E_x^{tmp} = C_{exe} E_x^n + \frac{\Delta t}{2\epsilon} C_{exh} \frac{\partial}{\partial y} (H_{zx}^{tmp} + H_{zy}^{tmp}).
\]  

(A.3)
We will consider the sub-time-step equations and full-time-step equations, one by one in following sections.

**Update-Equations for sub-time-step \( \text{tmp} \)**

We expand A.1 to read

\[
E_x^{\text{tmp}} = C_{exe} E_x^n + \frac{\Delta t}{2\epsilon} C_{exh} \frac{\partial}{\partial y}(H_z^{\text{tmp}} + H_y^{\text{tmp}}) \tag{A.4}
\]

\[
E_y^{\text{tmp}} = C_{eye} E_y^n - \frac{\Delta t}{2\epsilon} C_{eyh} \frac{\partial}{\partial x}(H_x^{\text{tmp}} + H_y^{\text{tmp}}) \tag{A.5}
\]

\[
H_x^{\text{tmp}} = C_{hzhx} H_x^n - \frac{\Delta t}{2\mu} C_{hzey} \frac{\partial}{\partial y}(E_x^{\text{tmp}}) \tag{A.6}
\]

\[
H_y^{\text{tmp}} = C_{hzhy} H_y^n + \frac{\Delta t}{2\mu} C_{hzex} \frac{\partial}{\partial x}(E_y^{\text{tmp}}) \tag{A.7}
\]

with \( C \)-coefficients defined as

\[
C_{exe} = 1 - \frac{\Delta t \sigma_x}{4\epsilon}, \quad C_{exh} = \frac{1}{1 + \frac{\Delta t \sigma_x}{4\epsilon}}, \quad C_{eye} = 1 - \frac{\Delta t \sigma_x}{4\epsilon}, \quad C_{eyh} = \frac{1}{1 + \frac{\Delta t \sigma_x}{4\epsilon}}, \\
C_{hzhx} = 1 - \frac{\Delta t \sigma_x}{4\mu}, \quad C_{hzey} = \frac{1}{1 + \frac{\Delta t \sigma_x}{4\mu}}, \quad C_{hzhy} = 1 - \frac{\Delta t \sigma_y}{4\mu}, \quad C_{hzex} = \frac{1}{1 + \frac{\Delta t \sigma_y}{4\mu}} \tag{A.8}
\]

Note that contrary to conventional FDTD, where the equations are explicitly defined, Equations A.3 - A.6 are not explicit equations, i.e., fields at time-step \( \text{tmp} \) are not defined explicitly in terms of fields at time-step \( n \). Using these, a simplified implicit equation for variable \( E_x^{\text{tmp}} \) can be obtained. After few algebraic substitutions, we obtain

\[
E_x^{\text{tmp}} - \frac{\Delta t^2}{2\epsilon} C_{exh} \frac{\partial}{\partial y} \left[ C_{hyex} \frac{\partial}{\partial y}(E_x^{\text{tmp}}) \right] = C_{exe} E_x^n + \frac{\Delta t}{2\epsilon} C_{exh} \frac{\partial}{\partial y} \left[ C_{hzhy} H_y^n + C_{hzhy} H_y^n \right] - \frac{\Delta t^2}{2\epsilon} C_{exh} \frac{\partial}{\partial y} \left[ C_{hzey} \frac{\partial}{\partial x}(E_y^n) \right] \tag{A.9}
\]
First order difference equation based on non-uniform space grid for above equation can be written as

\[
E_{x|_{i,j+1/2}}^{tmp} \left[ 1 + \Delta t^2 \left( \frac{C_{ehx}}{2\epsilon \Delta y} \right)_{i,j+1/2} \left( \frac{C_{zhy}}{2\mu \Delta y} \right)_{i,j+1} + \Delta t^2 \left( \frac{C_{ehx}}{2\epsilon \Delta y} \right)_{i,j+1/2} \left( \frac{C_{zhy}}{2\mu \Delta y} \right)_{i,j} \right] \\
+ E_{x|_{i,j-1/2}}^{tmp} \left[ -\Delta t^2 \left( \frac{C_{ehx}}{2\epsilon \Delta y} \right)_{i,j+1/2} \left( \frac{C_{zhy}}{2\mu \Delta y} \right)_{i,j+1} \right] \\
+ E_{x|_{i,j+1/2}}^{tmp} \left[ -\Delta t^2 \left( \frac{C_{ehx}}{2\epsilon \Delta y} \right)_{i,j+1/2} \left( \frac{C_{zhy}}{2\mu \Delta y} \right)_{i,j+1} \right] \\
= C_{exx|_{i,j+1/2}} E_{x|_{i,j+1/2}}^{n} + \Delta t \left( \frac{C_{exx}}{2\epsilon \Delta y} \right)_{i,j+1} \left[ C_{zhy|_{i,j+1}} H_{zy|_{i,j+1}}^{n} + C_{zhy|_{i,j+1}} H_{zy|_{i,j+1}}^{n} \right] \\
- C_{zhy|_{i,j}} H_{zy|_{i,j}}^{n} - C_{zhy|_{i,j}} H_{zy|_{i,j}}^{n} - \Delta t^2 \left( \frac{C_{ehx}}{2\epsilon \Delta y} \right)_{i,j+1} \left[ \left( \frac{C_{ehx}}{2\mu \Delta x} \right)_{i,j+1} \right] \\
- E_{y|_{i-1/2,j+1}}^{n} - \left( \frac{C_{ehx}}{2\mu \Delta y} \right)_{i,j} \left( E_{y|_{i+1/2,j}}^{n} - E_{y|_{i-1/2,j}}^{n} \right) \].
\]

(A.10)

Clearly, above is a tridiagonal system of linear simultaneous equations for variable \(E_{x|_{i,j+1/2}}^{tmp}\), with \(j\) varying from 1 to \(N\), \(N\) being number of cells in \(y\)-direction. The known quantities and variables involved in this equation are graphically represented in Fig. A.1. As noted, the stencil moves in vertical direction as \(j\) is varied from 1 to \(N\).

Having obtained \(E_{x|_{i,j+1/2}}^{tmp}\), we can then use difference equation corresponding to A.4 to A.7 to obtain other field values for sub-time-step \(tmp\)

\[
E_{y|_{i-1/2,j}}^{tmp} = C_{eye|_{i-1/2,j}} E_{y|_{i-1/2,j}}^{n} \\
- \Delta t \left( \frac{C_{eyh}}{2\epsilon \Delta x} \right)_{i-1/2,j} \left[ H_{zy|_{i,j}}^{n} + H_{zy|_{i,j}}^{n} - H_{zy|_{i-1,j}}^{n} + H_{zy|_{i-1,j}}^{n} \right] \]
\]

(A.11)

\[
H_{xz|_{i,j}}^{tmp} = C_{hzhx|_{i,j}} H_{xz|_{i,j}}^{n} - \Delta t \left( \frac{C_{hxyz}}{2\mu \Delta x} \right)_{i,j} \left[ E_{y|_{i+1/2,j}}^{n} - E_{y|_{i-1/2,j}}^{n} \right] \]
\]

(A.12)

\[
H_{zy|_{i,j}}^{tmp} = C_{hzhy|_{i,j}} H_{zy|_{i,j}}^{n} + \Delta t \left( \frac{C_{hzx}}{2\mu \Delta y} \right)_{i,j} \left[ E_{x|_{i,j+1/2}}^{tmp} - E_{x|_{i,j-1/2}}^{tmp} \right] \]
\]

(A.13)
Figure A.1: Mesh-grid scheme and stencil illustration for Equation A.10. Iterative terms only apply for Equations A.27

Thus, by using the steps A.9 - A.12, we can obtain the fields for the sub-time-step. Next we apply similar process obtain the fields at the time-step $n + 1$. 
Update-Equations for full-time-step \( n+1 \)

We use similar steps as before but now starting from Equation-A.2 to obtain the implicit difference equation in variable \( E_y|_{i-1/2,j}^{n+1} \)

\[
E_y|_{i-1/2,j}^{n+1} = 1 + \Delta t^2 \left( \frac{C_{eyh}}{2\epsilon \Delta x} \right)_{i-1/2,j}^{i,j} + \Delta t^2 \left( \frac{C_{hzy}}{2\mu \Delta x} \right)_{i-1/2,j}^{i,j} 
+ E_y|_{i+1/2,j}^{n+1} \left[ - \Delta t^2 \left( \frac{C_{eyh}}{2\epsilon \Delta x} \right)_{i-1/2,j}^{i,j} \right] 
+ E_y|_{i-1/2,j}^{n+1} \left[ - \Delta t^2 \left( \frac{C_{eyh}}{2\epsilon \Delta x} \right)_{i-1/2,j}^{i,j} \right] 
\]

\[
E_y|_{i-1/2,j}^{n+1} = C_{eyh} E_y|_{i-1/2,j}^{tmp} - \Delta t \left( \frac{C_{eyh}}{2\epsilon \Delta y} \right)_{i-1/2,j}^{i,j} \left[ C_{hzy} \left| H_y \right|_{i,j}^{tmp} + C_{hzy} \left| H_z \right|_{i,j}^{tmp} \right] 
- \Delta t^2 \left( \frac{C_{eyh}}{2\epsilon \Delta y} \right)_{i-1/2,j}^{i,j} \left[ \frac{C_{hzy}}{2\mu \Delta y} \right]_{i,j} \left( E_x \right)_{i,j+1/2}^{tmp} 
- E_x \left| (i,j-1/2) \right|^{tmp} \left( \frac{C_{hzy}}{2\mu \Delta y} \right)_{i-1,j}^{i,j} \left( E_x \right)_{i-1,j+1/2}^{tmp} - \left( E_x \right)_{i-1,j-1/2}^{tmp} \left( E_x \right)_{i-1,j-1/2}^{tmp} \right]. 
\]

(A.14)

Above is the set of simultaneous equations generated by varying \( i \) from 1 to \( M \), \( M \) being the number of cells in horizontal direction. Again, above can be graphically represented by a horizontally moving stencil on the mesh-grid as shown in Fig. A.2.

Update equations for \( E_y, H_zx \) and \( H_zy \) field components are

\[
E_x|_{i,j+1/2}^{n+1} = C_{exe} E_x|_{i,j+1/2}^{tmp} 
- \Delta t \left( \frac{C_{exe}}{2\epsilon \Delta y} \right)_{i,j+1/2}^{i,j+1/2} \left[ H_{zx} \left|_{i,j+1}^{tmp} + H_{yx} \left|_{i,j+1}^{tmp} - H_{zx} \left|_{i,j}^{tmp} - H_{yz} \left|_{i,j}^{tmp} \right] \right. \right. 
\]

(A.15)

\[
H_{zx}|_{i,j}^{n+1} = C_{hzx} H_{zx}|_{i,j}^{tmp} - \Delta t \left( \frac{C_{hxy}}{2\mu \Delta x} \right)_{i,j}^{i,j} \left[ E_y|_{i,j+1}^{n+1} - E_y|_{i,j-1/2}^{n+1} \right] 
\]

(A.16)

\[
H_{zy}|_{i,j}^{n+1} = C_{hzy} H_{zy}|_{i,j}^{tmp} + \Delta t \left( \frac{C_{hzy}}{2\mu \Delta y} \right)_{i,j}^{i,j} \left[ E_x|_{i,j+1}^{tmp} - E_x|_{i,j-1}^{tmp} \right] 
\]

(A.17)
Equations A.9-A.12 and A.13-A.16 are set of equations that be implemented using a computer code for developing the ADI-FDTD method. For our simulations, we used MATLAB for such implementation.

### A.2 Iterative ADI-FDTD-HD Difference Equations

For iterative ADI solution, we start with Equations 3.10 and 3.11

\[
\left( I - \frac{\Delta t}{2} [A] \right) \tilde{u}^\text{tmp}_{k+1} = \left( I + \frac{\Delta t}{2} [B] \right) \tilde{u}^n + \frac{\Delta t^2}{8} [A][B] \left( \tilde{u}^{n+1}_k - \tilde{u}^n \right). \quad (A.18)
\]

\[
\left( I - \frac{\Delta t}{2} [B] \right) \tilde{u}^{n+1}_{k+1} = \left( I + \frac{\Delta t}{2} [A] \right) \tilde{u}^\text{tmp}_{k+1} + \frac{\Delta t^2}{8} [A][B] \left( \tilde{u}^{n+1}_k - \tilde{u}^n \right). \quad (A.19)
\]
We note that in addition to terms in A.1 and A.2, the additional term

\[ \frac{\Delta t^2}{8} [A][B] \left( \vec{u}_{k+1}^{n+1} - \vec{u}^n \right) \]  

(A.20)
can be expanded as

\[ \frac{\Delta t^2}{8} \begin{bmatrix}
\frac{\sigma_x^*}{4\epsilon^2} & -1 \frac{\partial}{\partial y} \frac{\partial}{\partial x} & -1 \frac{\partial}{\partial y} \frac{\partial}{\partial x} & -1 \frac{\partial}{\partial y} \frac{\partial}{\partial x} \\
0 & \frac{\sigma_y^*}{4\mu} & 0 & 0 \\
-\frac{1}{\mu} \frac{\partial}{\partial y} \left( \frac{\sigma_y}{2\epsilon} \right) & 0 & 0 & \frac{(\sigma_y^*)^2}{4\mu^2}
\end{bmatrix}
\begin{bmatrix}
E_{x}^{n+1} |_{k} - E_{x}^{n} \\
E_{y}^{n+1} |_{k} - E_{y}^{n} \\
H_{x}^{n+1} |_{k} - H_{x}^{n} \\
H_{z}^{n+1} |_{k} - H_{z}^{n}
\end{bmatrix}. \]  

We do not apply the iterative correction for PML or metallic regions. Therefore, we set \( \sigma_x^* = \sigma_y^* = \sigma_x = \sigma_y = 0 \) and get

\[ \frac{\Delta t^2}{8} \begin{bmatrix}
-\frac{1}{\epsilon} \frac{\partial}{\partial y} \left( \frac{1}{\mu} \frac{\partial}{\partial x} \right) \left( E_{y}^{n+1} |_{k} - E_{y}^{n} \right) \\
0 \\
0 \\
0
\end{bmatrix} \]  

(A.21)

Thus, only first partial differential equation (or \( E_x \)-update equation) is modified due to the corrective term.

Using the above corrective term, we next derive the difference-equations for the sub-time-step and full-time-step for iterative-ADI-FDTD method.

**Update-Equations for sub-time-step tmp**

We start with equation A.17 and expand the terms after adding the iterative corrective terms as derived in A.20.

\[ E_x^{tmp} |_{k+1} = C_{exe} E_x^n + \frac{\Delta t}{2\epsilon} C_{exh} \frac{\partial}{\partial y} \left( H_{zx}^{tmp} |_{k+1} + H_{zy}^{tmp} |_{k+1} \right) \\
- C_{exh} \frac{\Delta t^2}{2\epsilon} \frac{\partial}{\partial y} \left( \frac{1}{\mu} \frac{\partial}{\partial x} \right) \left( E_{y}^{n+1} |_{k} - E_{y}^{n} \right) \]  

(A.22)
\[ E_{y}^{\text{tmp}}|_{k+1} = C_{eye}E_{y}^{n} - \frac{\Delta t}{2\epsilon}C_{eyh} \frac{\partial}{\partial x}(H_{zx}^{n} + H_{zy}^{n}) \] (A.23)

\[ H_{zx}^{\text{tmp}}|_{k+1} = C_{hzhx}H_{zx}^{n} - \frac{\Delta t}{2\mu}C_{hzvy} \frac{\partial}{\partial x}E_{y}^{n} \] (A.24)

\[ H_{zy}^{\text{tmp}}|_{k+1} = C_{hzhx}H_{zy}^{n} + \frac{\Delta t}{2\mu}C_{hzex} \frac{\partial}{\partial x}E_{x}^{\text{tmp}}|_{k+1} \] (A.25)

We apply similar steps as in previous sections to obtain the implicit differential equation with iterative term

\[
\begin{align*}
E_{x}^{\text{tmp}}|_{k+1} & - \frac{\Delta t^2}{2\epsilon}C_{exh} \frac{\partial}{\partial y} \left[ C_{hyst} \frac{\partial}{\partial y} (E_{x}^{\text{tmp}}|_{k+1}) \right] \\
& = C_{exe}E_{x}^{n} + \frac{\Delta t}{2\epsilon}C_{exh} \frac{\partial}{\partial y} \left[ C_{hzyH_{zy}^{n}} + C_{hzyH_{zy}^{n}} \right] - \frac{\Delta t^2}{2\epsilon}C_{exh} \frac{\partial}{\partial y} \left[ C_{hzey} \frac{\partial}{\partial x} (E_{y}^{n}) \right] \\
& - C_{exh} \frac{\Delta t^2}{8\epsilon} \frac{\partial}{\partial y} \left( \frac{1}{\mu} \frac{\partial}{\partial x} \right) (E_{y}^{n} |_{k+1} - E_{y}^{n})
\end{align*}
\] (A.26)
Likewise, difference equations can be obtained from above. The implicit difference

equation for $E_x$-update takes following form-

\[
E_x^{\text{tmp},k+1}_{i,j+1/2} = 1 + \Delta t^2 \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1/2} \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j+1} + \Delta t^2 \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1/2} \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j} \\
+ E_x^{\text{tmp},k+1}_{i,j-1/2} \left[ - \Delta t^2 \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1/2} \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j+1} \right] \\
+ E_x^{\text{tmp},k+1}_{i,j+3/2} \left[ - \Delta t^2 \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1/2} \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j+1} \right] \\
= C_{exe|i,j+1/2}^{n}E_x^{n}_{i,j+1/2} + \Delta t \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1/2} \left[ C_{hzhy|i,j+1}^{n}H_{zy}^{n}_{i,j+1} + C_{hzhy|i,j+1}^{n}H_{zy}^{n}_{i,j+1} \\
- C_{hzhy|i,j}^{n}H_{zy}^{n}_{i,j} + C_{hzhy|i,j}^{n}H_{zy}^{n}_{i,j} \right] - \Delta t^2 \left( \frac{C_{exh}}{2\epsilon\Delta y} \right)_{i,j+1} \left[ \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j+1} \left( E_y^{n}_{i,j+1/2} - E_y^{n}_{i,j-1/2} \right) \right] \\
- \frac{\Delta t^2}{8} \left( \frac{C_{exh}}{\epsilon\Delta y} \right)_{i,j+1/2} \left( \frac{1}{\mu\Delta x} \right)_{i,j+1} \left\{ \left( E_y^{n+1,k} - E_y^{n}_{i+1/2,j+1} \right) - \left( E_y^{n+1,k} - E_y^{n}_{i-1/2,j+1} \right) \right\} \\
- \frac{1}{\mu\Delta x} \left\{ \left( E_y^{n+1,k} - E_y^{n}_{i+1/2,j} \right) - \left( E_y^{n+1,k} - E_y^{n}_{i-1/2,j} \right) \right\} \\
\tag{A.27}
\]

Other update equations are similar to ADI-FDTD case,

\[
E_y^{\text{tmp},k+1}_{i-1/2,j} = C_{eyh|i-1/2,j}^{n}E_y^{n}_{i-1/2,j} - \Delta t \left( \frac{C_{eyh}}{2\epsilon\Delta x} \right)_{i-1/2,j} \left[ H_{zy}^{n}_{i,j} + H_{zy}^{n}_{i-1,j} - H_{zy}^{n}_{i+1,j} + H_{zy}^{n}_{i-1,j} \right] \tag{A.28}
\]

\[
H_{zx}^{\text{tmp},k+1}_{i,j} = C_{hzx|i,j}^{n}H_{zx}^{n}_{i,j} - \Delta t \left( \frac{C_{hzx}}{2\mu\Delta y} \right)_{i,j} \left[ E_y^{n}_{i+1/2,j} - E_y^{n}_{i-1/2,j} \right] \tag{A.29}
\]

\[
H_{zy}^{\text{tmp},k+1}_{i,j} = C_{hzhy|i,j}^{n}H_{zy}^{n}_{i,j} + \Delta t \left( \frac{C_{hzex}}{2\mu\Delta y} \right)_{i,j} \left[ E_x^{\text{tmp},k+1}_{i,j+1/2} - E_x^{\text{tmp},k+1}_{i,j-1/2} \right] \tag{A.30}
\]
Update-Equations for full-time-step $n+1$

Using the steps similar to previous section, we again derive the difference equation for the $n + 1$ time-step calculation from $tmp$ time-step. As expected for this step, implicit equation is for $E_y$-update, as shown below

$$E_{y|n+1/2,j}^{|n+k+1} + 1 + \Delta t^2 \left( \frac{C_{eyh}}{2 \epsilon \Delta x} \right)_{i-1/2,j} \left( \frac{CH_{zey}}{2 \mu \Delta x} \right)_{i,j} + \Delta t^2 \left( \frac{C_{eyh}}{2 \epsilon \Delta x} \right)_{i-1/2,j} \left( \frac{CH_{zey}}{2 \mu \Delta x} \right)_{i-1,j}$$

$$+ E_{y|n+1/2,j}^{n+k+1} - \Delta t^2 \left( \frac{C_{eyh}}{2 \epsilon \Delta x} \right)_{i-1/2,j} \left( \frac{CH_{zey}}{2 \mu \Delta x} \right)_{i,j}$$

$$+ E_{y|n+1/2,j}^{n+k+1} - \Delta t^2 \left( \frac{C_{eyh}}{2 \epsilon \Delta x} \right)_{i-1/2,j} \left( \frac{CH_{zey}}{2 \mu \Delta x} \right)_{i-1,j}$$

$$= C_{eyh} \left( E_{y|n+1/2,j}^{tmp,k+1} - \Delta t \left( \frac{C_{eyh}}{2 \epsilon \Delta x} \right)_{i-1/2,j} \left[ CH_{zey} | i,j \right] H_{zy} | i,j \right) + CH_{zey} | i,j \left[ CH_{zey} | i,j \right] H_{zy} | i,j \right)$$

$$- CH_{zey} | i-1,j \left[ CH_{zey} | i-1,j \right] H_{zy} | i-1,j \right) - \Delta t^2 \left( \frac{C_{eyh}}{2 \epsilon \Delta y} \right)_{i-1/2,j} \left[ \left( \frac{CH_{exy}}{2 \mu \Delta y} \right)_{i,j} \left( E_{x|n+1/2,j}^{tmp,k+1} \right) \right.$$}

$$- E_{x|n+1/2,j}^{tmp,k+1} \right) - \left( \frac{CH_{exy}}{2 \mu \Delta y} \right)_{i-1,j} \left( E_{x|n+1/2,j}^{tmp,k+1} - E_{x|n+1/2,j}^{tmp,k+1} \right).$$

(A.31)

Update equations $E_x$, $H_{xx}$ and $H_{zy}$ are shown below. Note that $E_x$ update equation are modified with iterative terms.

$$E_{x|n+1/2,j}^{n+k+1} = C_{exe} \left( E_{x|n+1/2,j}^{tmp,k+1} \right.$$}

$$- \Delta t \left( \frac{C_{exh}}{2 \epsilon \Delta y} \right)_{i,j+1/2} \left[ H_{zx} | i,j+1 \right] + H_{zy} | i,j+1 \right) - H_{xx} | i,j \right) - H_{zy} | i,j \right) \right)$$

$$- \Delta t^2 \left( \frac{C_{exh}}{2 \epsilon \Delta y} \right)_{i,j+1/2} \left( \left( \frac{1}{\mu \Delta x} \right)_{i,j+1} \left( E_{x|n+1/2,j} \right.\right.\right.$$}

$$- E_{x|n+1/2,j} \right) \right)$$

$$- \left( \frac{1}{\mu \Delta x} \right)_{i,j} \left( \left( E_{x|n+1/2,j} \right.\right.\right.$$}

(A.32)
\[ H_{z}^{n+1,k+1}_{i,j} = C_{h_{z}h_{x}}|_{i,j} H_{z}^{tmp,k+1}_{i,j} - \Delta t \left( \frac{C_{h_{z}e_{y}}}{2\mu\Delta x} \right)_{i,j} \left[ E_{y}^{n+1,k+1}_{i+1/2,j} - E_{y}^{n+1,k+1}_{i-1/2,j} \right] \] (A.33)

\[ H_{z}^{n+1,k+1}_{i,j} = C_{h_{z}f_{x}}|_{i,j} H_{z}^{tmp,k+1}_{i,j} + \Delta t \left( \frac{C_{h_{z}e_{x}}}{2\mu\Delta y} \right)_{i,j} \left[ E_{x}^{tmp,k+1}_{i,j+1/2} - E_{x}^{tmp,k+1}_{i,j-1/2} \right] \] (A.34)

Thus, by using the Equations A.27-A.34, iterative ADI-FDTD solution can be implemented.
Bibliography


151

