Nonlinear Optical Mechanisms in Semiconductors and Enhanced Nonlinearities at Epsilon-Near-Zero

Sepehr Ahmadzadeh Benis

University of Central Florida

Find similar works at: https://stars.library.ucf.edu/etd2020

University of Central Florida Libraries http://library.ucf.edu

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation
NONLINEAR OPTICAL MECHANISMS IN SEMICONDUCTORS AND ENHANCED NONLINEARITIES AT EPSILON-NEAR-ZERO

by

SEPEHR BENIS
B.S. Electrical Engineering, University of Tehran, 2014
M.S. Optics and Photonics, University of Central Florida, 2016

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in CREOL, the College of Optics and Photonics at the University of Central Florida Orlando, Florida

Summer Term 2020

Major Professors: David J. Hagan and Eric W. Van Stryland
ABSTRACT

Light does not interact with itself in linear optical materials. Such interactions occur only in non-linear optical (NLO) materials and typically require high intensity optical beams to be significant. The ever-increasing role of NLO, where intense light may change the properties of the medium, has created a pressing demand to invent materials for achieving more efficient light-light and light-matter interaction due to their potential capacity to augment and possibly replace current technologies with more efficient devices. There are numerous applications of NLO devices in fundamental science, technology, health, and defense such as all-optical computation and signal processing, ultrashort laser technology, photodynamic cancer therapy, and quantum communication and information.

The main objective of my Ph.D. dissertation is to investigate the interaction of laser pulses with an exciting class of material that has the dielectric constant close to zero, so-called epsilon-near-zero (ENZ). The ENZ materials and their scientific development have become a topic of interest owing to their fascinating nonlinear optical properties particularly in frequency ranges that the material is transitioning from dielectric to metal. The goal of my dissertation is the theoretical and experimental study of transparent conducting oxides such as Indium Tin Oxide (ITO) as a candidate material exhibiting the ENZ condition and utilizing this effect for nonlinear optical devices. The NLO effects in TCOs are dominated by carrier related nonlinearities. Additionally, this dissertation studies instantaneous third-order and non-instantaneous carrier nonlinearities in semiconductors such as GaAs and Silicon and fifth-order nonlinear absorption (three-photon absorption) in direct gap semiconductors.

In this work, we first introduce the development of NLO spectroscopy systems for the characterization of the NLO properties. In particular, the Beam-Deflection (BD) technique, which
allows us to simultaneously characterize the nonlinear refraction and absorption of the material. This technique enables us to characterize the sign, magnitude, temporal and polarization dependence of the nonlinearities, which all are of paramount importance in understanding the underlying physics of the material. We also extend BD to a cross-propagating geometry to measure off-diagonal nonlinear susceptibility matrix elements. Moreover, we employ BD to measure ultrafast and carrier-induced nonlinear absorption and refraction of common semiconductor materials and transparent conducting oxides. The ability of BD to measure the time-dynamics and polarization-dependence of the nonlinear phase shift becomes apparent in both subjects presented in this dissertation.
“Dedicated to my amazing wife, Behnaz, and to my wonderful family.”
ACKNOWLEDGMENTS

Perhaps the most difficult part of this dissertation is to find words to convey my appreciation to many people contributed to this work.

I truly appreciate my Ph.D. advisors Prof. Eric W. Van Stryland, Prof. David J. Hagan, and Prof. MJ Soileau for supporting me during my Ph.D. journey and giving me the opportunity to do research in the NLO group. They have been extremely helpful and supportive in many ways during these years and provided me not only the needed guidance to complete the dissertation work, but also numerous opportunities to advance my professional career. I would also like to thank my committee members Prof. Demetrios N. Christodoulides, Prof. Pieter Kik, and Prof. Arkadiy Lyakh for their valuable input and great interest in my dissertation research. I also thank Prof. Ayman Abouraddy for his guidance and numerous conversations that we had on the prospects of being a researcher.

I am extremely thankful to all the former and current members of the NLO group for their help and support, which are indispensable in every detail of this dissertation. I thank my colleagues and friends Dr. Trenton Ensley, Dr. Matthew Reichert, Dr. Himansu Pattanaik, and Dr. Peng Zhao who taught me most of the essential skills needed in my research. I would also like to thank Dr. Natalia Munera for her help throughout these years in conducting some of the experimental work and discussing ideas. Also, many thanks to other NLO group members whom I have worked with during my Ph.D., including Dr. Mike Bondar, Dr. Joel Hales, Dr. Scott Webster, Dr. Salimeh Tofighi, Munan Gao, Nicholas Cox, Hao-Jung (Wilbur) Chang, and Sanaz Faryadras. My appreciation goes the countless support of CREOL staff providing us with the best academic environment. Special thanks to Rachel, Alma, Gail, Amy, Mark, James, Denise, Evelyn, Maria, Matt,
Mike, Deon, Richard, April, and Sina. Many thanks to our collaborators for their valuable contributions and input from Northwestern University, Georgia Institute of Technology, University of Washington, Virginia Commonwealth University, University of Michigan, and Purdue University.

I would also like to thank all my friends who helped me, inspired me, had a drink with me, and made conversation with me during these years. I, particularly, thank Yolanda, Ramtin and Golareh, Nader and Elham, Murat, Isa, Abbas, Milad, Walker, Amin, and Amir Mahdi. I enjoyed every bit of time spent with any of you.

Finally, I would like to thank my father, Salman, and my mother, Raheleh, for their continuous and unparalleled love, their continued support, and their patience. Thank you both for giving me strength to reach for the stars and chase my dreams. I am extremely grateful to my sister, Samin, for always being there for me as a friend. Her unconditional love and kindness inspires me every day. My deepest appreciation goes to my wife, Behnaz, for her patience, support, inspiration, and love over all these years. She has stood by me through all my travails, my absences, and my impatience. She gave me support and help, discussed ideas in every aspect, and prevented several wrong turns. Cheers to all of them!
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... xi

LIST OF TABLES ........................................................................................................... xx

LIST OF ACRONYMS AND ABBREVIATIONS .............................................................. xxi

CHAPTER 1: INTRODUCTION ......................................................................................... 1

1.1. Bound-electronic and carrier-induced nonlinearities in semiconductors .................. 2

1.2. Enhanced Optical Nonlinearities at Epsilon-Near-Zero .......................................... 3

1.3. Dissertation Outline ............................................................................................... 5

CHAPTER 2: NONLINEAR ABSORPTION AND REFRACTION IN SOLIDS ............... 7

2.1. Nonlinear Wave Equation ...................................................................................... 8

2.1.1. Solutions of nonlinear wave equation .................................................................. 11

2.2. Nonlinear Absorption and Refraction of Bound Electrons in Semiconductors ......... 13

2.2.1. Power series expansion of nonlinear polarization ............................................... 13

2.2.2. Third-order bound-electronic nonlinearities ....................................................... 15

2.2.3. Fifth-order bound-electronic nonlinearities ......................................................... 17

2.3. Nonlinear Absorption and Refraction of Free-Carriers in Semiconductors and Metals... 19

2.3.1. Free-carrier absorption and refraction in semiconductors .................................. 19

2.3.2. Drude model and nonlinearities of metals ......................................................... 24

2.3.3. Extremely Large Nonlinear Optical Effects and Failure of Perturbative Analysis ... 26
CHAPTER 3: EXPERIMENTAL TECHNIQUES ................................................................. 27

3.1. Linear Spectroscopy Techniques ........................................................................ 27

3.1.1. Spectroscopic ellipsometry ........................................................................... 28

3.2. Ultrashort Light Sources ..................................................................................... 29

3.2.1. Laser Sources ................................................................................................. 29

3.2.2. Nonlinear optical parametric amplifiers ......................................................... 31

3.3. Nonlinear Spectroscopy Techniques .................................................................. 32

3.3.1. Z-scan ........................................................................................................... 32

3.3.2. Beam-Deflection ......................................................................................... 41

3.3.3. Cross-propagating Beam-Deflection .............................................................. 51

3.3.4. Temporal modulation of nonlinear phase ...................................................... 57

CHAPTER 4: BOUND-ELECTRONIC AND CARRIER-INDUCED NONLINEARITIES IN SEMICONDUCTORS ........................................................................ 59

4.1. Introduction and Motivations ............................................................................. 59

4.2. Third-order Bound-Electronic Nonlinearities of Direct-Gap Semiconductors .... 61

4.3. 2PA-induced Carrier Nonlinearities in Semiconductors .................................... 65

4.4. Measurements of the Instantaneous Third-order and Non-Instantaneous Fifth-Order Nonlinearities in GaAs and Silicon ......................................................... 66

4.5. Three-Photon Absorption in Direct Gap Semiconductors ................................ 80

4.6. Concluding Remarks ......................................................................................... 87
LIST OF FIGURES

Figure 3-1. Schematic setup for spectroscopic ellipsometry apparatus. ........................................ 28

Figure 3-2. System configuration of a regeneratively ultrafast laser system based on CPA mechanism. ................................................................................................................................. 30

Figure 3-3. Nonlinear paramteric conversion mechanism in a $\chi^2$ crystal. .......................... 31

Figure 3-4. Schematics of (a) open-aperture Z-scan to characterize NLA, and beam divergence in closed-aperture Z-scan using a sample with positive $n_2$ (b) after and (c) before the focus to characterize NLR. .................................................................................................................. 34

Figure 3-5. (a) OA Z-scan measurements on a 0.35 mm thick GaAs bulk material and fit to characterize 2PA $\alpha \approx 6 \text{ cm/GW}$. (b) CA Z-scan measurements on a 1 mm thick fused silica to characterize electronic NLR and fit $n_2 \approx 0.25 \times 10^{-15} \text{ cm}^2/\text{W}$. ...................................................... 37

Figure 3-6. Schematics for DA Z-scan with the main sample (solution or thin-film sample) at arm A and the background sample (solvent or substrate) at arm B. ................................................................. 39

Figure 3-7. DA Z-scan measurements of an organic film (AJCPE04-TCF04) of thickness 1.29 $\mu$m deposited on a 1 mm thick quartz substrate at a wavelength of 1300 nm using multiple irradiances. (a) and (b) are OA and CA/OA signals, respectively. The effects of substrate and LEB are subtracted from the signal. (c) Sequential measurements and (d) simultaneous measurements (DA Z-scan) show the comparison of the subtraction of the substrate CA/OA Z-scan from the film plus substrate CA/OA Z-scan in the case of 54 nJ input. The solid line in (d) is the best fit as in the top curve of (b). ................................................................................................................................. 40

Figure 3-8. (a) Experimental schematics of BD setup; (b) spatial distrubution of the irradiance of the pump beam (red) and displaced smaller probe beam (blue) at the position of sample; (c)
vertically deflected probe beam on a segmented quad-cell position sensitive detector with and without deflection. .......................................................................................................................................................................................... 44

Figure 3-9. BD measurements, $\Delta E/E$ (%) versus delay, of a 1 mm thick fused silica for relative polarizations between pump and probe of parallel (black) and perpendicular (red). ............. 47

Figure 3-10. Experimental configuration used for (a) pump and (b) probe beams to control the energy or polarization. Half-wave plate (HWP), polarizer (POL), and quarter-wave plate (QWP). .......................................................................................................................................................................................... 48

Figure 3-11. BD measurements of CS$_2$ solvent for parallel (black), perpendicular (red), magic angle (blue), and circular polarization (magenta). The magenta data is for the case of circular pump beam and linear probe beam, which is similar for all polarization combination with at least one beam being circular............................................................................................................................................................................................................. 50

Figure 3-12. (a) Conventional BD, (b) newly developed cross-propagating BD. The interaction geometry of the two beams are shown for each case........................................................................................................... 52

Figure 3-13. Interaction geometry of the cross-propagation BD method. ......................... 53

Figure 3-14. Polarization combinations possible in the cross-propagating beam deflection experiment. The strong excitation beam propagates in the z-direction, with the weak probe propagating in the y-direction............................................................................................................................................................................................................. 54

Figure 3-15: Cross propagating beam deflection results in CCl$_4$ and SnCl$_4$, for various polarization combinations.................................................. 55

Figure 3-16. (a) Cross-propagating beam-deflection, (b) extended geometry employing additions to adjust the wavefront tilt separately for excitation and probe beams................................. 56

Figure 3-17. (a) The interaction geometry and effect of wavefront tilt in spatio-temporal interaction, the illustrated tilt is 45 degrees (b) experimental observation of the optimization of the
cross-correlation width while wavefronts are tilted for 20 degrees for each beam compared to when they are perpendicular to the direction of propagation. The cross-correlation changes from ~ 6 ps (FWHM) to ~ 850 fs (FWHM). Note that the measurements were done for 20 degrees tilt for each beam.

Figure 4-1. Dispersion of the third-order degenerate nonlinear absorption (2PA) and refraction (Kerr effect) in GaAs.

Figure 4-2. (a) OA and (b) CA/OA Z-scan measurements on GaAs conducted at 1300 nm. Courtesy of my colleague Dr. Natalia Munera.

Figure 4-3. (a) Autocorrelation measurements for characterizing the pulsewidth of the pump pulse, (b) knife-edge scan measurements to characterize the beam profile (horizontal and vertical cross-section) of the probe beam and measuring the beam size of the pump beam.

Figure 4-4. (a) BD measurements on a 1 mm thick fused silica and fit values of $n^2 = 0.25 \times 10^{-19}$ cm$^2$/W for parallel polarization and (b) BD measurements on 1 mm thick CS$_2$ liquid after subtracting the contribution of the cuvette for three different polarization combinations.

Figure 4-5. (a) Nonlinear transmission and (b) BD measurements performed on a piece of GaAs for $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm.

Figure 4-6. (a) Nonlinear transmission and (b) BD measurements performed on a piece of GaAs for $\lambda_e = 1350$ nm and $\lambda_p = 1250$ nm.

Figure 4-7. Comparison of the measured values of third-order bound-electronic NLA and NLR for GaAs with the calculated dispersion using two-parabolic band model.

Figure 4-8. (a) Nonlinear transmission and (b) BD measurements performed on a piece of Silicon for $\lambda_e = 1350$ nm and $\lambda_p = 1250$ nm. (c) Nonlinear transmission and (d) BD measurements performed on a piece of GaAs for $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm.
Figure 4-9. Degenerate OA Z-scan measurements to characterize 3PA spectra of GaAs............ 82
Figure 4-10. Measured data for 3PA in GaAs obtained from the 8-band model (black dotted line). Scaling the calculated spectrum by a factor of 4.2 (solid black curve), provides the best agreement with experimental data. Red circles are experimental data taken with ~150 fs pulses. Blue circles are data for the ~ 10 ps pulses. (b) The FCA cross section from the 2-parameter fit for 3PA and the FCA........................................................................................................................................... 83
Figure 4-11. The contributions of starting in the HH (red), LH (blue) and SO bands (green) to the total (black) 3PA in several zinc-blende semiconductors, ordered from small to large Δ/Eg to show the continuous evolution of the 3PA spectral shape, and the 3PA energy sum starts just below the onset of 3PA and goes to where 2PA turns on................................................................. 85
Figure 4-12. Dependence of \( \alpha_3, \text{norm} \) on the normalized three-photon energy sum and \( \Delta/Eg \). \( \alpha_3, \text{norm} \) is normalized to the value of 3PA when \( \Delta = 0 \). The range of \( \Delta/Eg \) is limited to \( \Delta/Eg < 0.3 \) to better represent the effects of the SO band on the 3PA spectra. .................................................. 87
Figure 5-1. (a) Real and imaginary part of the electric permittivity, and (b) real and imaginary part of the complex refractive index for a Lorentz response of a single electron versus frequency.... 94
Figure 5-2. ITO built-in library model used in commercial software for analyzing ellipsometry data.................................................................................................................................................. 98
Figure 5-3. Experimentally measured values of \( \psi \) using spectroscopic ellipsometry (solid lines) at three different angles of incident. Fit with Lorentz oscillator model in the transparent region and Drude model in the lossy region. Fits are courtesy of my colleague Sanaz Faryadras.......... 98
Figure 5-4. Re-fit of ellipsometry data considering a B-spline fitting model and including non-idealities of the experimental condition............................................................................................................. 99
Figure 5-5. Optical constants of a Drude material (ITO) with $\omega p = 2\pi \times 469 \times 1012$ rad/s, $\gamma = 0.0410 \omega p$, and $\varepsilon_\infty = 3.65$ measured by the spectroscopic ellipsometry technique. The real and imaginary part of permittivity versus frequency, the blue shading is the ENZ region............. 100

Figure 5-6. (a) real and imaginary part of the complex index, i.e. refractive index and extinction coefficient, the red shading is the NZI region. (b) Values of $Re \eta = n$ in the complex permittivity space for ITO with optical parameters mentioned in Figure 5-5. The color bar shows values of $n = Re \varepsilon < 1$.......................................................... 101

Figure 5-7. FOM for various phononic, plasmonic, and metalic materials. Material data from refs [240-242]. The magenta data (ITO) is from this work. .......................................................... 102

Figure 5-8. Simple band structure for metals where d-band overlabs the sp-band. (a) the interband direct optical transition, and (b) the intraband optical transition that includes exchange of momentum. ......................................................................................................... 104

Figure 5-9. (Right axis) BD measurements, $\Delta E/E$ (%), versus delay for glass substrate and extracted BD signals for the ITO film at both parallel and perpendicular relative polarizations. (Left axis) Nonlinear phase change directly calculated from our BD measurements. .......... 114

Figure 5-10. (a) OA, (b) CA, and (c) CA/OA signals obtained from Z-scan measurements on the ITO film.................................................................................................................................................. 116

Figure 5-11. Irradiance dependence of $\Delta n_{eff}$ and $\Delta \phi$ calculated from BD (red) and Z-scan (blue) measurements.............................................................................................................. 116

Figure 5-12. Charge separation and induced electric field within the medium due to the longitudinal oscillation of free-electrons excited by a TM polarized light at an obliquely incident field. ............................................................................................................................... 119
Figure 5-13. (black) Real and (red) imaginary part of the Drude permittivity for ~320 nm ITO versus wavelength. The blue curve is the associated electron energy loss function for this material fitted by a Lorentzian with a maximum at \( \omega_p (\lambda_{ENZ} \approx 1240 \text{ nm}) \) and width of 113 nm. ...

Figure 5-14. Three-layer structure comprising of air, lossy film, and the semi-infinite substrate.

Figure 5-15. Transmission, reflection, and absorption versus angle of incident and wavelength for TM polarization (top) and TE polarization (bottom). Absorbed power exhibits an enhanced feature at oblique incident angles close to the ENZ wavelength.

Figure 5-16. \( A(\lambda, \theta) \) for (a) 2x larger, (b) 2x smaller, or (c) 10x smaller values of \( \text{Im}(\varepsilon) \) for all wavelengths.

Figure 5-17. \( R \) (black), \( T \) (red), and \( A \) (blue) calculated for different incident angles for TM polarization while (a) interference effects are not included, \textit{i.e.} Fresnel coefficients for intensity and (b) interference effects are included, \textit{i.e.} Fresnel coefficients for field, as discussed in APPENDIX B.

Figure 5-18. BD measurements for \( \lambda_e = 1250 \text{ nm} \) and \( \lambda_p = 775 \text{ nm} \) on (a) ITO + substrate and (b) bare glass substrate for various input beam irradiances. Irradiance dependence of the measured \( \Delta E/E(\%) \) for (c) ITO at two different delays where signal is maximum (A) and at a delay where the contribution of substrate is zero (B), and (d) the bare substrate.

Figure 5-19. The fitting procedure for each mechanism contributing to the BD measurements. The red solid line in (a-c) is the contribution of the electro-phonon mechanisms, the black solid line in (b-c) is the contribution of the electron-electron mechanisms. The sum of the two effects is the blue solid line in (c-d). In (d) we plot the contribution of ITO (blue) and glass (red) and sum of these two effects (black).
Figure 5-20. BD measurements and fit for (a) various irradiances and (b,c) two irradiances where an unusual dip is observed at delay time of ~250 fs. ......................................................... 132
Figure 5-21. Extracted values of $\Delta n_{\text{eff}}$ plotted for fits to BD measurements .................. 133
Figure 5-22. (a-e) BD measurements on a thin piece of ITO deposited on a glass substrate at $\lambda_{\text{probe}} = 775 \text{ nm}$ and $\lambda_{\text{pump}} = 1150 - 1300 \text{ nm}$. (f) The spectral dependence of the extracted effective NLR compared to the absorption spectra calculated at normal incident. .................... 134
Figure 5-23. Absorption spectra for (a) TE and (b) TM polarized light at various incident angles. (c) The correlation between the calculated absorption spectra and extracted NLR values from the BD measurements at $\sim 6^\circ$ and $30^\circ$. ................................................................. 135
Figure 5-24. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1150 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row......................................................... 137
Figure 5-25. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1200 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row......................................................... 137
Figure 5-26. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1250 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row......................................................... 138
Figure 5-27. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1300 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row......................................................... 138
Figure 5-28. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1350 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row................................. 139

Figure 5-29. (a) BD measurements for $\lambda e = 775$ nm $\lambda p = 1250$ nm for different pulse irradiances, (b) cross-phase modulation measurements at different delays between excitation and probe at $I e = 521$ GW/cm$^2$. ................................................................. 143

Figure 5-30. (Left axis) BD measurements and fit for $\lambda p = 1250$ nm at 521 GW/cm$^2$. (Right axis) The black solid line is the calculation of $\Delta \lambda p$ based on $\partial \Delta \varphi / \partial t$, and cross-phase modulation measurements data................................................................. 145

Figure 5-31. BD measurements and the extracted $\Delta \varphi$ for each probe wavelengths from 950 – 1550 nm. ........................................................................................................ 146

Figure 5-32. Spectral dependence of the measured ratio of the nonlinear phase change in ITO to the glass substrate for an excitation at 775 nm and comparison to the spectral dependence of $\Delta n_{ef f}$ obtained from the Drude model. ................................................................. 147

Figure 5-33. Degenerate Z-scan measurements at $\lambda p = 775$ nm, at $I 0 = 368$ GW/cm$^2$. (a) OA, (b) CA, and (c) CA/OA.................................................................................................................. 149

Figure 5-34. (Left axis) Values of $\Delta \varphi$ extracted from BD measurements (blue data and solid line fits), (right axis) experimentally predicted (black solid line) and experimentally measured (data points) magnitude of $\Delta \lambda$ imposed upon the probe beam via the pump beam at $I e = 521$ GW/cm$^2$. ........................................................................................................ 150

Figure 5-35. spectral dependence of the optically induced red shift and blue shift of the probe pulse for a 775 nm excitation irradiance at 521 GW/cm$^2$ predicted from BD measurements............ 151
Figure 5-36. (a) Dispersion of $\kappa$, $n$, and $ng$ for ITO, (b) dispersion of $n$, $ng$, and $S = ng/n$ for ITO. ...................................................................................................................................................... 155

Figure 5-37. (a) Dispersion of ITO without NLO effects (black) and after an optically-induced change in the plasma frequency $\Delta \omega_p/\omega_p \approx 5\%$ occurs (red). (b) Calculated $\Delta k$ at each frequency while only considering horizontal transitions. .............................................................................................................................................. 156

Figure 5-38. Dispersion of ITO without NLO effects (black) and after an optically-induced change in the plasma frequency $\Delta \omega_p/\omega_p \approx 5\%$ occurs (red). The inset shows the contributions of $\Delta \omega$direct and projection of intraband frequency shift while considering the group velocity to the total shift in frequency. ...................................................................................................................................................... 158
LIST OF TABLES

Table 4-1. Literature values of 2PA and third-order Kerr nonlinearity of GaAs around 1550 nm
....................................................................................................................................................................... 59

Table 4-2. Literature values of 2PA and third-order Kerr nonlinearity of Silicon around 1550 nm
......................................................................................................................................................................... 60

Table 4-3. Summary of the results obtained from Z-scan measurements at 1300 nm on GaAs .. 68

Table 4-4. Summary of the results obtained from BD measurements on CS₂ for λe = 1600 nm and λp = 1500 nm, while considering 20% error. ................................................................................................................. 71

Table 4-5. Extracted values for instantaneous third-order NLA and NLR of GaAs, estimated error is 20%.............................................................................................................................................................................. 76

Table 4-6. Values of NLA and NLR of GaAs measured by Z-scan and BD at 1300 nm.............. 77

Table 4-7. Values of NLA and NLR of GaAs measured by Z-scan and BD at 1550 nm.............. 77

Table 4-8. Values of NLA and NLR of Silicon measured by Z-scan and BD at 1300 nm. ........ 79

Table 4-9. Values of NLA and NLR of Silicon measured by Z-scan and BD at 1550 nm ........ 80

Table 4-10. Band-structure parameters for semiconductors in their zinc-blende form at 300 K (except ZnO, at 4.2 K) used for modeling of 3PA. The first three rows indicate materials where the SO band energy is relatively large (Δ/Eg > 1) and the last five rows indicate materials with a the SO band energy is relatively small (ΔEg ≪ 1). .................................................................................................. 84
### LIST OF ACRONYMS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronyms/Units</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PA</td>
<td>One-photon Absorption</td>
</tr>
<tr>
<td>2PA</td>
<td>Two-photon Absorption</td>
</tr>
<tr>
<td>3PA</td>
<td>Three-photon Absorption</td>
</tr>
<tr>
<td>3D</td>
<td>Three Dimensional</td>
</tr>
<tr>
<td>AOS</td>
<td>All Optical Switching</td>
</tr>
<tr>
<td>AZO</td>
<td>Aluminum-doped Zinc Oxide</td>
</tr>
<tr>
<td>BBO</td>
<td>Barium Borate</td>
</tr>
<tr>
<td>BD</td>
<td>Beam Deflection</td>
</tr>
<tr>
<td>CCl₄</td>
<td>Carbon Tetrachloride</td>
</tr>
<tr>
<td>CPA</td>
<td>Chirped Pulse Amplified</td>
</tr>
<tr>
<td>CS₂</td>
<td>Carbon Disulfide</td>
</tr>
<tr>
<td>DA</td>
<td>Dual-Arm</td>
</tr>
<tr>
<td>DFWM</td>
<td>Degenerate Four Wave Mixing</td>
</tr>
<tr>
<td>ENZ</td>
<td>Epsilon-Near-Zero</td>
</tr>
<tr>
<td>ESA</td>
<td>Excited-State Absorption</td>
</tr>
<tr>
<td>ESR</td>
<td>Excited-State Refraction</td>
</tr>
<tr>
<td>FOM</td>
<td>Figure Of Merit</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width Half Maximum</td>
</tr>
<tr>
<td>GVD</td>
<td>Group Velocity Dispersion</td>
</tr>
<tr>
<td>GVM</td>
<td>Group Velocity Mismatch</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HW1/eM</td>
<td>Half Width 1/e Maximum</td>
</tr>
<tr>
<td>ITO</td>
<td>Indium Tin Oxide</td>
</tr>
<tr>
<td>KK</td>
<td>Kramers-Kronig</td>
</tr>
<tr>
<td>ND</td>
<td>Nondegenerate</td>
</tr>
<tr>
<td>NLA</td>
<td>Nonlinear Absorption</td>
</tr>
<tr>
<td>NLO</td>
<td>Nonlinear Optical</td>
</tr>
<tr>
<td>NLR</td>
<td>Nonlinear Refraction</td>
</tr>
<tr>
<td>NZI</td>
<td>Near-Zero Index</td>
</tr>
<tr>
<td>NOPA</td>
<td>Non-collinear Optical Parametric Amplification</td>
</tr>
<tr>
<td>OA&amp;CA</td>
<td>Open &amp; Closed Aperture</td>
</tr>
<tr>
<td>OHD</td>
<td>Optical Heterodyne Detection</td>
</tr>
<tr>
<td>OKE</td>
<td>Optical Kerr Effect</td>
</tr>
<tr>
<td>OPA/G</td>
<td>Optical Parametric Amplifier / Generator</td>
</tr>
<tr>
<td>SFG</td>
<td>Sum Frequency Generation</td>
</tr>
<tr>
<td>SHG</td>
<td>Second Harmonic Generation</td>
</tr>
<tr>
<td>SVEP</td>
<td>Slowly Varying Envelope Approximation</td>
</tr>
<tr>
<td>THG</td>
<td>Third Harmonic Generation</td>
</tr>
<tr>
<td>WLC</td>
<td>White Light Continuum</td>
</tr>
<tr>
<td>ZnO</td>
<td>Zinc Oxide</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Nonlinear phenomena were discovered and have been exploited for many years in several research areas. In the field of electronics, nonlinear effects such as rectification and harmonic generation are essential phenomena that have been utilized in several applications. In the field of acoustics, the nonlinear response results in frequency mixing that may introduce undesirable distortion in speakers and amplifiers. In optics, the NLO effects are evidenced by changes in the absorption coefficient, refractive index, and light scattering properties of materials as the intensity of the incident light is increased. In the past six decades, after the invention of laser in 1960, there have been a considerable amount of interest in employing NLO effects in optical devices and photonic applications. The need for design and characterization of NLO materials and more efficient interaction geometries with large and fast nonlinear effects has attracted researchers and scientists to develop novel materials and study light-matter interaction schemes to improve NLO effects. Therefore, accurate modeling and characterization and reporting trustable values of nonlinear parameters are essential for NLO applications.

There are different NLO effects that can be of interest for device applications, but we primarily focus on those that influence the intensity and phase of the incident beam. The two principle nonlinear interactions influencing the intensity and phase are nonlinear refraction (NLR) and nonlinear absorption (NLA). This thesis covers two sections of fundamental study on NLA and NLR of semiconductors and transparent conducting oxides within the realm of nonlinear optical spectroscopy. The first section is focused on the development of a novel nonlinear optical spectroscopic technique, beam deflection, and its application to study the instantaneous third-order and non-
instantaneous fifth-order nonlinear dynamics of a direct-gap semiconductor, GaAs, and also Silicon, which is an indirect gap material. Additionally, measurements of the spectral dependence of instantaneous fifth-order NLA, three-photon absorption (3PA), of semiconductors is presented. The second part of this dissertation focuses on fundamental characterization of nonlinear refractive mechanisms of ITO at the frequency range where the real part of permittivity crosses zero.

1.1. Bound-electronic and carrier-induced nonlinearities in semiconductors

Semiconductor materials play a major role in nonlinear optics due to their large nonlinear optical responses and because of their applications in many areas of optics and their compatibility with complementary metal–oxide–semiconductor (CMOS) technology. This class of materials have some of the fastest nonlinearities for sub-bandgap photons. For near-gap excitations, nonlinearities are usually accompanied with linear absorption that is undesirable in some cases. In the transparency region of semiconductors, where the photon energy is far enough below the bandgap, bound electronic nonlinearities dominate the nonlinear response. Nonlinear refraction associated with the anharmonic response of bound valence electrons is known as the bound-electronic Kerr effect and is described by a change of refractive index proportional to the intensity of light. This type of nonlinearity is a result of virtual intermediate transitions as opposed to real resonant transitions. In the transparency region, the response time is very fast and can be considered instantaneous. As a consequence of causality in any linear system, a dispersion relation governs the real and imaginary parts of the frequency-domain complex response function via Hilbert transform pairs. In optical systems, this principle results in the Kramers-Kronig transformation, which relates the real and imagery parts of the linear susceptibility. For photon energies above the bandgap or when free-carriers are generated via absorptive processes, the nonlinearity will be dominated by the response of free-carriers, which are much larger but significantly slower than bound-electronic NLO effects.
This difference in the time-dynamics of nonlinear mechanisms becomes useful in distinguishing the two processes with time-resolved spectroscopic techniques such as transient absorption spectroscopy and beam-deflection. Additionally, for longer wavelengths, higher order bound-electronic effects such as 3PA can occur. In this case, the transmitted power will drop quadratically versus the input intensity.

To precisely characterize the strength of these nonlinearities and their physical origin, we use BD method to characterize magnitude, sign, and time dynamics of nonlinearities in telecommunication spectral range. Our measurements, together with Z-scan measurements provide a complete reference for third- and fifth-order nonlinearities at this regime. We also perform Z-scan to rigorously characterize the spectral dependence of the 3PA process in GaAs. We employ a Kane 8-band model to describe different transition pathways in the 3PA process. This model includes the quantum interference between different pathways originating from heavy-hole, light-hole, and split-off bands located in the valence band of direct-gap semiconductors. We used this model to predict the spectral dependence of 3PA in 15 different semiconductors in their zinc-blende form, which emphasizes the importance of spin-orbit interaction in the dispersion of the 3PA process.

1.2. Enhanced Optical Nonlinearities at Epsilon-Near-Zero

In this dissertation, we also study highly doped semiconductors such as transparent conducting oxides (TCO) as rewarding candidate materials for ultrafast nonlinear optical applications, due to their large and fast NLO response in the telecommunication band, low optical loss comparing to the conventional plasmonic materials such as gold (Au) and silver (Ag), tunable optical properties, and ease of fabrication. These materials are a unique class of wide-bandgap degenerate semiconductors which can support extremely high doping levels to demonstrate a metallic response in the near-infrared region. Of particular interest is their remarkable NLO properties at
wavelengths where the real part of the permittivity is near zero, so-called “Epsilon-Near-Zero” (ENZ) such as enhanced second and third harmonic generation [1, 2], enhancement in nonlinear refraction [3-9], and etc. These nonlinear optical processes, where the ultrafast index change can be larger than the linear index is a revolutionary development in nonlinear optics allowing a myriad of potential applications, such as rapid all-optical beam steering and switching, spatial mode conversion, and nonlinear metasurfaces for wavefront modulation. Further research in these materials can open new avenues to engineer nonlinearities on the nanoscale and potential device applications.

The intent of CHAPTER 5 of this dissertation is reviewing the basic mechanisms of the nonlinear optical properties of highly doped semiconductors such TCOs by focusing on nonlinear refraction phenomena. The work of particular interest here is to investigate novel materials with large nonlinear refraction for potential applications in all-optical signal processing and nanophotonics. We begin with a theoretical review of optical nonlinearities of interest, particularly in the case of doped semiconductor structures and contributing mechanisms to carrier dependent nonlinearities arising from interband and intraband transitions. We present a rigorous experimental study on NLO effects of ITO at ENZ condition. Our experimental work comprises of degenerate measurements using Z-scan and BD, and rigorous nondegenerate measurements for two combinations of pump and probe wavelengths. We first study how exciting carriers in vicinity of the ENZ region can improve NLO effects, and then investigate the optically-induced changes in nonlinear phase and wavelength of probe. BD becomes useful, particularly because of its ability to temporally resolve optical nonlinearities, to directly characterize the nonlinear phase change, and also to measure the anisotropy of the nonlinear phase change. Our experimental results from these two techniques show that the optically induced index changes in TCOs can be both very large and fast. Our
results also show that carrier redistribution effects dominate the nonlinear refraction, and we find a strong wavelength dependence of nonlinearities around the ENZ point.

Here, we also present a perspective to describe the NLO enhancement phenomena observed in ITO based on the dispersive optical properties of the material. This description is based on the reduced group velocity of the light pulses propagating in ITO, which facilitates a more efficient light-matter interaction. This perspective can be applied to other TCOs and may eventually help in the design and modeling of spectral responses of nonlinear plasmonic devices.

1.3. Dissertation Outline

This dissertation is organized as follows: CHAPTER 2: introduces the fundamentals of light-induced changes in optical properties of semiconductors and metals under short optical pulse illumination. We will discuss both instantaneous and non-instantaneous NLO effects and their physical origin. We will describe how wave propagation will be modified with any of these effects by understanding the effect of each mechanism on the irradiance and phase of the transmitted wave through the nonlinear medium. CHAPTER 3: presents a detailed introduction to the experimental capabilities developed before or during the work of this dissertation. We present the basic principles of two nonlinear spectroscopy techniques, namely Z-scan and BD method. We also describe advantageous of each method within the concepts of interest in this work. We also present the progress towards modifying the BD method to a new geometry, in which measuring off-diagonal susceptibility matrix elements becomes possible. CHAPTER 4: Reviews the fundamental principles of nonlinear optical properties of semiconductors, particularly instantaneous third-order nonlinear mechanisms such as 2PA and Kerr effect, non-instantaneous fifth-order nonlinear mechanisms such as 2PA-induced FCA and FCR. We show our capabilities in separately measuring each effect by using our BD methodology. We will also present our developed theory for calculating
the 3PA spectra by utilizing the Kane 8-band model, and our measurement of 3PA for GaAs. This theory will be then utilized to predict the 3PA spectra of 15 different semiconductors and provide insight on the effect of spin-orbit interaction. CHAPTER 5: Provides the fundamental principles of the optical nonlinearities in transparent conducting oxides, particularly in proximity to the ENZ condition, where nonlinear effects are significantly enhanced. We present our degenerate and nondegenerate characterizations, where they exhibit effective index changes close to unity. This includes transient beam-deflection (BD) spectroscopy and single beam Z-scan measurements. We will also discuss a perspective based on slow-light effects that greatly contributes to this enhanced optical nonlinear mechanism. CHAPTER 6: Summarizes the research performed within this dissertation and the conclusions of our studies.
CHAPTER 2: NONLINEAR ABSORPTION AND REFRACTION IN SOLIDS

The interaction of light with itself is only possible in the realm of nonlinear optics (NLO), in which a huge number of photons interact with NLO materials. The direct consequence of NLO interactions is induced changes in propagation of optical beams. A group of new phenomena arises from NLO interactions such as manipulating the transmission, phase, or polarization of an optical field (*i.e.* electric or magnetic field of an optical electromagnetic beam), and generating new frequencies. In general, this modification can be induced on the field by itself (self-action or single beam effects) or can be imposed on the field by means of another optical beam (cross-action or double beam effects). These phenomena are typically used in NLO devices where the output optical beam is controlled via modulation of the input beam realizing an all-optical signal processor or to synthesize new frequencies and to manipulate them on demand [10-13]. Nonlinear wave propagation models are used to accurately model these phenomena and to describe NLO interactions between light and matter. In this chapter, we primarily focus on exploring NLO phenomena modifying the intensity and phase of the transmitted beam via NLO interactions.

In a typical NLO device, nonlinearities are considered a perturbation to linear optical processes and approximations can be easily made to describe their behavior [14, 15]. Prominent examples are third-order and fifth-order nonlinearities of bound electrons that are nearly instantaneously influencing the intensity and phase of the optical field. We use a power series expansion of the induced polarization to model such nonlinearities and describe their physical origin. Depending on the physical mechanism and state of the material, non-instantaneous effects such as long-lived free-carrier induced nonlinearities also occur in NLO systems. The physical origin of NLO effects
from free-carriers is different than that of bound electrons. We also discuss the details of these phenomena in this chapter.

In general, nonlinear absorption (NLA) is referred to those nonlinearities that are responsible for well-known optical effects such as two-photon absorption (2PA), three-photon absorption (3PA), free-carrier absorption (FCA) in solids, saturable absorption, and etc. Similarly, nonlinear refraction (NLR) gives rise to optical phenomena such as self-focusing (defocusing), self-phase modulation, cross-phase modulation, free-carrier refraction (FCR) in semiconductors, and excited-state nonlinear refraction in organics.

In some cases, NLO processes can significantly modify material properties, hence perturbative analyses become inaccurate. In such circumstances, other approaches need to be considered and the notion of defining NLO coefficients as in their conventional meaning (NLA and NLR) are no longer entirely correct. We will discuss this case in more detail in CHAPTER 5.

2.1. Nonlinear Wave Equation

The evolution of the electric field in time and space coordinates is described via the wave equation that can be directly derived from Maxwell’s equation [16]. We begin from Maxwell’s equations in a source free, non-magnetic, and homogeneous medium. By taking the curl of the first of Maxwell’s equation we get

$$\nabla \times \nabla \times \vec{E}(t) = -\mu_0 \frac{\partial}{\partial t} \{\nabla \times \vec{H}(t)\},$$

where $\vec{E}(t)$ and $\vec{H}(t)$ are the real, time varying electric and magnetic fields, respectively, and $\mu_0$ is the permeability of free space. We then substitute the second of Maxwell’s equation into the right hand side of Eq. (2-1) and obtain the wave equation governing the time-dependent electric field
\[ \nabla \times \nabla \times \vec{E}(t) = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}(t)}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}(t)}{\partial t^2} \]  
(2-2)

where \( \vec{P}(t) \) is the induced polarization (C/m\(^2\)) by the electric field and \( \varepsilon_0 \) is the permittivity of free space. By expanding the left hand side using the identity [17]

\[ \nabla \times \nabla \times \vec{A} = \nabla (\nabla \cdot \vec{A}) - \nabla^2 \vec{A}, \]  
(2-3)

we simplify Eq. (2-2) to

\[ \nabla \left( \nabla \cdot \vec{E}(t) \right) - \nabla^2 \vec{E}(t) = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}(t)}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}(t)}{\partial t^2} \]  
(2-4)

By using the third of Maxwell’s equations we examine the divergence term as

\[ \frac{1}{\varepsilon_0} \nabla \left( \nabla \cdot \vec{P}(t) \right) - \nabla^2 \vec{E}(t) = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}(t)}{\partial t^2} - \mu_0 \frac{\partial^2 \vec{P}(t)}{\partial t^2} \]  
(2-5)

We can then neglect the first term, \( \left| \frac{1}{\varepsilon_0} \nabla \left( \nabla \cdot \vec{P}(t) \right) \right| \ll \left| \nabla^2 \vec{E}(t) \right| \), by assuming that the material polarization changes more slowly in space than the oscillation of electric field, and thus write

\[ \nabla^2 \vec{E}(t) - \frac{1}{c^2} \frac{\partial^2 \vec{E}(t)}{\partial t^2} = \mu_0 \frac{\partial^2 \vec{P}(t)}{\partial t^2}, \]  
(2-6)

where \( c = 1/\sqrt{\mu_0 \varepsilon_0} \) is the speed of light in free space. Eq. (2-7) relates the evolution of the electric field in time and space to the induced polarization of the material via both linear and nonlinear optical interactions. The right hand side of Eq. (2-6) comprises of two parts as

\[ \vec{P}(t) = \vec{P}_L(t) + \vec{P}_{NL}(t), \]  
(2-7)

where \( \vec{P}_L \) and \( \vec{P}_{NL} \) are the linear and nonlinear polarizations, respectively. The first term on the right hand side of Eq. (2-7) is responsible for all the phenomena in linear optics such as linear absorption (one-photon absorption – 1PA) and index of refraction. The second term gives rise to nonlinear optical phenomena that we will discuss in detail.
In a homogeneous, isotropic, and linear optical medium, the induced linear polarization on bound electrons can be described by

\[
\mathbf{P}(t) = \mathbf{P}_L(t) = \varepsilon_0 \int_{-\infty}^{\infty} \mathbf{\chi}^{(1)}(t - t_1) \cdot \mathbf{E}(t_1) dt_1,
\]

(2-8)

where \( \mathbf{\chi}^{(1)}(t) \) is the first order material susceptibility tensor. Eq. (2-8) can also be written in frequency domain by taking Fourier transform of two sides of the equation:

\[
\mathbf{P}(\omega) = \varepsilon_0 \mathbf{\chi}^{(1)}(\omega) \cdot \mathbf{E}(\omega),
\]

(2-9)

where \( \mathbf{E}(\omega) \) and \( \mathbf{P}(\omega) \) are Fourier transforms of the real, time varying electric field and polarization, respectively, and \( \omega \) is the optical frequency.

When an external electric field is applied to a material, bound charges to the molecule that are not free to move will experience a displacement, which is typically described by the displacement field, \( \mathbf{D}(\omega) \). In a linear material with a bound charge density of \( \rho_b \) and free charge density of \( \rho_f \) the total charge density can be written as \( \rho = \rho_f + \rho_b \) where from Maxwell’s equations \( \rho = \varepsilon_0 \nabla \cdot \mathbf{E}, \rho_b = -\nabla \cdot \mathbf{P}, \) and \( \rho_f = \nabla \cdot \mathbf{D} \). This will result in \( \mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \) and subsequently from constitutive relation we can derive \( \mathbf{\varepsilon}(\omega) = \varepsilon_0 \left( 1 + \mathbf{\chi}^{(1)}(\omega) \right) \) where \( \mathbf{\varepsilon} \) is the electric permittivity tensor.

As the electric field increases, the polarization described by Eq. (2-8) cannot indefinitely increase, and thus it cannot fully explain the total induced polarization. This is because the electric field starts to become same magnitude as the field between the particles forming the material and subsequently the material will be altered [14, 15, 17]. At this point the material is no longer linear in nature but starts to become weakly nonlinear. Treating nonlinearities as a perturbation, \( \mathbf{P}_{NL}(t) \) can be used to describe the modifications to the polarization.
2.1.1. Solutions of nonlinear wave equation

Nonlinearly induced modifications in the optical media directly influences the wave propagation. To understand how this nonlinear behavior influences the light-matter interaction, we aim to solve Eq. (2-6) with some approximations. We consider the solution of Eq. (2-6) is superposition of monochromatic plane waves, and thus we expand the electric field and polarization as

\[
\mathbf{E}(t) = \sum_n \mathbf{E}_n(t) = \frac{1}{2} \sum_n \mathbf{E}_i(\omega_n) e^{i(k_n \cdot \mathbf{r} - \omega_n t)} + c.c.,
\]

\[
(2-10)
\]

\[
\mathbf{P}(t) = \sum_n \mathbf{P}_n(t) = \frac{1}{2} \sum_n \mathbf{P}_i(\omega_n) e^{i(k_n \cdot \mathbf{r} - \omega_n t)} + c.c.,
\]

\[
(2-11)
\]

where \( \mathbf{E}_i \) and \( \mathbf{P}_i \) are the complex electric field and polarization amplitudes along the direction \( i \), respectively, \( k_n \) is the wavevector, \( \mathbf{r} \) is the unit direction vector. The complex linear refractive index can be defined for isotropic media as

\[
\eta(\omega) = \sqrt{\varepsilon(\omega)/\varepsilon_0} = \sqrt{1 + \chi^{(1)}(\omega)} = n_0(\omega) + i\kappa_0(\omega),
\]

\[
(2-12)
\]

where \( n_0 \) and \( \kappa_0 \) are the linear index of refraction and extinction coefficients, respectively. From Eq. (2-12) we get \( |k_n| = k_n = \text{Re}[\eta(\omega)] \omega_n / c = n_{0,i}(\omega_n) \omega_n / c \). Note that \( k_n \) a real wavevector that only determines the phase and the decay in amplitude characterized by \( \kappa_0 \) is incorporated in the electric field amplitude.

Now we assume a slowly-varying envelope approximation (SVEA) for the electric field and polarization. This allows us to neglect the time and space second derivatives in Eq.(2-6). Considering the plane waves to be propagating in the \( z \) direction (\( \mathbf{r} = \mathbf{z} \)), each term in Eq.(2-6) associated with a given \( n \) yields
\[
\left( i2k \frac{\partial E}{\partial z} + \frac{i2\omega\eta^2}{c^2} \frac{\partial E}{\partial t} - \left( k^2 - k_0^2(1 + \chi^{(1)}) \right) E \right) e^{ikz} \\
= -\mu_0\omega \left( i2 \frac{\partial P_{NL}}{\partial t} + \omega P_{NL} \right) e^{ik_{NL}z},
\]

where \( k = n_0k_0 = n_0\omega/c \) and is the real wavevector, \( k_0 \) is the wavevector in free space, and \( P_{NL} \) and \( k_{NL} \) are the nonlinear polarization at frequency \( \omega \) and the associated wavevector, respectively. We can also assume \( 2|\partial P_{NL}/\partial t| \ll \omega P_{NL} \) since \( P_{NL} \) is slowly varying in time. Also, note that \( \eta^2 = 1 + \chi^{(1)} \) and by substituting \( \eta = n_0 + i\kappa_0 \) and into Eq. (2-13) we get

\[
i2k \left( \frac{\partial E}{\partial z} + \frac{n_0^2 - \kappa_0^2}{cn_0} \frac{\partial E}{\partial t} + \frac{2\kappa_0^2}{c} \frac{\partial E}{\partial t} + i \frac{k_0\kappa_0^2}{2n_0} E + k_0\kappa_0 E \right) e^{ik_0n_0z} \\
= -\mu_0\omega \left( i2 \frac{\partial P_{NL}}{\partial t} + \omega P_{NL} \right) e^{ik_{NL}z}
\]

(2-14)

Typically, it is favorable to use materials with small linear loss in NLO applications, and thus in almost all situations we assume \( \kappa_0^2 \ll n_0^2, 2n_0\kappa_0 \). In this case Eq. (2-14) simplifies to

\[
i2k \left( \frac{\partial E}{\partial z} + \frac{n_0}{c} \frac{\partial E}{\partial t} + \frac{\alpha_0}{2} E \right) e^{ik_0n_0z} = -\mu_0\omega \left( i2 \frac{\partial P_{NL}}{\partial t} + \omega P_{NL} \right) e^{ik_{NL}z},
\]

(2-15)

where \( \alpha_0 = 2k_0\kappa_0 \) is the linear 1PA absorption coefficient. If the material is completely transparent we can also ignore \( \alpha_0 \). Also, by using a transformation to the reference frame of the moving pulse (\( \tau = t - nz/c \)) we get

\[
\frac{\partial E}{\partial z} = i \frac{\omega}{2nc\epsilon_0} P_{NL} e^{i(k_{NL}-k_0n_0)z}
\]

(2-16)

In lossy materials such as metals, one needs to consider Eq. (2-14) without ignoring \( \kappa_0 \) to fully describe the wave propagation; however, Eq. (2-16) can be safely used in semiconductors for photon energies below the bandgap, \( E_g \), where bound-electronic nonlinearities are dominant.
2.2. Nonlinear Absorption and Refraction of Bound Electrons in Semiconductors

Bound-electronic nonlinearities are the smallest nonlinear effect occurring in semiconductor materials. These nonlinearities are dominant in the transparency regime of semiconductors where free-carrier nonlinearities are not present. When an intense optical beam with photon energies less than the bandgap, $\hbar \omega_p < E_g = \hbar \omega_g$, is illuminated on a semiconductor, the bound valence electrons motion becomes anharmonic. Because of the low mass of electrons, they will respond to this excitation at optical frequencies ($> 10^{15}$ Hz), and thus these nonlinearities are usually referred to as instantaneous nonlinearities. The fast response time ($< 10^{-15}$ s) of these nonlinear effects will force them to show dispersive properties. We discuss the dispersion of bound-electronic nonlinearities in more detail in CHAPTER 4. Here, we only discuss how such nonlinearities will influence the nonlinear polarization and subsequently the nonlinear wave equation.

2.2.1. Power series expansion of nonlinear polarization

Of particular interest in semiconductors are nonlinear refraction (NLR) and nonlinear absorption (NLA) that directly influence the intensity and phase of the field interacting with the material. We study two cases where the beam acts upon itself (degenerate nonlinearities) or a second incident beam (non-degenerate nonlinearities) with a different frequency than the incident beam. Let us consider that $\omega_a$ is the frequency of the beam experiencing the nonlinearities upon itself or from another beam with the frequency $\omega_b$. We now use power series expansion and derive NLA and NLR for two cases of third-order and fifth-order nonlinearities.

The power series expansion approach of the polarization field in terms of the incident field fully characterizes the instantaneous response of electronic nonlinearities. For weak nonlinearities
that are considered a perturbation, the polarization can be expanded via the power series of the applied electric field as

$$\tilde{P}_{NL}(t) = \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(2)}(t - t_1, t - t_2) : \tilde{E}(t_1)\tilde{E}(t_2) dt_1 dt_2$$

$$+ \varepsilon_0 \iiint_{-\infty}^{\infty} \chi^{(3)}(t - t_1, t - t_2, t - t_3) : \tilde{E}(t_1)\tilde{E}(t_2)\tilde{E}(t_3) dt_1 dt_2 dt_3 + \cdots,$$

(2-17)

where $\chi^{(n)}(t - t_1, t - t_2, \ldots)$ is the $n^{th}$ order nonlinear susceptibility tensor, and : denotes the tensor product. Each order of nonlinear susceptibility is responsible for various NLO effects and contributes to the nonlinear polarization differently. For instance, second-order susceptibility results in generation of new frequencies rising to phenomena such as second-harmonic generation (SHG), sum-frequency generation (SFG), difference frequency generation (DFG), and optical rectification (OR) processes. Third-order susceptibility can be used to describe third-harmonic generation (THG), four-wave mixing (FWM), two-photon absorption (2PA), and third-order nonlinear refraction. Fifth-order susceptibility gives rise to higher order NLA such as three-photon absorption (3PA) and higher order NLR. Eq. (2-7) can also be written alternatively in frequency domain by taking a Fourier transform of Eq. (2-7) and Eq. (2-17). This representation is more favorable for bound-electronic nonlinearities

$$\mathbf{P}(\omega) = \varepsilon_0 \chi^{(1)}(\omega) \cdot \mathbf{E}(\omega) +$$

(2-18)
\[ \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(2)}(\omega; \omega_1, \omega_2) : \mathbf{E}(\omega_1) \mathbf{E}(\omega_2) \delta(\omega - \omega_1 - \omega_2) d\omega_1 d\omega_2 \]

\[ + \varepsilon_0 \int_{-\infty}^{\infty} \chi^{(3)}(\omega; \omega_1, \omega_2, \omega_3) : \mathbf{E}(\omega_1) \mathbf{E}(\omega_2) \mathbf{E}(\omega_3) (\omega - \omega_1 - \omega_2 - \omega_3) d\omega_1 d\omega_2 d\omega_3 \]

\[ + \cdots, \]

### 2.2.2. Third-order bound-electronic nonlinearities

We can write the magnitude of the third-order nonlinear polarization as

\[ P_a^{(3)}(\omega_a) = \frac{3\varepsilon_0}{4} \chi_{D,el}^{(3)}(\omega_a; -\omega_a, \omega_a) |\mathbf{E}_a(\omega_a)|^2 \mathbf{E}_a(\omega_a) + \]

\[ \frac{6\varepsilon_0}{4} \chi_{ND,el}^{(3)}(\omega_a; -\omega_b, \omega_a) |\mathbf{E}_b(\omega_b)|^2 \mathbf{E}_a(\omega_a) \]

(2-19)

where \( \chi_{D,el}^{(3)} \) and \( \chi_{ND,el}^{(3)} \) are the degenerate and non-degenerate third-order nonlinear susceptibilities, respectively. The constant coefficients multiplied to the right hand side of the above equations and twice the effect are from the permutation symmetry of the nonlinear susceptibility. Note that for single beam experiments such as Z-scan, the first term needs to be used; however, the second term is used for double beam experiments such as pump-probe where two beams have different frequencies.

By applying Eq. (2-19) to Eq. (2-16) and separating the field into amplitude and phase

\[ E_i(z) = \sqrt{\frac{2}{\varepsilon_0 n_{0,i}} I_i(z)} e^{i\Delta \varphi_i(z)}, \]

where \( I_i(z) \) is the irradiance, \( \Delta \varphi_i(z) \) is the induced nonlinear phase change, and \( n_{0,i} \) is the index of refraction at frequency \( \omega_i \), we get two equations for \( I_a(z) \) and \( \Delta \varphi_a(z) \) as
\[
\frac{\partial}{\partial z} I_a(z) = -\frac{\omega_a}{c n_{0,a}} \text{Im}\{\chi^{(1)}\} I_a(z) - \frac{3 \omega_a \mu_0}{2 n_{0,a}^2} \text{Im}\{\chi^{(3)}_{D,el}\} I_a^2(z)
\]

(2-20)

\[-3 \frac{\omega_a \mu_0}{n_{0,a} n_{0,b}} \text{Im}\{\chi^{(3)}_{ND,el}\} I_b(z) I_a(z),\]

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z) = k_{0,a} n_{0,a} + k_{0,a} n_{0,a}^2 I_a(z) + 2 k_{0,a} n_{0,a} I_b(z) + 2 k_{0,a} n_{0,a}^2 I_b(z),
\]

(2-21)

The first term on the right hand side of Eq. (2-20) is related to the linear absorption (1PA), the second term gives us the degenerate single-beam 2PA (D-2PA), and the third term is related to the non-degenerate double-beam 2PA (ND-2PA). Similarly, for Eq. (2-21), the first term gives the phase change due to linear refraction, the second term characterizes the degenerate third-order bound-electronic NLR, and the last term characterizes the non-degenerate third-order bound-electronic NLR. These equations can be also described as

\[
\frac{\partial}{\partial z} I_a(z) = -\alpha_{0,a} I_a(z) - \alpha_{2} D I_a^2(z) - 2 \alpha_{2} N I_b(z) I_a(z),
\]

(2-22)

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z) = k_{0,a} n_{0,a} + k_{0,a} n_{0,a}^2 I_a(z) + 2 k_{0,a} n_{0,a} I_b(z),
\]

(2-23)

where \(\alpha_{0,a}\) is the 1PA coefficient, \(\alpha_{2} D\) is the D-2PA coefficient, and \(\alpha_{2} N D\) is the ND-2PA coefficient.

Similarly, \(n_{0,a} D\) is the degenerate NLR coefficient, and \(n_{0,a} N D\) is the non-degenerate NLR coefficient.

We can think of the last two terms of the right hand side of Eq. (2-20) and (2-21) in terms of induced changes in refractive index and absorption coefficient and summarize them as

\[
\frac{\partial}{\partial z} I_a(z) = -(\alpha_{0,a} + \Delta \alpha^{(3)}_{a,el}(z)) I_a(z),
\]

(2-24)

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z) = k_{0,a} (n_{0,a} + \Delta n^{(3)}_{a,el}(z)),
\]

(2-25)
where $\Delta \alpha_{a,el}^{(3)}(z)$ and $\Delta n_{a,el}^{(3)}(z)$ are the magnitude of third-order NLA and NLR, respectively as described via

$$\Delta \alpha_{a,el}^{(3)}(z) = \alpha_2^D I_a(z) + 2 \alpha_2^{ND} I_b(z),$$

(2-26)

$$\Delta n_{a,el}^{(3)}(z) = n_2^D I_a(z) + 2 k_{0,a} n_2^{ND} I_b(z)$$

(2-27)

These two equations show that third-order nonlinearities are directly proportional to the irradiance. Also, at any given $z$ along the propagation direction, the magnitude of NLA or NLR depends upon the irradiance of the excitation beam at that exact position. This indicates that these effects are instantaneous meaning that there are no effects accumulated over time. These effects are the key mechanism of many nonlinear phenomena such as self-focusing [18-22], soliton propagation [23-28], phase conjugation [29-32], two-photon absorption [23, 33-39], and stimulated Raman scattering [23, 40, 41].

### 2.2.3. Fifth-order bound-electronic nonlinearities

We extend our derivation to a higher order term in nonlinear polarization that gives rise NLA and NLR that are quadratically scaled with the excitation irradiance. The magnitude of these effect is significantly smaller than third-order effects. However, there are numerous cases where fifth-order nonlinearities compete with third-order effects as we will show in following chapters. For instance, three-photon absorption (3PA) has shown both fundamental research importance and technological implications in many nonlinear optical (NLO) applications [42-46]. Some of them are stimulated emission and frequency upconversion fluorescence emission in organic optoelectronics [43-45], high-harmonic generation efficiency in semiconductors [47-49], and limiting two-photon lasing in direct gap semiconductors [50, 51]. There has also been a major interest in nonlinear pulse propagation in optical fibers over the past 20 year, and the main activity was centered around
nonlinear confinement of the pulse to avoid the distortions caused by dispersion and subsequently soliton propagation [52-54]. Higher order nonlinear mechanisms such as fifth-order nonlinear refraction become substantially important in such media [55-58]. Additionally, higher order Kerr nonlinearities may influence the propagation of intense ultrashort pulses in filamentation applications [59-63]. We overview the nonlinear wave propagation in such media.

To analyze how fifth-order nonlinear susceptibility influences the propagation of light we start with the fifth-order nonlinear polarization as

\[
\mathcal{P}_a^{(5)}(\omega_a) = \frac{5\varepsilon_0}{18} \chi_{D,el}^{(5)}(\omega_a; \omega_a, -\omega_a, -\omega_a, \omega_a)|\mathcal{E}_a(\omega_a)|^4 \mathcal{E}_a(\omega_a) + \frac{15\varepsilon_0}{18} \chi_{ND,el}^{(5)}(\omega_a; \omega_b, -\omega_b, -\omega_b, \omega_a)|\mathcal{E}_b(\omega_b)|^4 \mathcal{E}_a(\omega_a)
\]  

(2-28)

where \( \chi_{D,el}^{(5)} \) and \( \chi_{ND,el}^{(5)} \) are the degenerate and non-degenerate third-order nonlinear susceptibilities, respectively. The coefficients on the right hand side refers to the permutation symmetry of the fifth-order nonlinear susceptibility. Now, if we also apply Eq. (2-28) to the nonlinear wave equation, we get

\[
\frac{\partial}{\partial z} I_a(z) = -\left(\alpha_{0,a} + \Delta\alpha_{a,el}^{(3)}(z) + \Delta\alpha_{a,el}^{(5)}(z)\right) I_a(z),
\]

(2-29)

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z) = k_{0,a} \left( n_{0,a} + \Delta n_{a,el}^{(3)}(z) + \Delta n_{a,el}^{(5)}(z) \right),
\]

(2-30)

where

\[
\Delta \alpha_{a,el}^{(5)}(z) = \alpha_3^D I_a^2(z) + 3 \alpha_3^{ND} I_b^2(z),
\]

(2-31)

\[
\Delta n_{a,el}^{(5)}(z) = n_4^D I_a^2(z) + 3 n_4^{ND} I_b^2(z)
\]

(2-32)

Here, \( \alpha_3 \) and \( n_4 \) denotes to three-photon absorption (3PA) and 5-th order instantaneous bound-electronic NLR, respectively. We note that the 5-th order NLR is characterized by \( n_4 \) in the literature (as opposed to 3PA, which has a subscript of 3). This is perhaps because of the difference...
between order of nonlinearities contributing to \( n_2 \) (3\(^{rd} \) order effects) and \( n_4 \) (5\(^{th} \) order effects). \( \alpha_3 \) and \( n_4 \) are also related to fifth-order susceptibilities via

\[
\alpha_3^D = \frac{5}{2} \frac{\omega_a}{\varepsilon_0^2 c^3 n_{0,a}^3} \text{Im}\{\chi^{(5)}_{D,el}\} \quad \& \quad \alpha_3^{ND} = \frac{5}{2} \frac{\omega_a}{\varepsilon_0^2 c^3 n_{0,a}^2 n_{0,b}^2} \text{Im}\{\chi^{(5)}_{ND,el}\}
\]

\[
n_4^D = \frac{5}{4} \frac{\omega_a}{\varepsilon_0^2 c^2 n_{0,a}^3} \text{Re}\{\chi^{(5)}_{D,el}\} \quad \& \quad n_4^{ND} = \frac{5}{4} \frac{\omega_a}{\varepsilon_0^2 c^2 n_{0,a}^2 n_{0,b}^2} \text{Re}\{\chi^{(5)}_{ND,el}\}
\]

Eqs. (2-29) and (2-30) can be solved individually to find the irradiance and the induced phase change of the electric field while it propagates through the material. We will discuss the origins of bound electronic nonlinearities in more detail in CHAPTER 4.

2.3. Nonlinear Absorption and Refraction of Free-Carriers in Semiconductors and Metals

Nonlinearities of bound electrons are not the only mechanism that manipulates the propagation of light. Free-electrons and holes also alter the propagation of electromagnetic waves. Free-carriers in semiconductors need to be excited either optically or by electrical injection. Optical excitations typically include 1PA or 2PA. As opposed to semiconductors, free-carriers are always present in metals and doped semiconductors even without optical excitations. Thus, free-electron models such as Drude model are typically used to study them. Here, we discuss how free-carriers will influence the nonlinear wave propagation in both semiconductors and metals.

2.3.1. Free-carrier absorption and refraction in semiconductors

When semiconductors absorb a photon with an energy higher than its bandgap, \( \hbar \omega_p \geq E_g = \hbar \omega_g \), electrons start to partially fill the conduction band near the bottom edge and lower the number of filled levels at the top of the valence band. This phenomenon occurs in semiconductors when carriers are excited from top of the valence band to the conduction band through interband absorption and may happen either by absorbing one photon (e.g. 1PA) or multi-photons (e.g. 2PA, 3PA, etc.).
Subsequently, the bandgap of the material will be shifted from $E_g$ to $E_{g,\text{excited}}$. The extent of this shift depends upon number of photons or the intensity of the incident light, which are also called “band filling” effects [64-66]. The band filling effect will shift the absorption spectra predominantly close to $E_g$, and thus the intensity dependent absorption change will be larger close to the band edge. The direct consequence of the change in absorption spectra is also change in the refractive index, which its magnitude at each frequency can be mathematically derived via the Kramers-Kronig relation [67-69].

The electrons excited to the conduction band will eventually recombine and return back to the valence band. The recombination characteristic time, $\tau_r$, is typically around 10 ns for semiconductors at room temperature. The refraction will also change in accordance with the Kramers-Kronig relations. The electrons in the conduction band and holes in the valence band can absorb photons during this transitory time and make another transition to other energy bands through intraband absorption [65, 66, 70]. Absorption of holes between light and heavy hole bands also occurs. This process is typically referred to as free-carrier absorption (FCA). Subsequently, the refractive index will also be influenced by the FCA and its magnitude can be found by the Kramers-Kronig relations. The phenomenon of an induced change in the refractive index is also called free-carrier refraction (FCR).

In order to mathematically model the behavior of the propagation of light, we assign two parameters to the carriers that are generated through the interband absorption process. FCA cross-section, $\sigma_{\text{FCA}} \ (m^2)$, and FCR cross-section, $\sigma_{\text{FCR}} \ (m^2)$, which relates the change in irradiance and phase to the photo-excited charge-carrier density, respectively. In this case, we consider that each
free-carrier cross section corresponds to the sum of contributions from both electrons and holes.

The governing equations on irradiance and phase change will be

\[
\frac{\partial}{\partial z} I_a(z, t) = -\alpha_{0,a} I_a(z, t) - \Delta \alpha_{NL,inst} I_a(z, t) - \sigma_{FCA} N_{ex}(z, t) I_a(z, t),
\]

(2-35)

\[
\frac{\partial}{\partial z} \Delta \phi_a(z, t) = k_{0,a} n_{0,a} + k_{0,a} \Delta n_{NL,inst}(z, t) + \sigma_{FCA} N_{ex}(z, t),
\]

(2-36)

where \(N_{ex} \text{ (m}^{-3}\)) characterizes the photo-excited charge-carrier density, \(\Delta \alpha_{NL,inst}\) and \(\Delta n_{NL,inst}\) are instantaneous NLA and NLR discussed in §(2-2), respectively. Eqs. (2-35) and (2-36) are not complete without considering interband absorption by means of other absorptive processes. This process can be either due to absorption of photons via linear process or absorption of multi photons by means of NLA.

It is important to note that, as opposed to bound-electronic nonlinearities, free-carriers exhibit a non-instantaneous nonlinear response. To understand why this happens, we first study how carriers are generated by interband absorption processes. The simplest scenario is generation of carriers via linear absorption in the absence of instantaneous NLA that can be described by a set of significantly simplified rate equations as

\[
\frac{\partial}{\partial t} N_{ex}(z, t) = \frac{\alpha_{0,ex} I_{ex}(z, t)}{\hbar \omega_{ex}} - \frac{N_{ex}(z, t)}{\tau_{rec}}
\]

(2-37)

where \(\alpha_{0,ex}, I_{ex}(z), \) and \(\hbar \omega_{ex}\) are the 1PA coefficient, irradiance, and photon energy of the optical beam that excites carriers via interband absorption depending on whether the nonlinearities are single-beam or a dual-beam interactions. Typically, we ignore the second term in Eq. (2-37) and carrier diffusion because the recombination decay time is relatively long compared to the observation times. By substituting Eq. (2-37) in Eq. (2-35), the governing equations become
\[
\frac{\partial}{\partial z} I_a(z, t) = -\alpha_{0,a} I_a(z, t) - \sigma_{\text{FCA}} \frac{\alpha_{0,\text{ex}}}{\hbar \omega_{\text{ex}}} l_a(z, t) \int_{-\infty}^{t} I_{\text{ex}}(z, t') dt',
\]

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z, t) = k_{0,a} n_{0,a} + k_{0,a} \Delta n_{\text{NL,inst}}(t) + \sigma_{\text{FCR}} \frac{\alpha_{0,\text{ex}}}{\hbar \omega_{\text{ex}}} \int_{-\infty}^{t} I_{\text{ex}}(z, t') dt',
\]

where the subscript “ex” can either be “a” for the degenerate (single beam) interaction or “b” for the non-degenerate (double beam) interaction. We, therefore, define the fluence, \( F(t) \), as \( dF(t)/dt = I(t) \), and substitute it in Eq. (2-38). By integrating in time from \(-\infty\) to \(+\infty\) from both sides of the above equation, we obtain

\[
\frac{d}{dz} F_a(t) = -\alpha_{0,a} F_a(t) - \sigma_{\text{FCA}} \frac{\alpha_{0,\text{ex}}}{\hbar \omega_{\text{ex}}} \int_{-\infty}^{+\infty} dt F_{\text{ex}}(t) \left( \frac{d}{dt} F_a(t) \right),
\]

\[
\frac{\partial}{\partial z} \Delta \varphi_a(z) = k_{0,a} n_{0,a} + k_{0,a} \Delta n_{\text{NL,inst}}(t) + \sigma_{\text{FCR}} \frac{\alpha_{0,\text{ex}}}{\hbar \omega_{\text{ex}}} \int_{-\infty}^{t} I_{\text{ex}}(z, t') dt',
\]

For the degenerate case, Eq. (2-40) can be further simplified to

\[
\frac{d}{dz} F_a = -\alpha_{0,a} F_a - \sigma_{\text{FCA}} \frac{\alpha_{0,\text{ex}}}{2\hbar \omega_{\text{ex}}} F_a^2
\]

From (2-40) and (2-41), it becomes apparent that FCA and FCR of carriers generated by means of linear absorption are effectively scaling with fluence. This simply means that FCA and FCR can be considered as an effective third-order processes. These processes are also sometimes called a cascaded mechanism described by \( \chi^{(1)}: \chi^{(1)} \) processes [71, 72]. In order to accurately characterize the contributions of FCA and FCR to the nonlinear wave propagation, we used a numerical analysis based on discretizing the differential equations as discussed CHAPTER 4.
Another scenario that occurs in semiconductors is FCA and FCR of carriers generated by higher order processes than linear absorption, typically in our studies by 2PA. In this case the rate equation governing on $N_{ex}(z, t)$ becomes

$$\frac{\partial}{\partial t} N_{ex}(z, t) = \frac{\alpha_{2,ex} l_{ex}^2(z, t)}{2\hbar\omega_{ex}} - \frac{N_{ex}(z, t)}{\tau_{rec}},$$

(2-43)

where $\alpha_{2,ex}$ is the 2PA coefficient of the beam that excites the carriers (degenerate or non-degenerate). By ignoring the second term, we get

$$\frac{\partial}{\partial z} l_a(z, t) = -\alpha_{0,a} l_a(z, t) - \Delta\alpha_{NL,inst} l_a(z, t) - \sigma_{FCA} \frac{\alpha_{2,ex}}{2\hbar\omega_{ex}} \int_{-\infty}^{t} I_{ex}^2(z, t') dt',$$

(2-44)

$$\frac{\partial}{\partial z} \Delta\psi_a(z, t) = k_{0,a} n_{0,a} + k_{0,a} \Delta n_{NL,inst} (t) + \sigma_{FCR} \frac{\alpha_{2,ex}}{2\hbar\omega_{ex}} \int_{-\infty}^{t} I_{ex}^2(z, t') dt,$$

(2-45)

The factor of 2 in the denominator of the last term in both equations comes from the fact that two photons are needed to excite carriers. Similar to the previous case the subscript “ex” can either be “a” for the degenerate (single beam) interaction or “b” for the non-degenerate (double beam) interaction; however, in this case appropriate terms for $\Delta\alpha_{NL,inst}$ and $\Delta n_{NL,inst}$ need to be chosen from Eqs. (2-26) and (2-27) depending on what interaction we are considering. Similarly, from Eqs. (2-44) and (2-45) it becomes apparent that FCA and FCR induced by carriers generated via 2PA are quadratically scaling with the input irradiance. This means that these process are effectively fifth-order and they are sometimes referred to as cascaded effectively fifth-order mechanism explained via $\chi^{(3)}: \chi^{(1)}$ process [72, 73]. We also note that to accurately determine the contribution, a numerical integration is required, which we discuss more in CHAPTER 4.
2.3.2. **Drude model and nonlinearities of metals**

In this section, we briefly overview the nonlinear mechanisms that manipulate metallic materials. The transport properties of electrons in such materials can be explained by a simplified free-electron model developed by Paul Drude in 1900 [74-76].

In a general form, when electrons are bound to nucleus and assuming that the nucleus has an infinite mass compared to the electron, the equation of motion can be described via

\[ m \frac{d^2 \mathbf{r}}{dt^2} + m \gamma \frac{d \mathbf{r}}{dt} + m \omega_0^2 \mathbf{r} = -e E_{loc}, \]  

where \( m \) is the electron mass (\( m_e^* \) if we consider a finite mass for nucleus and consider all the screening effects), \( \gamma \) is the damping factor corresponding to the frictional force that causes the energy loss mechanism for accelerated electrons while excited, \( \omega_0 \) is the oscillator restoring force described by the Hooke’s law, \( \mathbf{r} \) is the position vector, \( e \) is the elementary charge, and \( E_{loc} \) is the local electric field on the particle. This model is called the Lorentz model.

The free-electron model, describing electrons of metals that are not bound, ignores the restoring effect from the nucleus. Therefore, Eq. (2-46) will be simplified further (discussed in detail in CHAPTER 5) and results in the Drude model, where the permittivity can be described via

\[ \varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma \omega} \]  

(2-47)

where \( \omega_p \) is the plasma frequency and \( \varepsilon_\infty \) is the permittivity at high frequencies. We can deduce the real and imaginary part of the permittivity, described as

\[ \text{Re}\{\varepsilon\} = \varepsilon'(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + \gamma^2} \]  

\[ \text{Im}\{\varepsilon\} = \varepsilon''(\omega) = \frac{\omega_p^2 \gamma}{\omega(\omega^2 + \gamma^2)} \]  

(2-48)  

(2-49)
In materials with relatively small loss ($\gamma \ll \omega$) and for high frequency, the Drude model simplifies to

$$\varepsilon'(\omega) \approx \varepsilon_\infty - \frac{\omega_p^2}{\omega^2}$$

(2-50)

$$\varepsilon''(\omega) \approx \frac{\gamma \omega_p^2}{\omega^3}$$

(2-51)

In the presence of nonlinearities, $\omega_p$ and $\gamma$ become irradiance dependent and induced changes in these parameters, $\Delta \omega_p$ and $\Delta \gamma$, will effectively manipulate the real and imaginary parts of the electric permittivity. In this case, for tiny nonlinear effects

$$\Delta\varepsilon' \approx \frac{\partial \varepsilon'}{\partial \omega_p} \Delta \omega_p = -2 \frac{\omega_p^2}{\omega^2} \frac{\Delta \omega_p}{\omega_p}$$

(2-52)

$$\Delta\varepsilon'' \approx \frac{\partial \varepsilon''}{\partial \omega_p} \Delta \omega_p + \frac{\partial \varepsilon''}{\partial \gamma} \Delta \gamma = \frac{2 \omega_p \gamma}{\omega^3} \Delta \omega_p + \frac{\omega_p^2}{\omega^3} \Delta \gamma$$

(2-53)

We can easily deduce NLR and NLA, $\Delta n_{\text{Drude}}$ and $\Delta \alpha_{\text{Drude}}$, respectively, from the Drude model, considering that $n_{\text{Drude}}^2 - \kappa_{\text{Drude}}^2 = \text{Re}\{\varepsilon\} \approx \varepsilon'$ and $2n_{\text{Drude}} \kappa_{\text{Drude}} = \text{Im}\{\varepsilon\} \approx \varepsilon''$. In this case, by taking derivatives from both sides, we get $2n_{\text{Drude}} \Delta n_{\text{Drude}} - 2 \kappa_{\text{Drude}} \Delta \kappa_{\text{Drude}} = \Delta \varepsilon'$ and $2n_{\text{Drude}} \Delta \kappa_{\text{Drude}} - 2 \kappa_{\text{Drude}} \Delta n_{\text{Drude}} = \Delta \varepsilon''$. We can then extract NLA and NLR of Drude material, which can be described via

$$\Delta\alpha_{\text{Drude}} \approx \frac{\omega n_{\text{Drude}}}{\varepsilon'} (n_{\text{Drude}} \Delta \varepsilon'' + \kappa_{\text{Drude}} \Delta \varepsilon')$$

(2-54)

$$\Delta n_{\text{Drude}} \approx \frac{1}{2 \varepsilon'} (n_{\text{Drude}} \Delta \varepsilon' + \kappa_{\text{Drude}} \Delta \varepsilon'')$$

(2-55)

By substituting $\Delta \varepsilon'$ and $\Delta \varepsilon''$, we find the dependence of NLA and NLR on $\Delta \omega_p$ and $\Delta \gamma$. We will discuss Drude nonlinearities and the physical origin of optically induced modifications in more detail in CHAPTER 5.
In summary, we conclude that NLA and NLR can be due to different physical mechanisms, including instantaneous effects (from bound-electrons), described by nonlinear susceptibilities, and non-instantaneous effects requiring temporal integration to characterize their contribution to the NLA and NLR. Therefore, considering all different phenomena discussed here, we can consider NLA and NLR from different mechanisms by

\[
\Delta \alpha_{NLA} = \left( \Delta \alpha_{el}^{(3)}(z) + \Delta \alpha_{el}^{(5)}(z) + \cdots \right) + (\Delta \alpha_{FCA} + \Delta \alpha_{Drude}) + \cdots \tag{2-56}
\]

\[
\Delta n_{NLR} = \left( \Delta n_{el}^{(3)}(z) + \Delta n_{el}^{(5)}(z) + \cdots \right) + (\Delta n_{FCA} + \Delta n_{Drude}) + \cdots , \tag{2-57}
\]

where the first part is due to bound-electronic instantaneous nonlinearities and the second part is from carrier-induced non-instantaneous nonlinearities. The magnitude of each individual term depends upon both the material under consideration and the optical frequency.

2.3.3. Extremely Large Nonlinear Optical Effects and Failure of Perturbative Analysis

Typically, bound-electronic nonlinearities are much smaller, but faster, than other types of nonlinearities. We usually consider \( \Delta n_{el} \) and \( \Delta \alpha_{el} \) to be significantly smaller than the initial linear values of \( n \) and \( \alpha \), and thus they can be considered as perturbations to the linear optical constants of the material. Similar arguments can also be made for carrier nonlinearities; however, typically carrier induced nonlinearities are larger than bound-electronic NLO effects. In the past decades, a huge research endeavor has been dedicated to increase nonlinear optical effects, particularly NLA and NLR, such as local enhancement of the irradiance [77-80], engineering material nonlinear properties [81-86], and optimizing the light-matter interaction schemes [1-4, 7, 9, 30, 87-92]. In such cases, the conventional definition of nonlinear optical coefficients are not entirely valid and other arrangements need to be considered. We discuss these effects in more detail in CHAPTER 5.
CHAPTER 3: EXPERIMENTAL TECHNIQUES

This chapter provides an introduction to experimental techniques used in this dissertation. We discuss both linear and nonlinear experimental apparatus used in our characterizations. Linear spectroscopy techniques characterize the basic linear optical properties of the material and help determine the strategy used in nonlinear experiments. For our nonlinear spectroscopy experiments, an intense light source is needed. In this work, we use femtosecond laser systems to generate ultrashort pulses and nonlinear conversion instruments to create new frequencies for our spectroscopic characterization. Depending on the material system and the physical mechanism of interest, we either use single beam techniques (e.g. Z-scan and nonlinear transmission) or double beam techniques (pump-probe, Beam-Deflection (BD), etc.). We also discuss self- and cross-phase modulation measurements in which we measure the induced shifts in the carrier frequency of the pulse.

3.1. Linear Spectroscopy Techniques

The knowledge of transmission, reflection, and absorption spectra are essential to understand the linear optical properties of materials and to design a proper nonlinear spectroscopy experiment. We normally use a CARY-500 spectrophotometer to measure the transmission and absorption spectra allowing us to characterize the spectral dependence of the absorptive properties of the material. Measuring the refractive index usually becomes more sophisticated particularly for measuring thin-films. We discuss here some of the basic principles of the spectroscopic ellipsometry and discuss this method in more detail in CHAPTER 5.
3.1.1. **Spectroscopic ellipsometry**

Spectroscopic ellipsometry allows us characterizing the spectral dependence of the dielectric response of thin films and bulk materials. This technique can also be used to characterize thickness, roughness, crystalline nature, and electrical conductivity of the material. The working principle of this technique is based on measuring the ellipticity of the polarization of the reflected wave. It is called spectroscopic ellipsometry particularly because it characterizes the spectral dependence of the reflected ellipticity. Here, we briefly explain the principles of this characterization technique.

Ellipsometry measures the change in polarization state of the reflected beam of light from the sample. Depending on the thickness, refractive index, absorption, and other optical properties of the sample, the polarization will be altered differently. The incident light is polarized by passing through a polarizer, as shown in Figure 3-1.

![Schematic setup for spectroscopic ellipsometry apparatus.](image)

The incident beam is composed of two distinct polarization components of $s$ and $p$, where they are perpendicular and parallel to the incident plane, respectively. The Fresnel reflection coefficients for each individual polarization components are denoted by $r_s$ and $r_p$, and their complex ratio is represented by an amplitude ($\tan \psi$) and phase ($\Delta$), described by

$$\rho = \frac{r_p}{r_s} = \tan(\psi) \cdot e^{i\Delta}$$

(3-1)
Upon reflection, each incident polarization component experiences a change in its amplitude and phase, which subsequently results in an elliptic polarization for the reflected wave. The incident angle, ϕ, is chosen in order to obtain the maximum difference between each polarization, which typically can be achieved near the Brewster angle. The raw data measured by the ellipsometry gives us ψ and Δ. These two parameters by themselves are not very informative and a detailed data analysis is required to extract the film thickness, optical properties, roughness, and etc. In some cases, the user needs to design a model to fit the experimental data, but in most of the cases the built-in fit models are sufficient to precisely characterize the optical constants. As the structure becomes more complicated (e.g. multilayer materials, lossy films, anisotropic structures, etc.), a new fit model is essential to extract the optical properties.

3.2. Ultrashort Light Sources

Before discussing the nonlinear spectroscopy methods used in this dissertation, it is important to note that, as opposed to linear spectroscopy techniques, for nonlinear spectroscopy we need high intensity laser sources. We first give a brief overview of the laser sources used in this work. The first section briefly discusses the working principle of amplified ultrafast laser sources and the second section discusses tunable nonlinear parametric conversion systems used to generate new frequencies for nonlinear spectroscopy.

3.2.1. Laser Sources

For the purpose of nonlinear measurements reported here, we use commercially available chirped pulse amplified (CPA) Ti:sapphire femtosecond laser systems. The main laser system used in this dissertation is a Clark MXR CPA-2010 source (for some of the experiments a different laser system, Coherent Legend Elite Due HE+, with almost similar working principle but different performance is used). The Clark system is composed of different sections, as shown in Figure 3-2. The
oscillator is a unidirectional, polarization rotation additively pulsed mode-locked Erbium doped fiber ring laser operating at 1550 nm. When modelocked, the repetition rate of the output pulse is 27 MHz. This fiber laser is pumped by a solid-state fiber coupled laser diode operating at 980 nm. The output of the fiber laser is doubled using a temperature-stabilized periodically poled Lithium Niobate (PPLN) crystal to generate output of 775 nm, which is then used as a seed for the CPA system. In order to avoid any potential damage to the gain medium, a diffraction grating pulse stretcher is used to temporally extend the 775 nm pulse before amplification. An electro-optic modulator (Pockel cell) is used to allow a single pulse to enter the cavity. The pulse inside the cavity gets amplified while propagating through a Ti:sapphire crystal, which is pumped by a Q-switched Nd:YAG laser operating at 532 nm. The Nd:YAG laser itself is pumped by a flash lamp and it generates 1064 nm pulses, which are subsequently doubled using a Lithium triborate (LBO) crystal. After reaching the desired amplification, another Pockel cell is used to switch the amplified pulse (~ 1 mJ energy) out of the cavity. At this point, the amplified extended pulse is routed towards a diffraction grating compressor to yield ~150 fs (FWHM) with a spectral bandwidth of ~6 nm (FWHM) pulses at the output.

Figure 3-2. System configuration of a regeneratively ultrafast laser system based on CPA mechanism.
3.2.2. **Nonlinear optical parametric amplifiers**

The amplified output of the CPA system is divided into two parts using a 90/10 plate beam-splitter. 10% of the 775 nm beam is reflected from the beam-splitter and used in some of our experimental setups. The remaining 90% is transmitted through the beam-splitter and precisely aligned into a nonlinear parametric conversion system (Light Conversion Ltd. – TOPAS C) operating as an optical parametric amplifier (OPA) to provide broadly tunable sources for nonlinear spectroscopy.

The 775 nm beam (photon energy of $\hbar \omega_p$) is used to generate a white-light continuum (WLC) by pumping a thick sapphire plate. The WLC beam is then chirped using a ZnSe plate to temporally separate different frequencies and overlapped with the pump in a low temperature $\beta$-Barium borate (BBO) second-order nonlinear crystal in a non-collinear geometry. This non-collinear optical parametric amplification processes (also called NOPA) generates the signal pulse (photon energy of $\hbar \omega_s$), as shown in Figure 3-3, with central wavelength between 1100 – 1550 nm depending on the rotation angle of the crystal and phase matching conditions. At the same time, the idler beam (wavelength range of 1550 – 2620 nm) with a photon energy of $\hbar \omega_i = \hbar \omega_p - \hbar \omega_s$ is also generated; however, at this stage of non-collinear amplification the idler beam is manually blocked.

![Figure 3-3. Nonlinear parametric conversion mechanism in a $\chi^{(2)}$ crystal.](image)

The amplified signal beam will interact with the remaining of the pump beam in another BBO crystal amplification stage in a collinear geometry. Upon this interaction the idler beam will also
be generated and amplified via the nonlinear parametric conversion process. At this stage the amplified signal and idler beam will propagate collinearly to the output of the OPA system. A pair of dielectric mirrors are then used to separate signal and idler beams at the output of OPA.

3.3. Nonlinear Spectroscopy Techniques

The ever-increasing role of nonlinear optics has created a pressing demand for the characterization of various nonlinear optical phenomena. From the invention of laser in 1960, different experimental configurations have been developed to characterize the NLO response of various materials such as Optical Kerr Effect (OKE) [93-96], degenerate four wave mixing (DFWM) [97-101], both of which require optical heterodyne detection (OHD) in order to tell the sign of the phase change [102, 103]. In this section, we overview two experimental configurations developed in our group, namely Z-scan and BD techniques [104-106]. Z-scan can be used for the purpose of characterizing the sign and magnitude of NLA and NLR. However, there are some limitations with this method that led us the development of the BD technique. Here, we primarily focus on time-resolved techniques, particularly BD method, which allows us to simultaneously characterize the temporal- and polarization-dependence of both NLR and NLA. Some modifications of the BD method will be discussed in more detail later.

3.3.1. Z-scan

The Z-scan technique, invented in 1989, is a simple method to characterize sign and magnitude of temporally averaged NLA and NLR [105, 106]. The advantage of this method is its simplicity requiring only a single beam to characterize both NLA and NLR simultaneously. Additionally, because of the interferometric mechanism of Z-scan, this method is very sensitive in measuring changes as small as $\lambda/200$ induced on the beam wavefront. However, Z-scan cannot directly give
us enough information about the physical origin of the NLA and NLR. In order to accurately determine the physical origin of the measured nonlinearity, a precise data analysis strategy based on the signal shape and the irradiance dependence is necessary [106]. Additionally, extracting NLR from Z-scan measurements for materials with large NLA or NLR background can become difficult in some cases. Z-scan has been expanded to several forms such as two-color Z-scan [107], eclipsing Z-scan [108], Z-scan for thick samples [109], dual-arm (DA) Z-scan for solutions and thin-films [110, 111]. These expansions have been to some extent successful to address some of the limitations of the Z-scan method.

3.3.1.1. Principles

The principle of Z-scan technique is based on continuously changing the irradiance of the incident pulse by focusing the beam to a small size and measuring the transmitted energy. In this method, the sample is placed on a motorized stage and scanned through the focus of a laser beam. While the sample is being moved, the energy of the transmitted beam in the far-field is measured with and without a partially closed aperture simultaneously. The scan of the portion of the beam not passing through any aperture is referred to as the open-aperture (OA) Z-scan and the portion that is passing through an aperture is referred to as the closed-aperture (CA) Z-scan. The signal from OA Z-scan is purely due to NLA, such as 2PA, 3PA, FCA, and etc. The signal from CA Z-scan includes a mixed contribution of both NLA and NLR. By calculating CA/OA, we isolate the NLR, such as bound-electronic NLR (Kerr effect) and FCR, and eliminate the contribution of NLA. In this method we consider the sample is thin enough that the effect of diffraction and NLR on the beam size within the sample is negligible. This assumption is also referred to as the “thin” sample approximation [106]. To fulfill this assumption it is usually enough (for small NLO effects) to use
a focusing geometry in which the sample thickness, \( L \), is smaller than the Rayleigh range of the beam in free-space, \( z_0 = \frac{\pi w_0^2}{\lambda_0} \).

![Diagram](image)

Figure 3-4. Schematics of (a) open-aperture Z-scan to characterize NLA, and beam divergence in closed-aperture Z-scan using a sample with positive \( n_2 \) (b) after and (c) before the focus to characterize NLR.

To understand the principles of this method, consider that the sample is scanned through a focusing beam, as shown in Figure 3-4. At the focus, the irradiance, \( I_0 \propto w_0^{-2} \), will be the largest and it decreases quadratically with the beam size. The NLA will be larger for higher irradiances, and thus the dip in transmission will be maximized at the focus. The shape of the normalized change in transmission, \( \Delta T/T \), versus the scan range depends on the spatial profile of the incident beam. The irradiance at each position, \( Z \), can be calculated for a Gaussian beam and used to fit the OA data to extract NLA coefficients such as \( \alpha_2 \), \( \alpha_3 \), and \( \sigma_{\text{FCA}} \). Measurement of NLR is typically more...
complicated and it requires a clean Gaussian TEM$_{00}$ incident beam. If the laser beam does not have a good beam profile, a spatial filtering telescope system is used to clean and obtain a TEM$_{00}$ beam [112]. This process becomes even more important in DA Z-scan method [104, 111], which will be discussed in more detail later. A partially closed aperture (typically ~ 33% transmission) is used in the far field, which makes the transmitted energy to be sensitive to focusing or de-focusing effects originated from NLR, as shown in Figure 3-4 for positive NLR. The refractive index change of the material follows the spatial Gaussian profile of the incident beam, and thus the change is larger at the center than the wings of the beam cross-section. The wavefront of the incident beam will be modified like a positive lens due to larger nonlinear phase change at the center of the beam, hence the divergence of the beam will be altered depending on the sample position, as shown in Figure 3-4. For instance, placing the sample before the focus, the positive NLR focuses the beam more and increases the divergence, resulting in a larger spot at the aperture. This is evident in CA Z-scan as a dip in the detected transmitted energy. If the sample is placed after the focus, the divergence of the incident beam will become less in presence of positive NLR, and thus a peak in transmission will be observed.

The process to calculate the transmission change at each position is first to calculate the complex electric field right at sample output plane, $\mathcal{E}_{\text{out}} = |\mathcal{E}_{\text{out}}|e^{i\Delta \phi_{\text{out}}}$ and then use Fresnel diffraction propagation to calculate the complex electric field on the detector or the partially closed aperture. For a beam propagating inside a material with a linear refractive index of $n_0$, the irradiance is defined as $I_{\text{out}} = 1/2\varepsilon_0 nc|\mathcal{E}_{\text{out}}|^2$. For a bound-electronic third-order nonlinear process with related NLO coefficients of $\alpha_2$ and $n_2$, in presence of linear absorption, $\alpha_0$, we have [105, 106]
\[ I_{\text{out}} = \frac{I_{\text{in}} e^{-\alpha_0 L}}{1 + \frac{\alpha_2}{\alpha_0} I_{\text{in}} (1 - e^{-\alpha_0 L})} \] (3-2)

\[ \Delta \varphi_{\text{out}} = k_0 n_2 \frac{\ln \left(1 + I_0 \frac{\alpha_2}{\alpha_0} (1 - e^{-\alpha_0 L})\right)}{\alpha_2} \] (3-3)

where \( I_{\text{in}} \) is the irradiance of the incident beam, \( k_0 \) is the wavevector in free space, and \( L \) is the thickness. These two relations can be easily obtained by solving the differential equation governing on the light intensity, in the presence of 1PA and 2PA, and then substituting the obtained intensity into the equation governing on the nonlinear phase change for Kerr nonlinearities. It is evident that the nonlinear phase change at the plane right after the sample is altered by both 1PA and 2PA process; however, we ignore the effect of \( n_2 \) in the beam size because of the thin sample approximation discussed earlier. Now that we know both irradiance and phase change, we can calculate the complex amplitude of the electric field via

\[ E_{\text{out}} = \frac{\sqrt{2 I_{\text{out}}(r, z, t)}}{\varepsilon_0 n c}, \]

where \( r, z, \) and \( t \) are the transverse spatial, longitudinal spatial, and temporal coordinates. The details of the Fresnel diffraction theorem and numerical calculations are reported in Ref. [105, 106, 112].

As an example, here we show OA Z-scan of a 0.35 mm thick bulk GaAs and CA Z-scan of a 1 mm thick fused silica at \( \lambda = 1100 \) nm, beam size of \( w_0 = 22 \) µm (HW1/\( e^2 \)M), and pulse-width of \( \tau = 215 \) fs (FWHM). These measurements were fit using the Z-scan theory discussed in detail in Ref. [105, 106].
OA Z-scan measurements on a 0.35 mm thick GaAs bulk material and fit to characterize $2\alpha \approx 6 \text{ cm/GW}$. (b) CA Z-scan measurements on a 1 mm thick fused silica to characterize electronic NLR and fit $n_2 \approx 0.25 \times 10^{-15} \text{ cm}^2/\text{W}$.

OA Z-scan on a two-photon absorbing material such as GaAs, and CA Z-scan on a purely refractive nonlinear material such as fused silica are the two preliminary scans usually done in every Z-scan experiments. The reason for this calibration measurement is that depending on the width of the OA scan and shape of the CA scan we can easily approximate some properties of the beam such as beam waist and pulsewidth. Typically, for relative measurements this method is very useful; however, for accurate characterization of the beam waist and pulsewidth other direct measurement methods are required.

Since fused silica does not show any NLA at 1100 nm, the CA measurement on fused silica will be purely due NLR effects. When material exhibits NLA properties, one needs to fit the OA data and find the contribution of NLA, and subsequently use this information to find a NLR parameter that fits the CA data. It is also possible to divide the CA to OA data to eliminate the NLA and fit CA/OA for NLR. However, when NLA is present this process decreases the signal to noise significantly, and thus analyzing the data becomes extremely difficult. To address this issue we briefly discuss advantages of DA Z-scan technique in the next section [110, 111].
3.3.1.2. Dual-arm Z-scan

In some cases, the background nonlinearity is large enough that it obscures the nonlinear effect of the sample under consideration. For instance, measuring the NLR of organic molecules dissolved in a solvent can become complicated when the NLR of solvent is dominant in the CA Z-scan. Usually, these measurements are performed sequentially where the CA Z-scan of the solvent is subtracted from that of the solution resulting in the signal from solute. Note that solvents do not possess NLA at the spectral range of interest for most of the nonlinear organic compounds, and thus sequential OA Z-scans for organic solutions usually results in clean signals. Similar situation also happens for thin-films particularly when the nonlinear contribution of the film to the CA Z-scan is relatively smaller than the thick substrate.

The DA Z-scan is an expansion of the original single arm Z-scan technique that was developed initially to measure the NLR of organic solutions [110]. This method improves the signal to noise ratio by an order of magnitude via simultaneously performing Z-scan measurements on the sample and the background material. In this method, a 50/50 beam splitter is used to divide the beam into two paths with identical pulse energies (additional filters may be used to control the energy of each path separately.). To use the full capability of DA Z-scan, exactly identical beam characteristics are needed. Therefore, the setup is designed in a way that every optics used in each individual arm and the optical path lengths are as similar as possible.
Figure 3-6. Schematics for DA Z-scan with the main sample (solution or thin-film sample) at arm A and the background sample (solvent or substrate) at arm B.

To match the Z-scan setup in both arms a series of calibration and equalization procedures are applied. First, by performing single arm OA Z-scans on a two-photon absorber material (such as GaAs or ZnSe depending on the wavelength) the beam waist is found and compared. Second, single arm CA Z-scans are performed on fused silica, CS$_2$, or any other purely refractive nonlinear material with relatively large signal. This step is necessary to compare the CA signals and equalize the incident beam energy and longitudinal position of each sample by separately controlling the position of each sample with 3-dimensional stages. The final step is to perform DA Z-scans simultaneously on the main sample and the background.
Figure 3-7. DA Z-scan measurements of an organic film (AJCPE04-TCF04) of thickness 1.29 µm deposited on a 1 mm thick quartz substrate at a wavelength of 1300 nm using multiple irradiances. (a) and (b) are OA and CA/OA signals, respectively. The effects of substrate and LEB are subtracted from the signal. (c) Sequential measurements and (d) simultaneous measurements (DA Z-scan) show the comparison of the subtraction of the substrate CA/OA Z-scan from the film plus substrate CA/OA Z-scan in the case of 54 nJ input. The solid line in (d) is the best fit as in the top curve of (b).

A comparison of sequential Z-scan and DA Z-scan of an organic film (AJCPE04-TCF04) deposited on a thick glass substrate is shown in Figure 3-7 (Figure from Ref. [111]). The OA Z-scan measurements of substrate is not required; however, from the two bottom curves, it is evident that sequentially subtracted Z-scan measurements are not useful in this material as opposed to DA Z-scan measurements, which allows effectively extracting the contribution of the organic film.
DA Z-scan addresses the complications of sequential Z-scan in measuring relatively small NLO response when large background effect is present. But, Z-scan in general, only measures the temporally averaged transmittance, and does not provide us with any information related to the temporal response of the NLO mechanism under study. This limitation can become problematic when studying for example non-instantaneous nonlinear effects such as nuclear third-order NLR, FCA, and FCR, particularly when pulsewidth is in the picosecond or nanosecond regime. To distinguish between bound-electronic and carrier induced nonlinearities in Z-scan measurements, it is required to perform irradiance dependent measurements and also to use numerical techniques to fit the shape of experimental curves. For example, in molecular liquids, the NLR essentially becomes pulsewidth dependence because of molecular nuclear reorientations, and thus Z-scan measurements with various pulsewidths are necessary to accurately determine the pulsewidth dependent NLR \[113, 114\]. Similarly, in presence of FCA and FCR, Z-scan measurements with several irradiances are needed to extract the scaling of the normalized change in transmission, which subsequently can be used to determine FCA and FCR. In such cases, we can use Eqs. (2-37) to (2-41) to fit the maximum change in transmission \[72, 73\] and find both instantaneous and non-instantaneous NLO coefficients or instead use time-resolved methods to separate the effect of each mechanism in temporal domain.

3.3.2. Beam-Deflection

As we discussed in §3.3.1. , Z-scan technique becomes complicated and somewhat inaccurate when carrier-related nonlinear mechanisms are dominant or when NLA background becomes very large. In particular, Z-scan measurements of bound-electronic nonlinearities can be less accurate for longer pulsewidth than with shorter pulsewidth because of the competing effect of FCA. This
can dominate the bound-electronic NLA when using picosecond and nanosecond pulsewidths. Similarly, carrier nonlinearities leading to FCR effects makes collecting all the transmitted energy difficult [115]. Double beam techniques such as pump-probe methods; however, are useful when temporal dynamics of the NLO mechanism are of interest [116]. As we discussed in CHAPTER 2, carrier nonlinearities are non-instantaneous, and thus time-resolved measurements are extremely useful to distinguish them from bound-electronic nonlinearities. Additionally, the possibility to separately controlling the polarization of each beam helps characterizing the anisotropy of the NLO response, which subsequently provides valuable information on the physical origin of the nonlinear mechanism. Double beam techniques also allow measuring non-degenerate nonlinear interactions where the frequency of the two beams are not the same.

BD technique, which is primarily used in this dissertation, is a double beam time-resolved method that directly measures the time-dynamics, sign, magnitude, and anisotropy of the NLR and NLA response. This method has a superior sensitivity compared to Z-scan and allows measuring changes as small as $\lambda/20,000$ induced on the beam wavefront. This value corresponds to an induced phase change of 0.3 mrad, or optical path length change on the order of Bohr radius [104]. BD can be used to measure NLA and NLR of solids (e.g. bulk semiconductors, plasmonic and metallic films, 2D layers, etc.) [4, 87, 88, 117-126], liquids (e.g. solvents, organic solutions, etc.) [21, 113, 114, 127, 128], and gases (solvents in gas phase, pressurized air, ambient air, etc.) [129, 130]. The high sensitivity of this method and the possibility of using Lock-in detection allows measuring NLO response of very dilute materials such as ambient air [129, 130] or very thin materials such as a monolayer of Graphene [124].

The principle of this method is very similar to pump-probe technique for characterizing the transient absorption of materials. In pump-probe, an intense beam excites a nonlinear mechanism
inside the sample and the weak probe beam experiences the induced changes. The NLA mechanism manipulates the transmitted energy of the weak probe beam and by measuring this change versus the time delay between the two pulses, we can extract time-dynamics of the NLA. The BD method, however, is similar to the photo thermal BD technique [131], in which slow thermal nonlinearities create an index gradient that deflects the weak probe beam. In ultrafast BD, the pump pulse also creates an index gradient via ultrafast NLR (discussed in detail in the previous chapter), which subsequently deflects the weak probe beam. By measuring the deflection angle versus the delay between the two pulses, the response dynamics of NLR can be characterized. Simultaneously, by measuring the total transmitted energy versus the delay between the two pulses, the dynamics of NLA can also be characterized. The index gradient that is created by the pump pulse is directly related to the irradiance of the pump beam, and its dependence is governed by the order and type of the NLO response.

Typically, the pump pulse is considered to be a TEM$_{00}$ Gaussian, as shown in Figure 3-8. In measurements requiring absolute characterization of nonlinearities, we usually spatially filter the beam to have a Gaussian distribution. However, in our experiments, where we typically do relative measurements, we first measure the BD signal of a well characterized material (e.g. fused silica or CS2) [104, 113, 114] to calibrate our experimental setup measurements. Thus, having a TEM$_{00}$ beam profile is not necessary for relative measurements; however, having a symmetric beam reduces the complexity of the measurements and analyses. The probe pulse is usually focused to a spot size $\sim 3-5 \times$ smaller than that of the pump beam. The probe is displaced to a place where the index gradient formed by the pump is nearly linear and maximum. The index gradient deflects the probe by a small angle, which is measured in the far-field using a sensitive segmented detector. In our experiments, a segmented detector with four quadrants is used and depending on
the experimental configuration, either the difference of the energy falling on left and right, \( \Delta E_H = E_{\text{left}} - E_{\text{right}} \), or falling on top and bottom, \( \Delta E_V = E_{\text{top}} - E_{\text{bottom}} \) is measured. The choice of vertical or horizontal deflection entirely depends upon how pump and probe are interacting. Consider the plane that includes the two beams in lab coordinates, if this plane is horizontal then we maximize the vertical deflection, but if this plane is vertical then we maximize the horizontal deflection. This is only because while two pulses are travelling through each other, the maximum index gradient occurs in the perpendicular plane to the plane that includes both pulses. The normalized signal to the total energy on the detector, \( \Delta E/E \), versus time directly characterizes the time-dynamics of the accumulated nonlinear phase-change, \( \Delta \phi_{\text{NL}} \), along the propagation direction, which also directly depends upon NLR.

Figure 3-8. (a) Experimental schematics of BD setup; (b) spatial distribution of the irradiance of the pump beam (red) and displaced smaller probe beam (blue) at the position of sample; (c) vertically deflected probe beam on a segmented quad-cell position sensitive detector with and without deflection.

The ability to directly measure the time-dynamics of NLR is the main advantage of BD compared to Z-scan or other dual beam methods such as OKE and FWM, which require OHD. Additionally, BD allows to resolve the polarization dependence, by varying the angle of relative
polarization between pump and probe pulses. This polarization dependence will be useful to determine different $\chi^{(3)}$ tensor elements in bound-electronic NLR, and also to separate the contribution of non-instantaneous nuclear contributions to NLR in liquids [113, 114]. As opposed to single beam techniques such as Z-scan, by varying the wavelengths of the pump and probe, BD characterizes nondegenerate NLO mechanisms, allowing characterization of nondegenerate NLR spectra [117].

To precisely fit the experimentally obtained signals, an accurate numerical model is required; however, for the experiments done in this dissertation analytical formulas can be derived with some approximations. Let us first consider there is no NLA and the sample is placed at the beam waist of the probe, with a spot size of $w_{0,p}$ (HW1/e$^2$M), and the beam size of the pump pulse on the sample is $w_e$ (HW1/e$^2$M), which is not necessarily the beam waist of the pump beam. The subscripts “p” and “e” denotes to probe and excitation (pump), respectively. Let us also consider that the wavelengths of pump, $\lambda_e$, and probe, $\lambda_p$, are nearly degenerate or the group velocity mismatch (GVM) of the material is negligible for these two wavelengths. Also, keeping the two beams collinear, the spatial and temporal walk-off along the propagation will be negligible. We also consider that the thickness $L$ is much less than the Rayleigh range of the probe pulse ($z_{0,p} = \pi w^2_{0,p}/2$) such that the irradiance of the pump is nearly uniform within the thickness of the sample. The induced NLR on the sample plane for a given time, $t$, and across the spatial irradiance Gaussian profile of the pump beam, $r$, can be described as

$$\Delta n(r, t) = \Delta n(t) \exp(-2r^2/w^2_e)$$

(3-4)

With no loss of generality, let us consider a vertical deflection meaning that the probe beam is initially displaced vertically to experience an index gradient along that direction. Therefore, the
maximum index gradient occurs at a position where the derivative of \( \exp(-2y^2/w_e^2) \) is maximized that is \( x = 0 \) and \( y = w_e/2 \). For a horizontal deflection, this condition will be fulfilled for \( x = w_e/2 \) and \( y = 0 \). Now, for a thin sample the deflection angle can be found via

\[
\theta (r, t) = \int_0^L \nabla[\Delta n(r, t)] \, dz \approx \nabla[\Delta n(r, t)]L = \frac{2L}{w_e \sqrt{e}} \Delta n(t)
\]  

(3-5)

Considering the Gaussian probe propagating a distance of \( d \) to the detector, the displacement of the beam on the detector will be \( \Delta = \theta(r, t) \cdot d \). By integrating over the \( xy \)-plane and some approximations [104], we can calculate the normalized deflection as

\[
\frac{\Delta E(\tau)}{E} = \frac{\sqrt{2} \, 2^{w_{0,p}}}{\sqrt{\pi \, e \, w_e}} \cdot k_{0,p} L < \Delta n(\tau) >
\]  

(3-6)

where \( <\Delta n(\tau)> = \int_{-\infty}^{\infty} \Delta n(t)I(t - \tau)dt / \int_{-\infty}^{\infty} I(t - \tau)dt \) is the temporal-dependence of the NLR, \( \tau \) is the relative delay between the two pulses, and \( k_{0,p} = 2\pi/\lambda_p \) is wavevector of the free space at the probe wavelength. This relation can become more generalized by substituting \( \Delta \varphi(\tau) = k_{0,p} L < \Delta n(\tau) > \), which indicates that the normalized BD signal directly characterizes the nonlinear phase change. Note that this relation can be used for both instantaneous bound-electronic and non-instantaneous NLR. For instance, considering \( \Delta n(t) = 2n_2(\lambda_p; \lambda_e)I_e(t) \) for a non-degenerate instantaneous NLR, we obtain

\[
\frac{\Delta E(\tau)}{E} = \frac{2\sqrt{2}}{\sqrt{e\pi}} \frac{w_{0,p}}{w_e} k_{0,p} L \frac{2n_2(\lambda_p; \lambda_e)I_{0,e}}{\sqrt{1 + \tau_e^2 / \tau_p^2}} \exp\left(-\frac{\tau^2}{\tau_e^2 + \tau_p^2}\right).
\]  

(3-7)

The equation for bound-electronic NLR indicates that \( \Delta E/E \) directly follows the temporal cross-correlation of the pump and probe pulses resulting in a Gaussian width (HW1/eM) of \( \tau_{BD} = \left(\tau_e^2 + \tau_p^2\right)^{1/2} \). This is the direct consequence of the instantaneous mechanism of bound-electronic NLR. By characterizing \( w_{0,p}, w_e, \) and \( \tau_e \), we can then extract the absolute values of \( n_2 \). As an
example, we show the BD measurement results of a 1 mm thick fused silica in Figure 3-9 for two relative polarizations of parallel (black) and perpendicular (perpendicular) while also characterizing beam parameters to extract absolute values. The left vertical axis corresponds to the BD signal, $\Delta E/E$ ($\%$), and the bottom horizontal axis corresponds to the delay between the two pulses. In this measurement, $\lambda_p = 1250 \, nm$, $\lambda_e = 775 \, nm$, $w_e = 200 \, \mu m$ (HW1/e^2M), $w_{0,p} = 70 \, \mu m$ (HW1/e^2M), and $\tau_e = 160 \, fs$ (FWHM). The pump and probe pulses are also spectrally filtered with a 10 nm bandpass filter. By fitting the data, we extract $n_{2,||} = (3 \pm 0.5) \times 10^{-16} \, cm^2/W$ and $n_{2,\perp} = (1 \pm 0.2) \times 10^{-16} \, cm^2/W$, which are in close agreement with the values previously reported in the literature [104]. The ratio of parallel to perpendicular polarization can be estimated to be around 3 as expected for the anisotropy of the isotropic instantaneous bound-electronic NLR.

![Figure 3-9. BD measurements, $\Delta E/E$ ($\%$) versus delay, of a 1 mm thick fused silica for relative polarizations between pump and probe of parallel (black) and perpendicular (red).](image)

It is important to note as long as the probe size is at least $\sim 3-5 \times$ smaller than that of the pump beam, we can use the BD measurements on fused silica to calibrate the experimental configuration without having any concerns on the absolute characterization of the beam sizes and pulsewidths.
Performing relative measurements is necessary particularly when the goal is to measure the nonlinear spectra. This is mainly because measurements in different wavelengths are not exactly identical, and thus having a calibration data for each wavelength is critical.

The ability to control the polarization of each beam allows characterizing different nonlinear mechanisms separately. In BD setups, a pair of half-wave plate (HWP) and polarizer is used to control the energy of the incident pump beam. Additionally, a HWP is used after the polarizer to control the linear polarization. In some cases, where we want a circular polarization, a quarter-wave plate is also used to change the linear polarization to a circular polarization. In most of the experiments, there is no need to have a circular polarization unless if some of the nonlinear mechanisms that are sensitive to circular polarization are under investigation. For the probe beam, typically we use a linear polarizer to purify the polarization and then a HWP to control it. Below, we show the configuration that we typically use in our BD setups. Additional, polarizing elements may be used for purifying the polarization or to control the energy of the probe beam.

![Diagram](image)

Figure 3-10. Experimental configuration used for (a) pump and (b) probe beams to control the energy or polarization. Half-wave plate (HWP), polarizer (POL), and quarter-wave plate (QWP).

The usefulness of controlling the polarization of each beam is evident for characterizing NLR of solvents. For small refractive index change, $\Delta n$ can be approximated as

$$\Delta n = \Delta n_\parallel \cos^2 \varphi + \Delta n_\perp \sin^2 \varphi,$$

(3-8)
where $\varphi$ is the angle between the field and the optic axis. The symmetry properties of different electronic and nuclear mechanisms contribute differently to the ultrafast NLR. In general, the third-order NLR of solvents (details in Ref. [113, 114]) can also be described as below,

$$\Delta n(\varphi) = \Delta n^{iso}(\varphi) + \Delta n^{re}(\varphi),$$  \hspace{1cm} (3-9)\

where superscripts iso and re denote to isotropic and re-orientational symmetries, respectively. There are various types of NLR mechanisms in solvents arising from the response of electrons that are isotropic (e.g. bound-electronic and collision induced response) or from the molecular orientation that have reorientational symmetry (e.g. libration and diffusive orientation) [113, 114]. These mechanisms contribute differently depending on the angle between the field and the optic axis. The combining equation governing on the angular dependent NLR can be derived as

$$\Delta n(\varphi) = \Delta n^{iso}_\parallel (\cos^2 \varphi + \frac{1}{3} \sin^2 \varphi) + \Delta n^{re}_\parallel (\cos^2 \varphi - \frac{1}{2} \sin^2 \varphi)$$  \hspace{1cm} (3-10)\

By substituting Eq. (3-10) in Eq. (3-8) and looking at $\varphi = 0^\circ$, $\varphi = 54.7^\circ$, and $\varphi = 90^\circ$ we get

$$\Delta n(0^\circ) = \Delta n^{iso}_\parallel + \Delta n^{re}_\parallel$$

$$\Delta n(54.7^\circ) = \frac{5}{9} \Delta n^{iso}_\parallel$$  \hspace{1cm} (3-11)\

$$\Delta n(90^\circ) = \frac{1}{3} \Delta n^{iso}_\parallel - \frac{1}{2} \Delta n^{re}_\parallel$$

The NLR at $\varphi = 54.7^\circ$, which is referred to as the “magic angle”, only depends upon the isotropic components contributions, and thus can be used to directly measure the bound-electronic and collision induced nonlinearities. The difference between the contribution of the circular and linear polarizations to the NLR also indicates the significance of measuring the polarization dependence.
of different mechanisms. By using the symmetry properties of the response function one can also derive that $\Delta n_{\text{circ}} = \frac{2}{3}\Delta n_{\parallel}^{(\text{iso})} + \frac{1}{4}\Delta n_{\parallel}^{(\text{re})}$. We used BD method to measure different polarization combinations of carbon disulfide (CS$_2$), which is a popular and well-studied material in NLO community. Figure 3-11 summarizes our measurements for parallel, perpendicular, magic angle, and circular polarization combinations versus delay.

![Figure 3-11](image.png)

Figure 3-11. BD measurements of CS$_2$ solvent for parallel (black), perpendicular (red), magic angle (blue), and circular polarization (magenta). The magenta data is for the case of circular pump beam and linear probe beam, which is similar for all polarization combination with at least one beam being circular.

Our measurements reveal the non-instantaneous nature of the optical nonlinearities in CS$_2$. The magic angle data is shown in blue, which exhibits an instantaneous NLR. This is expected from isotropic symmetry which only includes instantaneous electronic and collision induced nonlinearities. However, other combinations include both isotropic and reorientational symmetries, and thus follow Eq. (3-11) and the relation for circular polarization. Both isotropic and reorientational contributions can be found by knowing only two of these combinations, and thus one can reconstruct the other two mathematically. Since the NLO response of CS$_2$ is well known it can also be used to calibrate experimental setups such as BD and Z-scan.
To model the deflection of the probe more accurately, we can employ nonlinear wave equations. First we calculate the amplitude and phase of the electric field at the exit plane of the sample, $\tilde{E}_{\text{out}}(x, y, t, \tau)$ and then use Fresnel diffraction propagation to calculate the complex electric field on the detector at a distance $d$, $\tilde{E}_{\text{det}}(x_d, y_d, t, \tau)$ via

$$
\tilde{E}_{\text{det}}(x_d, y_d, t, \tau) = e^{ik_{0,p}d} e^{i\frac{k_{0,p}(x_d^2+y_d^2)}{2d}} \mathcal{F} \left[ \tilde{E}_{\text{out}}(x, y, t, \tau) e^{i\frac{k_{0,p}(x^2+y^2)}{2d}} \right],
$$

(3-12)

where $\mathcal{F}$ represents the two-dimensional spatial Fourier transform.

In this dissertation, we are primarily interested in $\tilde{E}_{\text{out}}$ from carrier-induced nonlinear mechanisms. For example, in CHAPTER 4, we use BD for slightly non-degenerate wavelengths to measure NLA and NLR of thick GaAs and Silicon bulk materials, where degenerate 2PA excites electrons to the conduction band and results in FCA and FCR. For those measurements, we employ discretized nonlinear wave propagation to account for accumulative process of exciting carriers via NLA. The details of our modeling will be discussed in more detail in CHAPTER 4.

3.3.3. Cross-propagating Beam-Deflection

As discussed earlier, various experimental techniques are devoted to measuring the NLO response of materials. However, the polarization combinations are often limited because of the interaction geometry. To fully exploit the third-order NLO properties of materials, we modified our beam deflection technique to study the elements of the third-order susceptibility tensor with field components along the propagation direction. In this modified method, as opposed to conventional BD technique, where pump and probe beams are nearly collinear, the interaction geometry is based on orthogonal propagation of pump and probe beams. The experimental configuration for both collinear BD (discussed in the previous section) and cross-propagating BD is shown below.
In the collinear geometry, the polarization of the pump and probe beams can only be modified in a plane perpendicular to the direction of propagation of both beams. However, in a cross-propagating geometry the BD method allows rotating the polarization of the probe to have the electric field polarized parallel to the wavevector of the pump beam (\( \vec{E}_{\text{probe}} \parallel \vec{k}_{\text{pump}} \)). This helps us to characterize additional non-diagonal susceptibility tensor elements in NLO materials as well as conventional parallel and perpendicular polarizations as in most other experimental methods. We note, however, that the interaction geometry is more complicated in the cross-propagation scheme. This is because the temporal dimension of one beam will cross-correlate the spatial dimension of the other beam, and thus a complex spatiotemporal interaction occurs. This is in contrast to the collinear BD method, where temporal and spatial dimensions are inherently independent of each other. This can be clarified by considering a probe beam with a pulsewidth of \( \tau_p \approx 100 \text{ fs} \) pulses, which corresponds to a physical extent of \( z_{\text{pulse}} \approx c \tau_p \approx 30 \mu m \). Now, for having the optimum
interaction in cross-propagation geometry, we need to have a spatial pump cross-section close to 30 $\mu m$. However, the size of the pump beam needs to be simultaneously large enough to have a reasonable linear index gradient that deflects the probe beam. This complicates the interaction geometry further, where a reasonably large pump beam is required, while the beam needs to be small enough to have an optimum overlap. We addressed these trade-offs by elliptically focusing the pump beam in one dimension to have a tightly focused beam and thus a large irradiance, while the other dimension is loosely focused to have a reasonable index gradient for BD measurements. The interaction geometry of this modified BD technique is shown below.

![Interaction geometry of the cross-propagation BD method.](image)

For cross-propagating geometries the relative polarizations of electric and magnetic fields can be independently controlled, as shown in Figure 3-14. The indices in the third-order susceptibility, $\chi^{(3)}$, are related to the orientation of the electric fields. It is well-known that for collinear geometry, the ratio of susceptibilities for parallel and cross-polarized beams is 3.
Figure 3-14. Polarization combinations possible in the cross-propagating beam deflection experiment.

The strong excitation beam propagates in the z-direction, with the weak probe propagating in the y-direction.

In Figure 3-15, we show the beam-deflection data, $\Delta E/E$, versus relative delay of pump and probe pulses for carbon tetrachloride (CCl₄) and Tin tetrachloride (SnCl₄). These two molecules both possess symmetric tetrahedral geometries, and thus exhibit isotropic nonlinear effects as opposed to CS₂, which also exhibits reorientational nonlinear response, as shown in the previous section. Clearly we see no difference between $\chi^{(3)}_{xxyy}$ and $\chi^{(3)}_{zzzz}$, which is expected for isotropic liquids. However the ratio of $\chi^{(3)}_{xxxx}$ and $\chi^{(3)}_{zzzz}$ is $4 \pm 0.2$, rather than the usual 3, which is expected for bound-electronic nonlinear refraction. This difference may be attributed (but not conclusively) to cascaded NLO effects induced by the charge separation along the propagation direction (parallel to the wavevector) of the pump beam and/or optical magnetization effects [132-135].
Figure 3-15: Cross propagating beam deflection results in CCl$_4$ and SnCl$_4$, for various polarization combinations.

The elliptical focusing helps improve the interaction geometry and enhance the signal because of tighter focusing; however, the cross-correlation of the two beams can also become narrower by controlling the wavefront tilts of the two beams. This limitation is apparent in Figure 3-15, where two pulses (pump and probe) with ~150 fs pulsewidths give us a cross-correlation width of ~2 ps, as measured via cross-propagation BD. The reason for this broadened correlation width is the complex spatiotemporal interaction geometry of the two beams.

In order to overcome some of the limitations in the interaction, we have employed tilted wavefronts for pump and probe beams independently to optimize the interaction in the nondegenerate cross-beam geometry. We use nondegenerate pump and probe beams to avoid complications that occur in degenerate experiments [136]. The benefit of this modification is the possibility to control the wavefront tilt to optimize the interaction geometry even when the two beams are nearly collinear. Additionally, if the two beams have tilted wavefronts, this modification can correct the tilt for a better interaction. In order to separately control the wavefront tilt in both excitation and probe beams, we have implemented two independent adjustments into our conventional cross-
beam geometry, Figure 3-16(a, b). The spatio-temporal interaction of two beams crossing each other depends upon the wavefront tilt of each individual beam. In a degenerate excite-probe geometry, the tilt can be adjusted properly before splitting the beam into pump and probe beams. However, for nondegenerate measurements, we need to separately control the wavefront tilt of each beam in order to optimize the interaction of the two beams crossing each other.

Figure 3-16. (a) Cross-propagating beam-deflection, (b) extended geometry employing additions to adjust the wavefront tilt separately for excitation and probe beams.

For instance, if both wavefronts are perpendicular to the direction of propagation, the interaction region will be a line oriented at 45°, which leads to a small interaction. However, if we adjust the wavefront tilts properly, we can optimize the interaction as shown in Figure 3-17(a). For example, if we tilt the wavefront for 20° in a correct direction, we can compress the cross correlation width compared to the case where there is no tilt in the wavefront, as shown in Figure 3-17(b) for normalized signals.
Figure 3-17. (a) The interaction geometry and effect of wavefront tilt in spatio-temporal interaction, the illustrated tilt is 45 degrees (b) experimental observation of the optimization of the cross-correlation width while wavefronts are tilted for 20 degrees for each beam compared to when they are perpendicular to the direction of propagation. The cross-correlation changes from ~ 6 ps (FWHM) to ~ 850 fs (FWHM). Note that the measurements were done for 20 degrees tilt for each beam.

Note that, the cross-correlation width is a function of both temporal and spatial parameters and compressing the width of the cross-correlation results in larger interaction signals. Here we have used a 45° tilted GaAs sample of 0.5 mm thickness to measure the nonlinear absorption of the probe induced by the pump beam via the 2PA process.

3.3.4. Temporal modulation of nonlinear phase

BD and Z-scan are capable of measuring NLA and NLR by measuring the nonlinearily induced changes in transmitted irradiance and induced nonlinear phase-changes. The nonlinear phase-change is typically referred to those changes in material that modify the phase, $\varphi(r, t)$, of the wave transmitted through the material in the spatial domain. Spatial modifications of the phase are indeed a direct consequence of spatial changes in refractive index, which subsequently alters the momentum of photons, $k$, and speed of light. In a nonlinear media, besides spatial modulation of the refractive index, the optical properties will also be manipulated in the temporal domain. The direct consequence of this temporal manipulation is induced changes in the energy of photon or the frequency, $\omega$. A simple equation explaining both phenomena is
\[ k = \frac{\partial}{\partial r} \varphi(r, t) \quad \& \quad \omega = -\frac{\partial}{\partial t} \varphi(r, t) \quad (3-13) \]

Nonlinear phase-change in temporal domain (sometimes referred to as time-refraction [137]) may change the frequency of the incident beam itself through self-phase modulation (SPM) or a second weak beam through cross-phase modulation (XPM). Since BD and Z-scan are limited to only measuring spatial nonlinear phase-changes, we employ an optical spectrum analyzer (OSA) in some of our experiments to simultaneously measure the central wavelength of the transmitted pulse while nonlinear phenomena are occurring in the material.
CHAPTER 4: BOUND-ELECTRONIC AND CARRIER-INDUCED NONLINEARITIES IN SEMICONDUCTORS

4.1. Introduction and Motivations

In the past decades, nonlinear optical processes in semiconductors has shown both fundamental research importance and technological implications in many NLO applications such as all-optical signal processing [11, 13, 138-143], optical limiting [43, 138, 143, 144], nonlinear integrated photonic devices [13, 139, 145-147], multiphoton processes [45, 48, 148-152], two-photon lasing and detection [50, 145, 153], and many other avenues of nonlinear optics that can be potentially used in NLO devices. There are numerous studies on characterizing the strength of third-order and fifth-order nonlinear effects in semiconductors; however, values reported in the literature typically exhibit significant spread by order of magnitude differences in their magnitude. This discrepancy can be attributed to many factors such as modeling, estimations, proper data analysis, difficulties in NLO characterization measurements including knowledge of the spatial and temporal profile of the interacting beam, sample-to-sample variations, anisotropy of the response, etc. As an example, the values of the third-order nonlinearities of Silicon and GaAs reported within the past three decades are listed in Table 4-1 and Table 4-2, respectively.

Table 4-1. Literature values of 2PA and third-order Kerr nonlinearity of GaAs around 1550 nm

<table>
<thead>
<tr>
<th>(\alpha_2) (cm/GW)</th>
<th>(n_2) (\times 10^{-5}) cm²/GW</th>
<th>Method</th>
<th>Wavelength</th>
<th>Wavelength</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>–</td>
<td>Pump-probe</td>
<td>1550</td>
<td>[154]</td>
</tr>
<tr>
<td>13</td>
<td>–</td>
<td>Pump-probe</td>
<td>1545</td>
<td>[155]</td>
</tr>
<tr>
<td>10.2</td>
<td>15.9</td>
<td>Z-scan</td>
<td>1540</td>
<td>[156]</td>
</tr>
<tr>
<td>2.5</td>
<td>30</td>
<td>Z-scan</td>
<td>1680</td>
<td>[150]</td>
</tr>
<tr>
<td>5.5</td>
<td>80</td>
<td>SPM*</td>
<td>1500</td>
<td>[157]</td>
</tr>
</tbody>
</table>
*SPM: Self-Phase Modulation

<table>
<thead>
<tr>
<th>$\alpha_2$ (cm/GW)</th>
<th>$n_2$ ($\times 10^{-5}$ cm$^2$/GW)</th>
<th>Method</th>
<th>Wavelength</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.48</td>
<td>2.4</td>
<td>Z-scan</td>
<td>1500</td>
<td>[158]</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>SPM</td>
<td>1559</td>
<td>[159]</td>
</tr>
<tr>
<td>0.79</td>
<td>4.5</td>
<td>Z-scan</td>
<td>1540</td>
<td>[156]</td>
</tr>
<tr>
<td>0.75</td>
<td>4.8</td>
<td>Z-scan</td>
<td>1550</td>
<td>[160]</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>SPM</td>
<td>1500</td>
<td>[161]</td>
</tr>
<tr>
<td>0.45</td>
<td>6.0</td>
<td>NLT*, SPM</td>
<td>1540</td>
<td>[162]</td>
</tr>
<tr>
<td></td>
<td>6.5</td>
<td>Z-scan</td>
<td>1550</td>
<td>[163]</td>
</tr>
<tr>
<td>1.03</td>
<td>6.7</td>
<td>Z-scan</td>
<td>1550</td>
<td>[163]</td>
</tr>
<tr>
<td>0.9</td>
<td>7.0</td>
<td>NLT, SPM</td>
<td>1530</td>
<td>[164]</td>
</tr>
<tr>
<td></td>
<td>9.0</td>
<td>FWM**</td>
<td>1547</td>
<td>[165]</td>
</tr>
<tr>
<td>0.6</td>
<td>14.5</td>
<td>NLT, SPM</td>
<td>1550</td>
<td>[166]</td>
</tr>
<tr>
<td>0.44</td>
<td>–</td>
<td>NLT</td>
<td>1560</td>
<td>[167]</td>
</tr>
<tr>
<td>0.5</td>
<td>–</td>
<td>NLT</td>
<td>1455</td>
<td>[168]</td>
</tr>
<tr>
<td>0.67</td>
<td>–</td>
<td>NLT</td>
<td>1547</td>
<td>[169]</td>
</tr>
</tbody>
</table>

*NLT: Nonlinear Transmission
**FWM: Four-Wave Mixing

There has also been a considerable interest in fifth-order semiconductor NLO effects such as 2PA-induced FCA and FCR, and 3PA. In particular, in organic optoelectronics, 3PA can be used as an excitation to observe efficient stimulated emission and frequency upconversion fluorescence emission [43-45]. Additionally, 3PA in semiconductor quantum dots and nanocrystals has attracted major interest for applications in bio-labeling and imaging agents due to the possibility of using longer excitation wavelengths to achieve deeper penetration depths for super-resolution imaging [42, 170-173].
The main conclusion by briefly examining the previous studies proves that accurate modeling and experimental verification of nonlinear mechanisms in semiconductors are still needed to address the discrepancies reported in the literature as listed in the above tables. We, therefore, employed our BD technique to simultaneously measure NLA and NLR of direct-gap semiconductors such as GaAs, which has myriad of applications in NLO.

The nonlinear response of semiconductors depends upon the electronic band structure and transition dipole moments between them. Thus, to accurately determine the magnitude, sign, and the dispersion of the nonlinear response, it is essential to have a knowledge of the energy levels. We briefly review the fundamentals of third-order nonlinear response of semiconductors that includes 2PA and instantaneous bound-electronic Kerr effect. Note that carrier nonlinearities that effectively scale similar to third-order processes, where they are generated via linear absorption, and subsequently either absorbed or refracted can also be included in the category of “effectively” third-order nonlinearities. We also review instantaneous fifth-order processes in semiconductors that includes 3PA that modifies the transmission via a quadratic dependence on the irradiance. We do not discuss fifth-order nonlinear refraction because of difficulties that are present in experimental characterization while they are smaller in magnitude compared to the third-order NLR. There are also carrier nonlinearities that are scaling similar to fifth-order processes as well. These carriers are typically generated via 2PA, and subsequently absorbed via a 1PA process or contribute to the NLR, which will be discussed too.

4.2. Third-order Bound-Electronic Nonlinearities of Direct-Gap Semiconductors

We briefly describe the relationship between the third-order bound-electronic NLR and 2PA and the quantum mechanical wave functions of the material developed by Sheik-Bahae [36, 174]. The main approach to obtain the dispersion of the third-order bound-electronic process that includes
two photons is to first calculate the nonlinearly induced change in the absorption, $\Delta \alpha$, for the entire spectrum, and subsequently use the Kramers-Kröning relation to determine the dispersion of the bound-electronic Kerr effect, $\Delta n$, [37, 174-176].

To calculate the dispersion of nonlinear absorption, first-order perturbation theory with a dressed state for the two bands (valence and conduction band) can be used. In this method, the tunneling effect, which contributes to the acceleration of the electrons due to the oscillating electric field, is included [174]. To formulate the interaction Hamiltonian, the dipole approximation described by the $A \cdot \mathbf{p}$ method can be used, which can be described as

$$H_{int} = -\frac{e}{m_0} A \cdot \mathbf{p}, \quad (4-1)$$

where $A$ is the magnetic vector potential, $\mathbf{p}$ is the electron momentum, $e$ is the electron charge, and $m_0$ is the electron mass. For a linearly polarized beam

$$A = \hat{\mathbf{a}} A_0 \cos(\omega t), \quad (4-2)$$

where $\hat{\mathbf{a}}$ is the polarization unit vector, $A_0$ is the vector field amplitude, and $\omega$ is the frequency of the incident beam. In this formulation, a dressed state for both valence and conduction bands is used that can be approximated as [177-180]

$$\psi_{v,c}(\mathbf{k}, \mathbf{r}, t) = u_{v,c}(\mathbf{k}, \mathbf{r}) \exp \left\{ i \mathbf{k} \cdot \mathbf{r} - i \hbar^{-1} \int_0^t E_i(\tau) d\tau \right\}, \quad (4-3)$$

where $u_{v,c}$ are the Bloch wavefunctions. Upon applying an electric field, the energy of the electrons and holes in the initial and final states will be manipulated. The transition rates can be calculated using the $S$ matrix formalism, which is described as
By substituting the wave functions and the interaction Hamiltonian into the above equation and performing time integral, the final relation for $S$ will be found. The transition rates can be then written based on the $S$-matrix by calculating

$$W = \lim_{t \to \infty} \int \frac{d^3k}{(2\pi)^3} \frac{d}{dt} |S|^2,$$  \hspace{1cm} (4-5)$$

which for the degenerate case the results will be

$$\Delta W = \sum_{\text{spin}} \int \frac{d^3k}{(2\pi)^3} \left\{ \frac{\pi e^2 A_0^2}{2 m_0 m_{vc} c^2 \omega^2} \right\}^2 |\hat{a} \cdot p_{vc}|^2$$

$$\cdot |k \cdot \hat{a}|^2 \frac{1}{2\pi \hbar} \delta \left(E_g + \frac{\hbar^2 k^2}{2m_{vc}} - 2\hbar \omega\right).$$  \hspace{1cm} (4-6)$$

Now to find the 2PA coefficient, we first substitute $A_0^2 = 8\pi c I_0 / n \omega^2$, where $I_0$ is the irradiance and $n$ is the linear index of the material at frequency $\omega$. The 2PA coefficient, $\alpha_2(\omega) = 2\hbar \omega \Delta W I_1^{-2}$, can be written as

$$\alpha_2(\omega) = K \sqrt{\frac{E_p}{n_0^2 E_g}} F_2 \left( \frac{\hbar \omega}{E_g} \right),$$  \hspace{1cm} (4-7)$$

where $E_p = 2|p_{vc}|^2 / m_0 \approx E_g / m_c$ [36, 174] and

$$F_2(x) = \frac{(2x - 1)^{3/2}}{(2x)^5},$$  \hspace{1cm} (4-8)$$

while $2x - 1 > 0$. This function effectively considers the intermediate states assumed in calculating the 2PA transition. The Kane energy, $E_p$, is also considered independent of the material and
possesses the value of \( E_p \approx 21 \text{ eV} \) for most direct gap semiconductors [181]. The value of \( K \) is also material-independent and only depends upon electron effective mass and charge. There are debates on the value of \( K \); however, note that the value of \( K \) only scales the magnitude of the results and does not modify the spectral dependence of 2PA. Previous results for direct gap semiconductors led to the value \( K = 3100 \text{ eV}^{-5/2} \text{ cm/GW} \) while averaging for various materials [37]. Second-order perturbation calculations from Weiler gives \( K = 5200 \text{ eV}^{-5/2} \text{ cm/GW} \) for four-band model [182]. As we will show later, our measurements with two different techniques together with measurements done in another facility were nicely fit with the predicted spectral dependence while we considered \( K \approx 4600 \text{ eV}^{-5/2} \text{ cm/GW} \) for a single material.

The NLR can also be calculated via the Kramers-Krönig relations, which will result in

\[
n_2(\omega) = K \frac{\hbar c \sqrt{E_p}}{n_2 E_g^4} G_2\left(\frac{\hbar \omega}{E_g}\right),
\]

(4-9)

Where the dispersion function \( G_2 \) is given by

\[
G_2\left(\frac{\hbar \omega}{E_g}\right) = \frac{2}{\pi} \int_0^\infty \frac{F_2(x_1; x_2)}{x_1^2 - x_2^2} dx_1,
\]

(4-10)

where \( F_2(x_1; x_2) \) is calculated for nondegenerate photons with \( x_1 = \hbar \omega_1 / E_g \) and \( x_2 = \hbar \omega_2 / E_g \) [36, 174, 175]d. It is important to note that for calculating the Kramers-Krönig transformation, the nondegenerate spectra for 2PA is needed. We use this theory to plot the frequency dependence of the degenerate third-order nonlinearities of GaAs considering \( K = 4600 \text{ eV}^{-5/2} \text{ cm/GW} \). For other direct gap semiconductors, while considering two parabolic band model, the frequency dependence of the third-order nonlinear response will be similar to GaAs with different magnitude and spectral range.
In addition to the bound-electronic nonlinear processes, excited carriers can also act as they are free (with reduced mass depending on the band structure), and create electron (or hole) plasma. Since these carriers are generated via absorption (either 1PA or 2PA), the absorption or refraction of these carriers are inherently a nonlinear mechanism. The free-carrier nonlinearities are cumulative with a decay time given by the carrier lifetime. The contribution of free-carriers becomes dominant for longer pulsewidths, where carriers absorb light within the pulse, and are least important for short pulsewidths and low irradiances, where less number of carriers are generated. As discussed in CHAPTER 2, FCA and FCR modifies the irradiance, $I_p$, and phase, $\varphi_p$, of the probe, in presence of bound-electronic nonlinearities, via

\[
\frac{d}{dz} I_p = -2\alpha_{2,p}^{ND} I_e I_p - (\sigma_{FCA,el} N_e + \sigma_{FCA,hl} N_h) I_p, \tag{4-11}
\]

\[
\frac{d}{dz} \varphi_p = -2k_0 n_{2,p}^{ND} I_e + (\sigma_{FCR,el} N_e + \sigma_{FCR,hl} N_h), \tag{4-12}
\]
where $\alpha_{2,p}^{ND}$ and $\beta_{2,p}^{ND}$ correspond to bound-electronic nondegenerate nonlinearities, $I_e$ is the irradiance of the pump (excitation) beam. Additionally, $N_e$ and $N_h$ are the density of excited electrons and holes, respectively, and $\sigma_{FCA,i}$ and $\sigma_{FCR,i}$ are the FCA and FCR cross-sections associated with electrons and holes $(i = e, h)$. In two-parabolic band model, we assume $N_e = N_h = N_{ex}$, $\sigma_{FCA,e} + \sigma_{FCA,h} = \sigma_{FCA}$, and $\sigma_{FCR,e} + \sigma_{FCR,h} = \sigma_{FCR}$. The equation governing on the total density of excited carriers can also be written as

$$\frac{d}{dt} N_{ex} = \frac{\alpha_2^D I_e^2}{2\hbar\omega_e} - \frac{N_{ex}}{\tau_{ex}},$$

where $\alpha_2^D$ is the degenerate 2PA coefficient of the pump beam with photon energy of $\hbar\omega_e$ and $\tau_{ex}$ is the decay constant of the excited carriers. To understand the spectral dependence of FCA and FCR cross-section, we can employ the Drude model (discussed in CHAPTER 2); however, excited carriers in semiconductors do not exhibit identical behavior as the Drude model would predict, mainly because Drude model is accurate for a free-electrons while not considering other screening effects, as opposed to semiconductors where an energy band needs to be considered. Additionally, for non-resonant photon energies, the conservation of momentum via phonon-assisted transitions needs to be considered. The contribution of these processes cannot be properly analyzed by simple Drude model for free-electrons.

**4.4. Measurements of the Instantaneous Third-order and Non-Instantaneous Fifth-Order Nonlinearities in GaAs and Silicon**

In order to accurately characterize the third-order bound electronic nonlinearities and fifth-order carrier-induced nonlinearities of GaAs, we and our collaborators have performed series of measurements by employing Z-scan and BD with different laser systems and at two different institutions.
(CREOL and Georgia Institute of Technology). We performed these measurements on a single piece of GaAs (500 μm) thick. The measurements were conducted independently to account for any systematic errors. We note that both Z-scan and BD have their own advantages, but they can be considered complementary in the sense that they can each provide a unique set of information [116]. While Z-scan is a simple methodology to measure degenerate NLA and NLR of materials, it lacks information on the transient NLO response. Therefore, separating instantaneous bound-electronic process from non-instantaneous carrier nonlinearities can become problematic, particularly when the magnitude of one of them is much smaller than the other one. To accurately distinguish these two effects, higher energy measurements are required to investigate the scaling of the detected signal versus irradiance. Similarly, when the magnitude of NLR is very small compared to NLA, separating them can be difficult in Z-scan measurements. This separation is possible BD, which inherently decouples NLR from NLA. As an example, in Figure 4-2, we show OA and CA/OA Z-scan measurement results on a 500 μm thick GaAs sample.
For these measurements, the beam size at the focus is $w_0 = 36.5 \mu m$ (HW1/e$^2$M) and the pulse-width is $\tau = 186$ fs (FWHM). The experiments were fit numerically and values for bound-electronic and carrier nonlinearity are found and reported below.

Table 4-3. Summary of the results obtained from Z-scan measurements at 1300 nm on GaAs

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PA</td>
<td>$\alpha_2$(cm/GW)</td>
<td>30</td>
</tr>
<tr>
<td>FCA</td>
<td>$\sigma_{FCA}$ (m$^2$)</td>
<td>(0.5 to 30) $\times$ 10$^{-21}$</td>
</tr>
<tr>
<td>Bound-electronic Kerr effect</td>
<td>$n_2$(cm$^2$/GW)</td>
<td>6 $\times$ 10$^{-5}$</td>
</tr>
<tr>
<td>FCR</td>
<td>$\sigma_{FCR}$ (m$^3$)</td>
<td>($-15$ to $-30$) $\times$ 10$^{-21}$</td>
</tr>
</tbody>
</table>

As shown in Table 4-3, it is very difficult to obtain accurate values for $\sigma_{FCA}$ and $\sigma_{FCR}$. Additionally, the magnitude of the bound-electronic NLR is very small (as expected from the theory and calculations reported in Figure 4-1). The small magnitude of NLR results in an increase in the CA/OA signal, which is mostly due to the drop in the transmission obtained in OA, and the contribution of NLR becomes negligible unless the irradiance is high enough that FCR becomes significant. We, therefore, used BD to measure these parameters for GaAs at 1300 nm and 1550 nm.
As discussed earlier, BD allows us to simultaneously measure both transient NLR and NLA. It is important to note that the deflection signal originates from the spatial energy redistribution. In BD, as opposed to the NLR that generates an index gradient, NLA will introduce an attenuation gradient, which is due to having a larger absorption near the center and less at the wings. This attenuation gradient reshapes the probe beam in a way that translates the centroid of the beam towards the wing. Therefore, the contribution of the NLA looks like a negative NLR; however, since the deflection is originated laterally, it will be dependent of the propagation distance. For NLR, there will be an angular deflection mechanism that makes the detected signal independent of the propagation distance (sample to detector). This means that at the far-field the contribution of the NLA will be eliminated, and thus it will be possible to detect a signal that is purely from NLR mechanisms. Additionally, the instantaneous bound-electronic and non-instantaneous carrier effects can be effectively separated in temporal domain, thus making it much simpler to simultaneously characterize each effect.

For BD measurements, we first performed calibration measurements on a piece of fused silica sample (1 mm thick) and a cuvette of CS₂ (1 mm liquid), which the nonlinear response of both materials are well-known. We performed these measurements to scale our results and to eliminate the effect of any systematic errors originated from irradiance characterization. Typically, a precise knowledge of the beam profile and pulse duration is required for absolute characterization of nonlinear parameters; however, it is very difficult to have this knowledge particularly for high power lasers. Additionally, performing measurements for different wavelengths or at different laboratories can increase these errors, therefore, it is needed to perform the final measurements relative to well-characterized materials.
We performed two sets of BD measurements: 1) $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm, 2) $\lambda_e = 1350$ nm and $\lambda_p = 1250$ nm, where $\lambda_e$ and $\lambda_p$ are wavelengths of the pump and probe beams, respectively. We chose slightly nondegenerate wavelengths to avoid coherent artifacts of degenerate experiments [136]. At each set of measurements, we first characterized the beam characteristics of the both beams via knife-edge scans and also measured the pulsewidth of the pump beam. As an example, we show the results of these measurements for $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm in Figure 4-3. We measured autocorrelation width of 335 fs, which corresponds to $\tau = 236$ fs (FWHM) width for the 1600 nm pulse. We also measured the beam size of the probe and pump beams to find the $M^2$ parameter of the probe, which for our beam is $M^2 \approx 1.2$. The ratio of the beam size for pump, $w_e$, to that of the probe, $w_p$, is close to 5, which is appropriate for BD measurements. We need to know this ratio for analyzing the signal obtained from BD measurements. We note, however, that as long as this ratio is large enough (typically more than 3), we can use our relative measurements to scale our results accordingly.

![Figure 4-3](image_url)

Figure 4-3. (a) Autocorrelation measurements for characterizing the pulsewidth of the pump pulse, (b) knife-edge scan measurements to characterize the beam profile (horizontal and vertical cross-section) of the probe beam and measuring the beam size of the pump beam.

As mentioned earlier, we also conducted measurements on fused silica sample and CS$_2$ solvent. These measurements were used as a reference to calibrate our experiment. We fit our fused silica
measurements with \( n_2 = 0.25 \times 10^{-19} \text{ cm}^2/\text{W} \) at parallel polarization, which is in an excellent agreement with previously reported values [104, 183].

![BD measurements on a 1mm thick fused silica](image)

Figure 4-4. (a) BD measurements on a 1 mm thick fused silica and fit values of \( n_2 = 0.25 \times 10^{-19} \text{ cm}^2/\text{W} \) for parallel polarization and (b) BD measurements on 1 mm thick CS\(_2\) liquid after subtracting the contribution of the cuvette for three different polarization combinations.

The measurements for CS\(_2\) were fit using the response function for molecular third-order nonlinearities developed and reported in [21, 113, 114] and discussed briefly in previously in CHAPTER 3. Summary of our results are shown in Table 4-4, which are in agreement with values reported in [113, 114].

Table 4-4. Summary of the results obtained from BD measurements on CS\(_2\) for \( \lambda_e = 1600 \text{ nm} \) and \( \lambda_p = 1500 \text{ nm} \), while considering 20\% error.

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>( n_{2,m} (\times 10^{-19} \text{ m}^2/\text{W}) )</th>
<th>( \tau_{r,m} ) (fs)</th>
<th>( \tau_{f,m} ) (fs)</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic</td>
<td>2.3</td>
<td>Instantaneous</td>
<td></td>
<td>Isotropic</td>
</tr>
<tr>
<td>Collision</td>
<td>1.3</td>
<td>200</td>
<td>140</td>
<td>Isotropic</td>
</tr>
<tr>
<td>Libration</td>
<td>8.8</td>
<td>( \omega_0 = 8.5 \text{ ps}^{-1} )</td>
<td>420</td>
<td>Reorientational</td>
</tr>
<tr>
<td>Diffusion</td>
<td>22</td>
<td>200</td>
<td>1500</td>
<td>Reorientational</td>
</tr>
</tbody>
</table>

After conducting calibration measurements, experiments were performed to measure NLA and NLR of a GaAs sample by using the BD technique. BD allows us to simultaneously measure both
induced NLA and NLR. The working principle of this method was explained in detail in Chapter 3. To briefly review it, the deflected probe beam was Gaussian propagated to the detector and integrated over the $x - y$ plane to calculate the difference in probe power incident on the left and right sides of the quad segmented detector. Next, we integrate over time to determine the normalized signal on the detector, $\Delta E/E$, which directly gives the nonlinear phase change, and subsequently the NLR. Additionally, if we perform the time integral over the total transmitted power of the probe beam reaching the four quadrant, we will get the normalized total transmission, $\Delta T/T$, which gives us the NLA similar to the transient absorption spectroscopy methodology.

Measurements of normalized transmission and $\Delta E/E$ versus delay for $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm are shown below.

![Graphs showing measurements](image)

Figure 4-5. (a) Nonlinear transmission and (b) BD measurements performed on a piece of GaAs for $\lambda_e = 1600$ nm and $\lambda_p = 1500$ nm.

We performed similar measurements on GaAs for $\lambda_e = 1350$ nm and $\lambda_p = 1250$ nm and plotted the results. Measurements were done for three different pulse energies listed in the plots presented below.
Figure 4-6. (a) Nonlinear transmission and (b) BD measurements performed on a piece of GaAs for $\lambda_e = 1350$ nm and $\lambda_p = 1250$ nm.

Note that NLA and NLR will influence the irradiance and phase of the probe calculated right after the sample. Therefore, to analyze our experimental results, we use the nonlinear wave equations (presented in CHAPTER 3) to determine irradiance and phase of the probe beam at the exit plane of the sample. The governing equations on the irradiance and phase can be written as

$$\frac{d}{dz} I_e(z, t) = -\alpha_{2e}^D I_e^2(z, t),$$  \hspace{1cm} (4-14)

$$\frac{d}{dz} I_p(z, t) = -2\alpha_2^{ND} I_p(z, t) I_e(z, t) - \sigma_{FCA} N_{ex}(z, t) I_e(z, t),$$  \hspace{1cm} (4-15)

$$\frac{d}{dz} \varphi_p(z, t) = 2k_0 n_2^{ND} I_e(z, t) + \sigma_{FCR} N_{ex}(z, t),$$  \hspace{1cm} (4-16)

$$\frac{d}{dt} N_{ex}(z, t) = \frac{\alpha_{2e}^D I_e^2(z, t)}{2\hbar\omega} - \frac{N_{ex}(z, t)}{\tau_{ex}},$$  \hspace{1cm} (4-17)

where $I_p$ and $\varphi_p$ are the irradiance and phase of the probe beam after the sample, $N_{ex}$ is the density of excited carriers via 2PA, $\tau_{ex}$ is the decay time associated to carriers to relax to their initial states,
\( \alpha_{2}^{ND} \) and \( n_{2}^{ND} \) are the nondegenerate 2PA and electronic NLR coefficient, \( \alpha_{2}^{D} \) is the degenerate 2PA coefficient of the pump beam, \( I_{e} \) is the irradiance of the pump (excitation) beam, \( \sigma_{FCA} \) and \( \sigma_{FCR} \) are the free-carrier absorption and refraction cross-sections, respectively. In above equations, \( z \) and \( t \) are the spatial and temporal coordinates, respectively. Eq. (4-14) quantifies the magnitude of irradiance of the pump beam, which is modified via the degenerate 2PA of the pump within the nonlinear medium. For relatively low pump irradiances, this modification can be ignored while considering an undepleted pump beam within the material. Eq. (4-17) characterizes the density of excited carriers, which are generated via degenerate 2PA of the pump beam, which is quantified by \( \alpha_{2,e}^{D} \). The density of carriers can decrease via recombination mechanisms with a decay time of \( \tau_{ex} \). Eqs. (4-15) and (4-16) also characterizes the NLA and NLR imposed upon the probe beam, respectively. Our goal is to find values of \( \alpha_{2}^{ND}, n_{2}^{ND}, \sigma_{FCA}, \) and \( \sigma_{FCR} \) by fitting our experimental measurements. We simplify our fitting procedure by considering that \( \alpha_{2}^{ND} = \alpha_{2,e}^{D} \), which is only true in this case where the pump and probe pulses are nearly degenerate. Otherwise, first \( \alpha_{2,e}^{D} \) of the pump needs to be measured by degenerate techniques such as Z-scan, and then \( \alpha_{2}^{ND} \) can be determined from nondegenerate measurements.

To accurately calculate the irradiance and phase of the probe beam at the exit plane of the sample, we discretized Eqs. (4-15) to (4-17) in temporal and spatial domains. For this, consider that the thickness of the material is divided into \( N_{z} \) segments each having a thickness of \( \Delta z \) and the domain associated with the delay between to pulses is divided into \( N_{\tau} \) segments each having a duration of \( \Delta \tau \). We first determine the density of excited carriers at each time delay via

\[
N_{ex}(z, \tau + \Delta \tau) = N_{ex}(z, \tau) + \left( \frac{\alpha_{2,e}^{D} I_{e}^{2}(z, \tau)}{2 \hbar \omega} - \frac{N_{ex}(z, \tau)}{\tau_{ex}} \right) \cdot \Delta \tau, \quad (4-18)
\]
We, therefore, use the vector of $N_{ex}$ to calculate irradiance and phase at each spatial coordinate by using

$$I_e(z + \Delta z, \tau) - I_e(z, t) = \alpha_{z,e}^2 I_e^2(z, t) \cdot \Delta z,$$

(4-19)

$$I_p(z + \Delta z, \tau) = I_p(z, \tau) - \{2\alpha_2^{ND} I_p(z, t)I_e(z, t) + \sigma_{FCA} N_{ex}(z, t)I_e(z, t)\} \cdot \Delta z,$$

(4-20)

$$\varphi_p(z + \Delta z, \tau) = \varphi_p(z, \tau) - \{2k_0 n_2^{ND} I_e(z, t) + \sigma_{FCA} N_{ex}(z, t)\} \cdot \Delta z.$$

(4-21)

Finally, the complex magnitude of the probe field at the output plane of the sample, $\mathcal{E}_{p, exit}$, with a thickness of $L$ and at a given delay $\tau$ can be determined via

$$\mathcal{E}_{p, exit}(L, \tau) = \sqrt{\frac{2I_p(L, \tau)(1 - R_{exit})}{c\varepsilon_0 n_p}} \exp(i\varphi_p(L, \tau)),$$

(4-22)

where $R_{exit}$ is the Fresnel reflection coefficient of the intensity calculated for the exit plane of the sample. Now, we can use Eq. (3-12) to propagate the probe field to the detector and calculated $\Delta E/E$ similar to what has been discussed in Ref. [104].

We used this analysis to fit our experimental measurements, shown in Figure 4-5 and Figure 4-6, to determine NLA and NLR parameters of GaAs at 1300 nm and 1550 nm. The fitting procedure is based on the fact that BD effectively decouples the instantaneous nonlinearities from carrier related nonlinearities in temporal domain. We first fit the data of normalized transmission at longer delay times by finding the best fit for $\sigma_{FCA}$. The second step is to use $\sigma_{FCA}$ and find $\alpha_2^{ND}$ by fitting the signal around the zero-delay, which gives us the instantaneous NLA. These two steps are necessary in order to have a complete information on pump and probe irradiances and also the
density of excited carriers. Finally, we fit the BD signal at longer delays, which is directly related to FCR. The magnitude of this signal depends upon $\alpha_{z,e}^D \times \sigma_{FCR}$ that can be found by substituting $N_{ex}$ from Eq. (4-17) into Eq. (4-16). We note that the sign of $\sigma_{FCR}$ is negative, and therefore, decreases the effective NLR, and thus the nonlinear phase change. This can be seen from our measurements as well. At this stage, the only missing parameter is bound-electronic third-order NLR, which can be easily found by fitting our analysis to the BD signal close to the zero-delay. As expected, third-order NLA and NLR gives us instantaneous effects that follow the cross-correlation of the Gaussian temporal profile of the pump and probe beams. This is not the case for FCA and FCR, which for semiconductors have relatively long decay times and contribute to NLA and NLR as a constant signal in the temporal range of our measurements.

Table 4-5. Extracted values for instantaneous third-order NLA and NLR of GaAs, estimated error is 20%.

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>$\lambda_e = 1350$ nm &amp; $\lambda_p = 1250$ nm</th>
<th>$\lambda_e = 1600$ nm &amp; $\lambda_p = 1550$ nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{z,e}^{ND}$ (cm/GW)</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>$n_{ND}^{2}$ (cm/GW$^2$)</td>
<td>$5 \times 10^{-5}$</td>
<td>$14 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma_{FCA}$ (m$^2$)</td>
<td>$20 \times 10^{-21}$</td>
<td>$24 \times 10^{-21}$</td>
</tr>
<tr>
<td>$\sigma_{FCR}$ (m$^2$)</td>
<td>$-15 \times 10^{-21}$</td>
<td>$-26 \times 10^{-21}$</td>
</tr>
</tbody>
</table>

We have performed degenerate Z-scan measurements at 1300 nm and 1550 nm to characterize the NLA and NLR of GaAs with multiple laser systems and at two different institutions. As we mentioned earlier, we have performed these measurements to compare our results and to establish a self-consistent reference for the nonlinear mechanisms contributing the NLA and NLR of GaAs at telecommunication wavelengths. The summary of the results for 1300 nm (and slight nondegenerate measurements at 1300 nm) are shown in Table 4-6. The values for $\sigma_{FCA}$ and $\sigma_{FCR}$ are left empty because of the difficulties in determining these values for Z-scan measurements.
Table 4-6. Values of NLA and NLR of GaAs measured by Z-scan and BD at 1300 nm.

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>GT* − Z-scan</th>
<th>CREOL − Z-scan**</th>
<th>CREOL − BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2^{ND}$ (cm/GW)</td>
<td>23</td>
<td>30</td>
<td>16</td>
</tr>
<tr>
<td>$n_2^{ND}$ (cm/GW²)</td>
<td>$3 \times 10^{-5}$</td>
<td>$6 \times 10^{-5}$</td>
<td>$5 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma_{FCA}$ (m²)</td>
<td>−</td>
<td>$(0.5 \text{ to } 30) \times 10^{-21}$</td>
<td>$20 \times 10^{-21}$</td>
</tr>
<tr>
<td>$\sigma_{FCR}$ (m²)</td>
<td>−</td>
<td>$(-15 \text{ to } -30) \times 10^{-21}$</td>
<td>$-15 \times 10^{-21}$</td>
</tr>
</tbody>
</table>

*Georgia Institute of Technology (Courtesy of Dr. Joel Hales).
**Courtesy of Dr. Natalia Munera.

Similarly, the summary of the results for 1500 nm (and slightly nondegenerate measurements at 1550 nm) are shown in Table 4-7. The values for $\sigma_{FCA}$ and $\sigma_{FCR}$ in the columns related to Z-scan measurements are left empty mainly due to the difficulties in determining these values for Z-scan measurements. We attempted to fit the experimental data of FCA and FCR measured by Z-scan at our institution, however, the errors on these measurements were large. BD can address these difficulties by separating the instantaneous and non-instantaneous signals in the temporal domain.

Table 4-7. Values of NLA and NLR of GaAs measured by Z-scan and BD at 1550 nm.

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>GT* − Z-scan</th>
<th>CREOL − Z-scan**</th>
<th>CREOL − BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2^{ND}$ (cm/GW)</td>
<td>17</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>$n_2^{ND}$ (cm/GW²)</td>
<td>$20 \times 10^{-5}$</td>
<td>$18 \times 10^{-5}$</td>
<td>$14 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma_{FCA}$ (m²)</td>
<td>−</td>
<td>−</td>
<td>$24 \times 10^{-21}$</td>
</tr>
<tr>
<td>$\sigma_{FCR}$ (m²)</td>
<td>−</td>
<td>−</td>
<td>$-26 \times 10^{-21}$</td>
</tr>
</tbody>
</table>

*Georgia Institute of Technology (Courtesy of Dr. Joel M. Hales).
**Courtesy of Dr. Natalia Munera.

We summarized all of these measurements and compared our results with the frequency dependence of instantaneous bound-electronic third-order NLA and NLR.
In addition to our measurements for GaAs, we also performed similar measurements on a 400 μm thick Silicon. The theory that we presented in previous sections are not applicable to Silicon as this material is an indirect gap semiconductor. Therefore, phonon assisted transitions that conserve the momentum needs to be considered as well. However, because of the interest in NLO community in this material we attempted to perform NLO measurements to report the associated NLA and NLR coefficients. Below, we show the normalized transmission and BD measurements of Silicon for similar pump and probe conditions as mentioned above.
Figure 4-8. (a) Nonlinear transmission and (b) BD measurements performed on a piece of Silicon for \( \lambda_e = 1350 \text{ nm} \) and \( \lambda_p = 1250 \text{ nm} \). (c) Nonlinear transmission and (d) BD measurements performed on a piece of GaAs for \( \lambda_e = 1600 \text{ nm} \) and \( \lambda_p = 1500 \text{ nm} \).

We also fit our experiments with the same procedure used for GaAs and compared our results with degenerate Z-scan measurements performed in two different laboratories. The summary of the results for 1300 nm are reported in Table 4-8.

Table 4-8. Values of NLA and NLR of Silicon measured by Z-scan and BD at 1300 nm.

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>GT* – Z-scan</th>
<th>CREOL – Z-scan**</th>
<th>CREOL – BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha^N_D ) (cm/GW)</td>
<td>1.1</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>( n^N_D^2 ) (cm/GW²)</td>
<td>( 4 \times 10^{-5} )</td>
<td>( 5 \times 10^{-5} )</td>
<td>( 5 \times 10^{-5} )</td>
</tr>
<tr>
<td>( \sigma_{FCA} ) (m²)</td>
<td>( 1 \times 10^{-21} )</td>
<td>( 0.5 \times 10^{-21} )</td>
<td>( 1 \times 10^{-21} )</td>
</tr>
<tr>
<td>( \sigma_{FCR} ) (m²)</td>
<td>( -8 \times 10^{-21} )</td>
<td>( -5 \times 10^{-21} )</td>
<td>( -6.2 \times 10^{-21} )</td>
</tr>
</tbody>
</table>

*Georgia Institute of Technology (Courtesy of Dr. Joel Hales).
**Courtesy of Dr. Natalia Munera.
Table 4-9: Values of NLA and NLR of Silicon measured by Z-scan and BD at 1550 nm

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>GT* − Z-scan</th>
<th>CREOL − Z-scan**</th>
<th>CREOL − BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2^{ND}$ (cm/GW)</td>
<td>1.0</td>
<td>1.35</td>
<td>0.8</td>
</tr>
<tr>
<td>$n_2^{ND}$ (cm/GW²)</td>
<td>$5 \times 10^{-5}$</td>
<td>$5.5 \times 10^{-5}$</td>
<td>$5.0 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma_{FCA}$ (m²)</td>
<td>$1.5 \times 10^{-21}$</td>
<td>$35 \times 10^{-21}$</td>
<td>$24 \times 10^{-21}$</td>
</tr>
<tr>
<td>$\sigma_{FCR}$ (m²)</td>
<td>$-12 \times 10^{-21}$</td>
<td>$-10 \times 10^{-21}$</td>
<td>$-12 \times 10^{-21}$</td>
</tr>
</tbody>
</table>

*Georgia Institute of Technology (Courtesy of Dr. Joel M. Hales).
**Courtesy of Dr. Natalia Munera.

4.5. Three-Photon Absorption in Direct Gap Semiconductors

In this section, we use the Kane 8-band model (4 bands with double spin degeneracy) [49, 181, 184, 185] to predict the 3PA spectra of 15 different semiconductors and to compare our degenerate 3PA measurements of GaAs, conducted by the Z-scan method, with the theory. The details of the theory, developed and modified by previous students in our group [184-186], is presented in the APPENDIX A. The 8-band model allows us to calculate the quantum interference between the various pathways available for valence to conduction transitions. This calculation is not possible to be conducted for simple 2-band models. We decided to investigate this mainly because of discrepancies in previous comparisons for 3PA. Previous studies were based on two parabolic bands, developed by Brandi and de Araujo using Keldysh’s tunneling approach in 1983 [177, 187], and by Wherrett using perturbation theory in 1984 [176]. These two theories give similar results for 2PA; however, they disagree in the spectral dependence of the 3PA process. We, therefore, employed Kane 8-band model that includes the quantum interference between the various pathways and obtained a more richly featured 3PA spectrum compared to that obtained from simple 2-band models. The Kane 8-band model is sometimes referred to as the 4-band model where it includes the valence band comprising of heavy-hole (HH), light-hole (LH), and split-off bands.
(SO) and an individual conduction band. However, since each bands includes double spin degeneracy, it is more accurate to refer to it as the 8-band model. The energy difference between the zone center of the valence and conduction bands is the bandgap, \( E_g \), and the splitting between the SO band and the zone center is defined as the spin-orbit coupling parameter, \( \Delta \), or split-off separation. As the value of \( \Delta \) increases, it means that the strength of the spin-orbit coupling is larger.

We performed degenerate Z-scan measurements to characterize the 3PA spectra of GaAs for photon energies ranging \( E_g / 3 < h\omega < E_g / 2 \) to avoid 2PA and only isolate the absorption mechanism to 3PA. The data and theory for GaAs and previously measured spectra for ZnSe and ZnS are in excellent agreement for the shape of the spectra [49, 184, 185, 188-190]. For these measurements, we performed autocorrelation pulsewidth measurements and CA Z-scans on a piece of fused silica sample (considering no nonlinear dispersion at this spectral range) to measure \( n_2 \). These measurements are conducted to calibrate our experiments while using a well-known reference material.

Previously, similar measurements were performed using picosecond laser systems and significant contribution from free-carriers were observed [188, 191]. The picosecond measurements have the advantage of having a narrow band pulses (~ 1 nm) that is beneficial for resolving the spectra. However, it is important to note that we needed to keep our pulsewidth short enough to avoid NLO effects associated with free-carriers. Therefore, the measurements were performed using a femtosecond laser system (Clark-MXR) by pumping an OPA system and generating idler beams (1550 nm < \( \lambda \) < 2620 nm) and spectrally filtering the beams to obtain narrow bandwidths to resolve the 3PA spectra effectively. Femtosecond OA Z-scan measurements on GaAs are summarized and shown in Figure 4-9.
Figure 4-9. Degenerate OA Z-scan measurements to characterize 3PA spectra of GaAs

We fit our experimental measurements by solving the nonlinear wave equation and calculating the irradiance of the beam at the exit plane of the sample. For this calculation we consider
\[
\frac{d}{dz} I = -\alpha_3 I^3,
\]  

(4-23)

where \( \alpha_3 \) is the 3PA coefficient. We plotted our measurement results together with picosecond data previously measured in our group in Figure 4-10(a) and extracted FCA coefficients from picosecond measurements in Figure 4-10(b). In the end, we find relatively good agreement between picosecond and femtosecond 3PA measurements. The error bars in both graphs indicate relative errors between measurements, and the overall absolute errors are greater by a factor of 2. The FCA values are in qualitative agreement with those reported in Ref. [154].

![Figure 4-10. Measured data for 3PA in GaAs obtained from the 8-band model (black dotted line). Scaling the calculated spectrum by a factor of 4.2 (solid black curve), provides the best agreement with experimental data. Red circles are experimental data taken with ~150 fs pulses. Blue circles are data for the ~ 10 ps pulses. (b) The FCA cross section from the 2-parameter fit for 3PA and the FCA.](image)

We also used our 8-band model to predict the 3PA spectra of various direct-gap zinc blende semiconductors. For this prediction we only need the band structure parameters of each semiconductor, which are gathered from the literature as shown in the table below. We include a separate column for the ratio \( \Delta/E_g \) as it becomes apparent later that it significantly contributes to the 3PA spectral shape. This ratio spans from the order of unity in narrow-gap materials to \( 10^{-2} \) to \( 10^{-3} \) in wide-gap materials.
Table 4-10. Band-structure parameters for semiconductors in their zinc-blende form at 300 K (except ZnO, at 4.2 K) used for modeling of 3PA. The first three rows indicate materials where the SO band energy is relatively large ($\Delta/E_g > 1$) and the last five rows indicate materials with a the SO band energy is relatively small ($\Delta/E_g \ll 1$).

<table>
<thead>
<tr>
<th>Material</th>
<th>$E_g (eV)$</th>
<th>$\Delta (eV)$</th>
<th>$E_p (eV)$</th>
<th>$\Delta/E_g$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>InSb</td>
<td>0.174</td>
<td>0.81</td>
<td>23.3</td>
<td>4.65</td>
<td>[181, 192]</td>
</tr>
<tr>
<td>InAs</td>
<td>0.354</td>
<td>0.39</td>
<td>21.5</td>
<td>1.10</td>
<td>[192]</td>
</tr>
<tr>
<td>GaSb</td>
<td>0.726</td>
<td>0.80</td>
<td>22.0</td>
<td>1.10</td>
<td>[193, 194]</td>
</tr>
<tr>
<td>InP</td>
<td>1.344</td>
<td>0.11</td>
<td>20.4</td>
<td>0.08</td>
<td>[195, 196]</td>
</tr>
<tr>
<td>GaAs</td>
<td>1.42</td>
<td>0.34</td>
<td>28.9</td>
<td>0.24</td>
<td>[192, 197]</td>
</tr>
<tr>
<td>CdTe</td>
<td>1.44</td>
<td>0.91</td>
<td>20.7</td>
<td>0.63</td>
<td>[198]</td>
</tr>
<tr>
<td>CdSe</td>
<td>1.90</td>
<td>0.42</td>
<td>20.0</td>
<td>0.22</td>
<td>[38, 198]</td>
</tr>
<tr>
<td>ZnTe</td>
<td>2.28</td>
<td>0.92</td>
<td>19.0</td>
<td>0.40</td>
<td>[38, 198]</td>
</tr>
<tr>
<td>CdS$^a$</td>
<td>2.42</td>
<td>0.07</td>
<td>21.0</td>
<td>0.029</td>
<td>[195, 199]</td>
</tr>
<tr>
<td>ZnSe</td>
<td>2.67</td>
<td>0.42</td>
<td>24.2</td>
<td>0.16</td>
<td>[34, 38]</td>
</tr>
<tr>
<td>GaN</td>
<td>3.24</td>
<td>0.017</td>
<td>25.0</td>
<td>0.005</td>
<td>[192]</td>
</tr>
<tr>
<td>ZnO$^b$</td>
<td>3.44</td>
<td>0.0087</td>
<td>21.0</td>
<td>0.002</td>
<td>[200]</td>
</tr>
<tr>
<td>ZnS</td>
<td>3.54</td>
<td>0.06</td>
<td>20.4</td>
<td>0.017</td>
<td>[34, 38]</td>
</tr>
<tr>
<td>AlP</td>
<td>3.56</td>
<td>0.07</td>
<td>17.7</td>
<td>0.02</td>
<td>[192]</td>
</tr>
<tr>
<td>AlN</td>
<td>4.90</td>
<td>0.02</td>
<td>27.1</td>
<td>0.001</td>
<td>[192, 198]</td>
</tr>
</tbody>
</table>

$^a$Table is sorted based on the values of $E_g$. However, note that CdS belongs to the $\Delta/E_g \ll 1$ section.

$^b$8-band model parameters for ZnO are given for $T = 4.2$ K (Liquid helium temperature).

We employed our 8-band model to plot the 3PA spectra of each semiconductor listed in the table above for three-photon energy sum between $E_g$ and $3/2E_g$ (2PA edge). The predicted 3PA spectra, sorted from the lowest $\Delta/E_g$ to the highest $\Delta/E_g$ is shown below. We explicitly show the contributions from transitions that start in each of the three valence bands: HH, LH, and SO. These bands serve as the initial state "v" in the sum-over-states described in the APPENDIX A. Note that the contribution to the 3PA coefficient from a given initial valence band depends on the behavior of every other band due to the interference of 3PA pathways which pass through each band.
Figure 4-11. The contributions of starting in the HH (red), LH (blue) and SO bands (green) to the total (black) 3PA in several zinc-blende semiconductors, ordered from small to large $\Delta/E_g$ to show the
The ratio of $\Delta/E_g$, qualitatively determines the spectral dependence of the contributions of the initial states of LH, HH and SO bands to the 3PA. For example, narrow-gap semiconductors exhibit a rapid increase in the contribution of the transitions starting in the HH band (red curves) the contribution rises rapidly for photon energies, $3\hbar \omega$, to a peak just above the gap and then slowly decreases up to the turn-on of 2PA without quantum interference which is seen as a dip in the blue curves when starting from the LH band. Here the SO band (green curves) is separated so far ($\Delta/E_g \sim 1$) that its contribution occurs for energies above the turn-on of 2PA. For intermediate-gap materials, the contribution of the SO band occurs in the region between 2PA and 3PA and shows quantum interference between the various pathways to get from the initial SO band to the conduction band. The effect of the quantum interference can be seen as a dip in green curves. The effects of starting in the SO band move closer to the band edge as $E_g$ increases and its effects begin to dominate for large photon energies near the turn-on of 2PA, see e.g. going from GaAs to CdSe to ZnSe. For wide-gap materials, e.g. GaN and AlN, the contributions from starting in the LH and HH band becomes almost identical and quantum interference is no longer observed. At the same time, the contribution from starting in the SO band becomes similar in shape to that for the LH band in narrow-gap semiconductors and shows similar but much stronger quantum interference.

In Figure 4-12, we show the dependence of the normalized $\alpha_3$ on $\Delta/E_g$ to investigate the effects of spin-orbit coupling. For this plot, we fixed the value of bandgap, $E_g$, and considered $E_p = 21$ eV, while tuning $\Delta$ to obtain values of $10^{-4} < \Delta/E_g < 0.3$. This modeling exhibits that as $\Delta/E_g$
decreases, the contribution of the SO band becomes more significant close to the peak 3PA coefficient and the spectral shape. Additionally, as $\Delta/E_g$ increases, the resulted peak from the contribution of SO band goes closer to the 2PA edge [49].

Figure 4-12. Dependence of $\alpha_{3,norm}$ on the normalized three-photon energy sum and $\Delta/E_g$. $\alpha_{3,norm}$ is normalized to the value of 3PA when $\Delta = 0$. The range of $\Delta/E_g$ is limited to $\Delta/E_g < 0.3$ to better represent the effects of the SO band on the 3PA spectra.

4.6. Concluding Remarks

In conclusion, we presented how BD technique allows us to separately measure bound-electronic and carrier nonlinearities in semiconductors. This can provide us with complete information on the nonlinear processes in the telecommunication spectral range. Particularly, we show BD measurement results of GaAs and Silicon at 1300 nm and 1550 nm and compared our results with the measurements obtained via the Z-scan technique. These set of measurements are particularly useful for those research communities and applications that employ the third-order NLO response of semiconductors in the near-IR including silicon photonics, pulsed-laser radiation testing of electronic circuits, and deep lithography in silicon.
Additionally, we presented measurement of 3PA in GaAs in the spectral range where 2PA is absent. Our analysis, based on an 8-band model, separately characterizes the contributions of the transitions originating in the HH, LH, and SO bands to 3PA spectra. This analysis is imperative to account for quantum interference between different pathways that leads to a more richly featured 3PA spectrum compared to that obtained from simple 2-band models. In our analysis, the importance of spin-orbit coupling in determining the spectral shape becomes apparent. We also presented predictions of the 3PA spectra for 15 different direct-gap semiconductors in their zinc-blende form, which can be extremely useful in the case of low-dimensional systems such as quantum wells, quantum dots, nanocrystals, and two-dimensional semiconductors where spin-orbit coupling and inter-subband transitions are present [201-205].
CHAPTER 5: ENHANCED LIGHT-MATTER INTERACTION IN Epsilon-Near-Zero Regime

5.1. Introduction and Motivations

A long-term goal in NLO has been the development of materials with large and fast nonlinear response. The main conclusion by briefly examining the previous studies on various materials such as organic structures [86, 146, 206, 207] and inorganic semiconductor NLO materials [117, 140, 147, 153, 208] is that ultrafast NLO for practical device applications is still material limited. There is a large interest in developing novel materials for NLO device applications such as all-optical signal processing [73, 117], optical limiting [209-212], harmonic generation [213-215], nonlinear beam shaping [216-218], quantum nonlinear phenomena [219-222], and adaptive all-optical planar optics [79, 223] which has created a demand for development of materials whose refractive index or absorption coefficient can be dramatically changed using a low-power optical field with sub-picosecond response. These challenges may be addressed using a broad range of possibilities in the field of plasmonics, metamaterial, and metasurfaces [79, 224]. Recently, the spectral region where the refractive index, \( n \), or complex permittivity, \( \varepsilon \), of a material is near zero has become a topic of large interest due to their fascinating applications in boosting optical nonlinearities [3, 5, 7-9, 89, 91, 225, 226], ability to control the frequency [5, 89, 227], and harmonic generation [1, 2, 90, 92], etc.

In this section, we briefly explain the essential theoretical platform that is required to understand the enhanced nonlinear optical processes in ENZ regime. This includes both linear and nonlinear optical properties of highly doped semiconductors particularly TCOs around the plasma
frequency. From the linear optical perspective, the main characteristic of these large bandgap materials is that the plasma frequency is in the NIR spectral region and can be tailored via controlling the doping concentration. At this spectral region, the dominant nonlinear mechanism that manipulates the propagation of light is the redistribution of carriers excited via intraband transitions that can be accurately described by carrier transport mechanisms, or with phenomenological approaches like intraband transitions or two-temperature model. There are also different classes of optical materials exhibiting ENZ condition in other wavelength ranges. For example, polaritonic materials such as Silicon Carbide (SiC) possess a vanishing real part of the permittivity in the mid-infrared (mid-IR) due to the phonon-polariton resonances [228, 229], and topological insulators such as Bi$_{1.5}$Sb$_{0.5}$Te$_{1.8}$Se$_{1.2}$ (BTSE) single crystals have the ENZ condition in the visible and ultraviolet (UV) [230]. In terms of their NLO effects and performance, our interest is mainly in NLR effects enabling the possibility of all-optical switching and other applications operating based on nonlinear phase change. The two main characteristics of light-matter interaction at ENZ, enhanced longitudinal coupling to oscillating electrons and enhanced nonlinear phase change due to dispersive nature of TCOs, will be explained in detail in this chapter.

To describe induced nonlinear changes in optical properties of heavily doped semiconductors, it is first necessary to understand the transitions between electronic states which in turn leads to a redistribution of the population between different electronic levels. A complete picture of these processes requires information from the band structure of the solid, thus, methods such as density functional theory (DFT), which replaces the many-body electronic wave-function with a simpler problem, are required. We do not study this \textit{ab initio} approach and only limit our studies to transitions between electronic states from a phenomenological standpoint, but with a brief introduction to the underlying condensed matter physics background. The two main linear optical consequences
that emerge at ENZ will be utilized to explain the enhanced NLO properties. This theoretical background will serve as a tool to implement the analysis of the experimental data presented in this chapter.

5.2. Redistribution of Carriers in Transparent Conducting Oxides

TCOs are usually wide bandgap ($E_g > 3$ eV), highly doped ($\sim 10^{21}$ cm$^{-3}$) semiconductors. Prominent examples are ZnO doped with any of the group III elements (B, Al, Ga, In), SnO$_2$ doped with fluorine (FTO) or a mixture of In$_2$O$_3$ and SnO$_2$ (ITO), CdO, and Cd$_2$SnO$_4$ (cadmium stannate). For moderately doped semiconductors, the dopant atoms create localized states which can donate electrons or holes by thermal excitation or optical transitions, depending on how close the energy levels of the foreign atoms are to the band edge. In the case of highly doped semiconductors, the doping level is so high that the impurity atomic levels are close enough to form impurity bands, and the system ends up showing metal-like properties; however, the doping is always lower than the case of a true metal, thus the material has properties between those of semiconductors and metals. The Fermi level either shifts towards the conduction band or the valence band depending on the doping type [66, 231, 232]. At this point, the majority of Fermi-Dirac function overlaps with the band edge, and the material is called a degenerate semiconductor.

A realistic electric conductivity can be obtained by doping TCOs to degeneracy [233]. For example, for ITO resistivity of less than 5 $\mu\Omega$m is achievable, which makes them good conductors [233]. Metals are only one or two orders of magnitude below this. Titanium and lead have resistivity of 0.45 and 0.2 $\mu\Omega$m, respectively, larger than that of silver (0.015 $\mu\Omega$m) [66, 70, 231]. The physics of TCOs, however, is easier to understand by formally treating them as plasmas, with the plasma frequency in the near-IR spectrum. Hence, their electromagnetic response is heavily influenced by plasmon dynamics.
In order to understand the fundamental optical properties of TCOs, the dynamics that occur following absorption of photons must be studied, which is a vastly expanded subject. Here, we first discuss in detail how optical beams interact with plasmas. At the plasma frequency, electrons are free and there is no restoring force between the electrons and nuclei. Thus, the simplest realistic model is based on the nonlinear response of the Drude model to the laser pulse. We will discuss this approach in detail in the following sections.

5.2.1. The Drude-Free Electron Model

Recalling Eq. (2-46), the motion of a single electron bound to the nucleus can be described via a classical equation of motion. By solving the equation of motion in the frequency domain, we can easily obtain the position vector associated with the electron movement as below

\[ \mathbf{r} = -\frac{eE_{\text{loc}}/m_e^*}{\omega_0^2 - \omega^2 - i\gamma\omega}, \]  

(5-1)

The induced microscopic dipole moment for a single electron is defined as \( \mathbf{p} = -e\mathbf{r} \), which for a sufficiently small \( \mathbf{r} \), it linearly depends upon the electric field by \( \mathbf{p} = \hat{\alpha}(\omega)E_{\text{loc}} \), where \( \hat{\alpha}(\omega) \) is the frequency dependent atomic polarizability, which can be described via the following relation

\[ \hat{\alpha}(\omega) = \frac{e^2/m_e^*}{\omega_0^2 - \omega^2 - i\gamma\omega}, \]  

(5-2)

The atomic polarizability is complex because of the inclusion of a damping term. As a result, the polarization differs in phase from the local field at all frequencies. If there are \( N \) atoms per unit volume, the macroscopic polarization can be written as

\[ \mathbf{P} = N < \mathbf{p} >= N\hat{\alpha}(\omega) < E_{\text{loc}} > = \varepsilon_0\chi(\omega)E \]  

(5-3)

where \( \chi(\omega) \) is the electric susceptibility and \( E \) is the macroscopic electric field, as defined in CHAPTER 2. Except for some cases, the relationship between microscopic and macroscopic fields
can be very complex. For free-electron metals, which is also called the Drude metal, we can argue that since the conduction electrons are not bound, the field felt by the electrons is on the average just the macroscopic field $E$ [66, 231]. The other consequence of the free-electron model is that we should let $\omega_0 = 0$ in Eq. (5-2) since electrons are not bound.

Here, we consider the microscopic and macroscopic fields are similar, and we will keep the restoring force term to be non-zero to first study different spectral regions of the Lorentz model. The electric permittivity of the material can be found from the relation between the complex displacement field and the electric field, which can be described by

$$\epsilon(\omega) = 1 + \chi_{\text{background}} + \chi(\omega) = \epsilon_{\infty} + \frac{N \hat{\alpha}(\omega)}{\epsilon_0} = \epsilon_{\infty} + \frac{N e^2/m_e \epsilon_0}{\omega_0^2 - \omega^2 - i\gamma \omega}$$ \hspace{1cm} (5-4)

where $\epsilon_{\infty}$ is the electric permittivity far away from the Lorentz response (i.e. $\omega \gg \omega_0$), i.e. background permittivity, and $\omega_p^2 = Ne^2/\epsilon_0 m_e$ is defined as the plasma frequency. Consequently, the real and imaginary part of the electric permittivity can be extracted from Eq. (5-4), resulting in

$$\text{Re}(\epsilon) = \epsilon'(\omega) = \epsilon_{\infty} + \omega_p^2 \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}$$ \hspace{1cm} (5-5)

$$\text{Im}(\epsilon) = \epsilon''(\omega) = \omega_p^2 \frac{\gamma \omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}$$ \hspace{1cm} (5-6)

The complex refractive index, $\eta = n + i\kappa = \sqrt{\epsilon}$, can also be defined as below

$$n(\omega) = \sqrt{\frac{1}{2} \epsilon'(\omega) + \frac{1}{2} \sqrt{(\epsilon'(\omega)^2 + \epsilon''(\omega)^2)}}$$ \hspace{1cm} (5-7)

$$\kappa(\omega) = \sqrt{-\frac{1}{2} \epsilon'(\omega) + \frac{1}{2} \sqrt{(\epsilon'(\omega)^2 + \epsilon''(\omega)^2)}}$$ \hspace{1cm} (5-8)
The frequency dependencies of Eqs. (5-5) to (5-8) are plotted in Figure 5-1. Figure 5-1(a), shows the dispersion of the real and imaginary part of the electric permittivity, and Figure 5-1(b), shows the dispersion of the refractive index and extinction coefficient. In region I, $\omega \ll \omega_0$, $\varepsilon'' \approx 0$, and $\varepsilon' > 1$. In this region, we can conveniently conclude that $n > 1 \gg \kappa$ and $\varepsilon' = n^2$. This region describes the optical properties of lossless dielectric materials. Region II is characterized by strong absorption for dielectric materials. In Region III, $\omega \gg \omega_0$, and the electrons of the insulator respond as if they were free-electrons, which is because the photon energy is much greater than the binding energy of the electron [66, 231]. For semiconductors like Ge and Si, the band gap (i.e. $\sim \omega_0$) lies in the IR and region III is in the visible. The transition from region III to region IV happens at a frequency, $\omega_p$, called the plasma frequency, where $\varepsilon' (\omega_p) \approx 0$. This frequency is also called the dielectric to metal transition frequency or in electromagnetic waves community as the ENZ wavelength.

![Figure 5-1](image_url)

Figure 5-1. (a) Real and imaginary part of the electric permittivity, and (b) real and imaginary part of the complex refractive index for a Lorentz response of a single electron versus frequency.

At this spectral region, the restoring effect from the nucleus can be neglected since the electrons of a metal are not bound and the wave function for a free-electron is distributed uniformly throughout the
metal. Therefore, Eq. (5-4) will be simplified further and results in the Drude model, where the permittivity can be described via

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}$$  \hspace{1cm} (5-9)

Similar to Eqs. (5-5) and (5-6), we can extract the real and imaginary part of the permittivity, which can be described as

$$\text{Re}(\varepsilon) = \varepsilon'(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + \gamma^2}$$  \hspace{1cm} (5-10)

$$\text{Im}(\varepsilon) = \varepsilon''(\omega) = \frac{\omega_p^2\gamma}{\omega(\omega^2 + \gamma^2)}$$  \hspace{1cm} (5-11)

In materials with a very small loss, Eq. (5-11) is simplified to

$$\text{Im}(\varepsilon) = \varepsilon''(\omega) \approx \frac{\gamma\omega_p^2}{\omega^3}$$  \hspace{1cm} (5-12)

For a lossless material, where $\gamma = 0$, both Eqs. (5-10) and (5-11) will be simplified as below

$$\text{Re}(\varepsilon) = \varepsilon'(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2}$$  \hspace{1cm} (5-13)

$$\text{Im}(\varepsilon) = \varepsilon''(\omega) = 0$$  \hspace{1cm} (5-14)

Let us now take a closer look at the damping factor, $\gamma$, and understand the origins of this parameter in this model. The origin of the damping term for a free-electron metal is the scattering of electrons associated with electrical resistivity. An electron in a metal makes random collisions with the defects and other electrons. These are all inelastic scattering mechanisms where electrons have their paths slightly and randomly deflected. Now, assume that the probability that an electron makes a collision in a time interval $dt$ is $dt/\tau$, where $\tau$ is a constant. If we consider $n(dt)$ to be the number of electrons at time $t = dt$ before they make a collision, then $n(0) - n(dt)$ will be the number of
electrons at time $t = dt$ that have experienced collisions with other particles. Therefore, we can say

$$n(0) - n(dt) = n(0)dt/\tau$$  \hspace{1cm} (5-15)

This equation can also be written as

$$\frac{1}{dt}\{n(dt) - n(0)\} = -\frac{1}{\tau}n(0)$$  \hspace{1cm} (5-16)

By solving this equation between time $t = 0$ and $t = t$,

$$n(t) = n(0)e^{-t/\tau}$$  \hspace{1cm} (5-17)

The constant $\tau$ is the relaxation time, and it is possible to show that it is also the mean free time between collisions. Now that we know the mean free time, we can look at the interaction between electrons and the electric field in more detail. Electrons accelerate upon interaction with electric field. This is because photons transfer energy to the electrons, mainly in the form of the kinetic energy. This acceleration is the reason of collisional effects between electrons and lattice. If the electric field is constant and only applied for a time interval $dt'$, with no energy loss mechanism (i.e. while $\gamma = 0$), the velocity increment in time $dt'$ will be $(-eE/m_e^*)/dt'$. This increment will maintain for any time after $t$ when there is no loss. Now, if there is an energy loss mechanism present, the velocity increment will be damped, at a time $t$, as a result of scattering and by the factor $exp(-(t - t')/\tau)$. Thus, for a constant electric field, the average velocity of electrons can be described by

$$\bar{v} = \int_{-\infty}^{t} (-eE/m_e^*) e^{-t'/\tau} dt' = -eE\tau/m_e^*$$  \hspace{1cm} (5-18)
Now if we consider a time-harmonic electric field, \( \mathbf{E}(t) = \mathbf{E}(0)e^{i\omega t} \), the average velocity can be found by a simple convolution

\[
\bar{v}(\omega) = \int_{-\infty}^{\infty} \left( -e\mathbf{E}(0)/m_e^* \right) e^{i\omega t'} e^{-\frac{t-t'}{\tau}} dt' = -\frac{e}{m_e^*} \mathbf{E}(0)e^{i\omega t} e^{\frac{-\tau}{1-i\omega \tau}}
\]

(5-19)

This gives us the average velocity of electrons in a system where there is a probability for collision, meaning that there is an electron loss mechanism. We can easily relate this equation to Eq. (5-9) by writing the average velocity in terms of the position in frequency domain as

\[
\mathbf{r}(\omega) = \frac{1}{i\omega} \bar{v}(\omega) = -\frac{e\mathbf{E}/m_e^*}{\tau (i\omega + \omega^2)}
\]

(5-20)

Now, we can conclude that \( \tau = 1/\gamma \). In this approach, we consider an effective mass for electron, \( m_e \), because the screening forces from other electrons and lattice must be also considered. This is because we consider a homogeneous gas of electrons as opposed to our previous derivation, Eq. (2-46), where we assumed a single electron. We can also define the mobility of electron, \( \mu \), as

\[
\mu = \frac{|v|}{|E|} = \frac{e\tau}{m_e^*} = \frac{e}{\gamma m_e}
\]

(5-21)

5.2.2. Linear characterization

We performed spectroscopic ellipsometry on a commercially bought thin layer of ITO (~ 320 nm) deposited on a glass substrate (~ 1.1 mm) to characterize the real and imaginary part of the permittivity of this material. Since ITO exhibits dielectric (transparent) behavior at shorter wavelengths and metallic (lossy) behavior at longer wavelengths, a combination of Lorentz oscillator model and Drude model must be used to fit the ellipsometry data. For example, the ITO generalized built-in model in the library of the commercial software of the ellipsometry equipment, uses these two combinations to fit the data as shown below
Figure 5-2. ITO built-in library model used in commercial software for analyzing ellipsometry data. The values for $\psi$, as defined in §3.1.1., for our ITO ellipsometry measurements are shown in Figure 5-3, where the dielectric transparent region is considered to be at $\lambda < 1000$ nm, whereas the metallic lossy region occurs at $\lambda > 1100$ nm.

Figure 5-3. Experimentally measured values of $\psi$ using spectroscopic ellipsometry (solid lines) at three different angles of incident. Fit with Lorentz oscillator model in the transparent region and Drude model in the lossy region. Fits are courtesy of my colleague Sanaz Faryadras. The tiny noisy behavior around $\lambda = 1100$ nm can be attributed to non-idealities of the sample and measurements, which can be addressed using a B-spline fit that interpolates multiply connected polynomial functions [234, 235] and includes a bandwidth that models the smearing of wavelength at the detection spectrometer. This model does not consider the underlying physical phenomena that results in the observed optical behavior. We attempted to use this approach and resolved the issue with fitting the non-idealities, as shown below for the wavelengths longer than 900 nm.
The resulting experimentally measured dispersion of ITO, which can be considered a Drude-type material in the near-IR is shown in Figure 5-5. The blue shaded region is the frequency range where $|\text{Re}(\varepsilon)| < 1$, which is also called the ENZ region. The ENZ condition occurs at the plasma frequency in absence of loss. The plasma frequency can be controlled via the doping rate [236] by inducing a non-parabolicity in the band structure and altering the free-carrier density dependent effective mass of electron, $\omega_p^2 = Ne^2/\varepsilon_0 m_e(N)$ [236]. By manipulating $N$, it is possible to effectively control the value of $\omega_p$ to span a wide frequency range in the near-IR. We discuss this in more detail in §5.2.4.

For our ITO sample, we fit he extracted dispersion from ellipsometry measurements by a Drude model with $\omega_p = 2\pi \times 469 \times 10^{12}$ rad/s, which corresponds to a zero-crossing at $\lambda_{ENZ} \approx 1240$ nm. The doping concentration of this material can be approximated to be $N \approx 7.63 \times 10^{26}$ m$^{-3}$, considering the effective mass of electron for ITO to be $m_e^* = 0.28 m_e$ [237, 238]. The value of the free-electron density of ITO is two orders of magnitude smaller than that of gold and silver [66, 231, 237, 238], which is the main reason of the smaller plasma frequency compared to typical metals and consequently a dielectric to metal transition happening at shorter wavelengths.
As we mentioned earlier, there is a spectral region where the real part of the permittivity crosses zero, i.e. $|\varepsilon'| < 1$, so-called the ENZ region. This is primarily a condition upon $\varepsilon'$, and the imaginary part of the permittivity, $\varepsilon''$, may take any value, although large values of $\varepsilon''$ creates significant optical loss which may be detrimental to certain applications. However, another definition is based on having a near-zero refractive index, which is attainable for low loss ENZ materials.

![Diagram of optical constants](image)

Figure 5-5. Optical constants of a Drude material (ITO) with $\omega_p = 2\pi \times 469 \times 10^{12}$ rad/s, $\gamma = 0.0410\omega_p$, and $\varepsilon_\infty = 3.65$ measured by the spectroscopic ellipsometry technique. The real and imaginary part of permittivity versus frequency, the blue shading is the ENZ region.

We may also impose a similar condition upon the refractive index such that $n < 1$, which is also sometimes called the near-zero index (NZI) region [7]. The NZI condition can be realized when

$$\text{Re}(\sqrt{\varepsilon}) \approx \frac{1}{2} \sqrt{\text{Re}(\varepsilon) + \sqrt{\text{Re}(\varepsilon)^2 + \text{Im}(\varepsilon)^2}} < 1.$$ Practicaly the plane loci of the permittivity values satisfying the NZI condition is governed by $\text{Im}(\varepsilon) < 2\sqrt{1 - \text{Re}(\varepsilon)}$, which is shown below. This condition creates a range of potential combinations which extends the NZI conditions to $\text{Re}(\varepsilon)$ values much less than zero. However, the loss will be very large at this condition. Clearly, NZI can also be achieved when $|\text{Re}(\varepsilon)| < 1$ and $|\text{Im}(\varepsilon)| < 1$, which is a more flexible constraint and a useful condition in terms of loss.
Another parameter that is typically used for such materials with refractive index below unity and reasonable loss is a figure of merit (FOM) that simultaneously quantifies the loss, $\kappa$, and the difference of $n$ from unity [239]. This parameter is typically defined as $\eta_{FOM} = \kappa/(1 - n)$. For comparison, we plot $\eta_{FOM}$ for silver (Ag) with zero-crossing in UV [240], AZO [241], ITO (this work), with plasma resonances in near-IR, and fluoride materials such as BaF$_2$ and CaF$_2$, which are materials with optical-phonon resonances in the mid-IR [242]. Materials with negligible loss and NZI condition exhibit smaller FOM values. In this definition, we only consider spectral regions where $n < 1 - \kappa$ where the FOM remains between 0 and 1. For instance, BaF$_2$ reaches the smallest value of $\eta_{FOM}$ [242], and exhibits a relatively broad spectral region where $0 < \eta_{FOM} < 1$. In general, fluorides are good candidate materials for ENZ applications, particularly, because values of loss are smaller for phononic resonances, where collective oscillation of the lattice occurs, compared to plasmonic materials, where electrons oscillate collectively [9, 224], hence reaching smaller values for $n$ is more feasible, as also shown in Figure 5-6. FOM for TCOs such as AZO and ITO are also shown, where it reaches to larger values and closer to the unity, with ITO having a slightly
better performance. The loss in TCOs are larger than fluorides; however, their operating spectral region in the near-IR makes them more attractive choices. Similarly, silver shows larger FOM values in UV with much narrower spectral bandwidth. Larger loss for silver deteriorates its performance in nonlinear optical applications.

Figure 5-7. FOM for various phononic, plasmonic, and metallic materials. Material data from refs [240-242]. The magenta data (ITO) is from this work.

In this work, considering a vanishing real part of permittivity we assume NZI and ENZ conditions are identical, and thus we only use the term “ENZ” here.

Both ENZ and NZI regions happen very close to the bulk plasma frequency. In a lossy material, this dielectric to metal transition happens at \( \omega = \sqrt{\omega_0^2 / \epsilon_\infty - \gamma^2} \). For an idealized material with zero loss, \( i.e. \gamma = 0 \), both \( \text{Re}(\epsilon) \) and \( n \) become zero exactly at the bulk plasma frequency. At this very special case, the material exhibits properties of a perfect conductor (PEC), where there is an infinite conductivity and a constant static magnetic flux. The constraint of constant magnetic flux can be easily verified by the Ampere’s law when the electric permittivity is zero. The consequence of an infinite conductivity is a complete reflectivity from an ideal ENZ material, \( i.e. \).
\( \text{Re}(\varepsilon) = \text{Im}(\varepsilon) = 0 \). Hence, for the light to propagate in the medium, a minimum amount of loss is required. However, as in many other cases, dielectric losses also present a significant problem by limiting the useful applications of ENZ materials. There have been many different approaches to mitigate these limitations by engineering the loss via multilayer stack of metal and dielectric \([225, 243]\), longitudinal and hyperbolic metamaterials \([244, 245]\), and material engineering and chemistry in conducting oxides such as ITO and nitride materials \([246, 247]\), etc.

Recalling Eqs. (2-52) to (2-55), upon nonlinearly induced changes in plasma frequency, \( \Delta \omega_p \), and damping factor, \( \Delta \gamma \), the real and imaginary permittivity will be modified, and thus NLA and NLR of Drude materials can be derived. In the two following sections, we first look closer into the optical transitions in metals and phenomenological models to describe the origins of optically induced changes in Drude materials.

5.2.3. Intraband Absorption and Two-Temperature Model

To understand the interaction of light with doped semiconductors, first, we present a simplified review on optical transitions in such materials. In n-type heavily doped semiconductors such as ITO, which exhibits a metallic type behavior, the Fermi level is located in the conduction band. The conduction band in oxides such as ITO is formed from the \( s \) orbitals of the cation and \( p \) orbitals of the oxygen atoms that participate in the charge transport in doped materials. The overlap of these orbitals is highly dependent on oxygen coordination \([248]\). The valence band, however, is formed from the \( d \) orbital, similar to Indium \((4d^{10}5s^25p^1)\) and Tin \((4d^{10}5s^25p^2)\). Additionally, the electrons have completely occupied the \( d \) band \([248]\).
Similar to other solids, upon absorption of photons, there are interband and intraband optical transitions. These transitions are illustrated in Figure 5-8 considering a simplified band structure. The interband transition, Figure 5-8(a), is a direct transition exciting an electron to another band. This transition only involves the excitation of an electron by a photon, and the value of the wavevector \( k \) of the electron is essentially unchanged under the conservation of momentum. This transition has a threshold energy analogous to that for the excitation of an electron across the band gap in an insulator. The intraband transition, Figure 5-8(b), corresponds to optical excitation of an electron from below the Fermi energy, \( E_F \), to another state above the Fermi energy but within the same band. These transitions are described by the Drude model and do not have any threshold energy.

In intraband transitions, the absorption process requires a phonon or an imperfection to conserve the momentum. This results in electron-phonon \( (e - ph) \) scattering mechanism in metals, which leaves two “hot” carriers – electrons and hole – close to the Fermi energy due to the conservation of energy. The electron-electron \( (e - e) \) scattering mechanism also leads to an absorption process while the momentum is conserved, as sketched in Figure 5-8(b). The momentum conservation in this case results in the generation of hot carriers close to the Fermi level. However, for interband
absorption processes, no hot carriers are created. This is because hot electrons are around the Fermi level, and the hole in the $d$-band has a large effective mass and small velocity [66]. Once the hot carriers have been generated, they lose their energy via scattering mechanisms transferring the energy to electrons in the form of kinetic energy which will be then transferred to the lattice (phonons) in the form of thermal energy.

This redistribution of carriers upon excitation with ultrashort (sub picosecond) pulses can be phenomenologically described via the two-temperature model (TTM) [65, 249-256]. In this model, the large density of electronic states around the Fermi energy allows us to describe the electron response by a quasi-continuum of states. The femtosecond pulse excitation will smear the electron distribution, and the temporal dynamics of the system is governed by the TTM model. The two-step process of $e-e$ and $e-ph$ scattering can be modeled with series of semiclassical relations based on the Boltzman transport equations by means of two coupled differential equations relating the temperature of electrons and phonons:

$$ C_e \frac{\partial T_e}{\partial t} = -g(T_e - T_{ph}) + aN, $$ (5-22)

$$ C_l \frac{\partial T_{ph}}{\partial t} = +g(T_e - T_{ph}) + bN, $$

$$ \frac{\partial N}{\partial t} = -aN - bN + P(z,t), $$

where $C_e$ and $C_l$ are the electron and lattice heat capacitance, $T_e$ and $T_{ph}$ are the electron and phonon (lattice) temperature, $g$ is the electron-phonon coupling constant which is a constant of a material dependent on the Fermi velocity and Fermi energy, $N$ is the energy density in the non-thermalized electronic distribution, $a$ is the electron gas heating rate, $b$ is the electron-phonon coupling rate, and $P(z,t)$ is the absorbed energy density and is given by
\[ P(z, t) = (1 - R - T)ae^{-az}I(t), \]  

(5-23)

where \( R \) and \( T \) are the reflection a transmission coefficients of intensity, respectively, \( a \) is the linear absorption coefficient, and \( I(t) \) is the intensity of the incident light. In this approach, we consider that the total electronic distribution is separated into a thermalized part (characterized by \( T_e \)) and a nonthermalized part (characterized by \( N \)), which is verified numerically in Ref. [249, 257]. We also consider that the beam size is much larger than the electrons diffusion length, meaning that we can ignore the transverse dependence.

Upon absorption of photons, the material experiences a delayed thermalization process, which is usually described by a thermalization response function, \( h_{th}(t) = e^{-t/\tau_{th}} \cdot H(t) \). In this relation \( H(t) \) is the Heavyside function ensuring causality, and \( \tau_{th} = 1/(a + b) \) [258]. The solution of the rate equation of the nonthermalized energy density \( N \), based on Eq. (5-22) can be calculated via

\[ N(t) = \int_{-\infty}^{t} P(t')h_{th}(t - t')dt' \]  

(5-24)

\( a \) and \( b \) can be obtained from the Lindhard dielectric function and using the Fermi liquid theory [232], which for the excitation pulse with a photon energy of \( E_e = \hbar\omega_e \) is described by

\[ a = \frac{1}{2\tau_{e-e}} = C \left( \frac{\omega_e^2}{8\pi^2\omega_p} + \frac{2\pi^2k_B^2m_e^2}{\hbar^2\omega_p} \right), \]  

(5-25)

\[ b = \frac{1}{2\tau_{e-p\hbar}} = \frac{1}{4} \frac{g}{C_e}, \]  

(5-26)
where $\tau_{e-e}$ is the electron-electron energy relaxation time, $\tau_{e-ph}$ is the electron-phonon energy relaxation time, respectively, and $C$ is a constant scaling factor [232, 253]. For an intense excitation, the heat capacitance, $C_e$, also becomes temperature dependent. For an uncorrelated contribution of each mechanism, the damping parameter for the Drude model, $\gamma$, can be described as

$$\gamma(T_e) = \frac{1}{\tau_{e-e}(T_e)} + \frac{1}{\tau_{e-ph}(T_e)}$$  \hspace{1cm} (5-27)

For a material in a thermally equilibrium state, the $\tau_{e-ph}(T_e)$ is much smaller, and therefore, the contribution of $\tau_{e-ph}$ to $\gamma$ becomes dominant. The temperature dependence of $C(T_e)$ makes the damping factor nonlinear, and thus the total decay time will not be exponential. The analytical solution of the TTM model is difficult, since it requires a precise knowledge of the injected energy. However, this model gives us useful insights on the nonlinear dependence of the damping factor of Drude permittivity on the incident light irradiance. Recalling Eq. (2-53), these changes in damping factor results in a nonlinearly induced change in the imaginary part of the permittivity, $\Delta\varepsilon''$, and thus inducing both NLA and NLR. It is important to note that, in addition to the changes in the damping factor, there are optically induced changes in the plasma frequency, $\Delta\omega_p$, that modify both real and imaginary parts of the electric permittivity. While phenomenological modeling based on TTM can be insightful, the Drude model intrinsically treats all electrons as identical in free-electron systems such as metals [17, 66, 231]. A more accurate approach to model the behavior of electrons is to consider a band structure, in which electrons have different energies.
5.2.4. Non-parabolic band structure

In materials, such as TCOs, a non-parabolic band structure needs to be considered to account for higher energy levels [238, 259, 260]. In this case, the dispersion law deviates from parabolic relation close to band edge of semiconductors, and the first-order non-parabolic band can be described via [66, 70]

\[
\frac{\hbar^2 k^2}{2m_e^*} = E + \alpha E^2,
\]

(5-28)

where \( k \) and is the wavevector, \( E \) is the electron energy in the conduction band, and \( \alpha \approx 0.4191 \text{ eV}^{-1} [260, 261] \) is a constant for modeling the first-order non-parabolic factor. In addition, \( m_e^* \) is the electron effective mass at the conduction band edge where \( E = E_C \), in which \( E_C \) is energy of the bottom of the conduction band. The electron effective mass at each energy, \( m_e^*(E) \) can then be described as

\[
\frac{1}{m_e^*(E)} = \frac{1}{\hbar^2 k} \frac{dE}{dk} = \frac{1}{m_e^*} \left( 1 + 2\alpha E \right)
\]

(5-29)

Upon absorption of photons by free-carriers, the kinetic energy of the electron (described by the electron temperature, \( T_e \)) will change. However, for sub-bandgap photons, only intraband band absorption occurs and the electron concentration, \( n \), will remain the same. The electron concentration can be written as

\[
n = \int_{E_C}^{\infty} f(E) \cdot \rho(E) dE,
\]

(5-30)

where \( f(E) = \{\exp((E - E_F)/kBT_e ) + 1\}^{-1} \) is the Fermi-Dirac probability function, and \( \rho(E) \) is the density of states. To derive \( \rho(E) \), first we deduce \( k(E) \) from Eq. (5-28) as
\[ k(E) = \frac{1}{\hbar} (2m_e^*E)^{1/2}(1 + \alpha E)^{1/2} \]  

(5-31)

Now, we can write \( \rho(E) \) as

\[ \rho(E) = \frac{1}{3\pi^2} k^3(E) = \frac{1}{3\pi^2 \hbar^3} (2m_e^*E)^{3/2}(1 + \alpha E)^{3/2} \]  

(5-32)

The conservation of the electron concentration means that the occupied state distribution, \( f(E) \cdot \rho(E) \), will be smeared to keep the integration in Eq. (5-30) constant. This smearing manipulates the band structure, and subsequently the overall effective mass will be changed. This is the direct consequence of the non-parabolicity of the conduction band, in which the higher energy states have a different effective mass than the lower energy states. Now, recalling the Drude model, the plasma frequency and damping factor can be written as

\[ \omega_p^2 = \frac{Ne^2}{\varepsilon_0 m_e^*(E)} \]  

(5-33)

\[ \gamma = \frac{e}{\mu m_e^*(E)} \]  

(5-34)

Upon an induced increase in the electron effective mass, \( \Delta m_e^* \), we can deduce the optically induced changes in plasma frequency and damping factor as

\[ \Delta \omega_p \approx -\frac{1}{2m_e} \Delta m_e^* \]  

(5-35)

\[ \Delta \gamma \approx -\frac{\gamma}{m_e} \Delta m_e^* \]  

(5-36)
These changes manipulate the electric permittivity, and subsequently will induce NLA and NLR effects. Since the electron effective mass increases upon absorption of photons, the plasma frequency will decrease, and thus the real part of the electric permittivity, $\Delta \varepsilon'$, described by Eq. (2-52), increases. Similarly, decrease in both plasma frequency and damping factor results in reduction of the imaginary part of the electric permittivity, $\Delta \varepsilon''$. These effects will subsequently increase the NLR and decrease the NLA.

**5.3. Experimental Characterization of Nonlinear Mechanisms at Epsilon-Near-Zero**

Here, we present the progress towards linear and nonlinear optical characterization of a 320 nm thin film of ITO deposited on a 1.1 mm glass substrate. The ENZ wavelength of this sample is located at $\lambda_{ENZ} = 1240$ nm, characterized by ellipsometry as discussed before. TCOs are, in general, popular particularly because the ENZ wavelength is easily accessible within the spectral range of OPA systems.

We performed series of BD and Z-scan measurements to study the enhancement mechanism and physical origin of optical nonlinearities. Here, we used a Ti:sapphire laser system at a repetition rate of 1 kHz with ~150 fs, 775 nm pulses to pump an OPA (Light Conversion, TOPAS-C) whose output can be tuned in the wavelength range of 1060 nm to 2630 nm. The working principle of this laser was described in CHAPTER 3. For our slightly non-degenerate BD measurements, we used another Ti:sapphire laser system with ~35 fs pulsewidth and output beam at 800 nm (Coherent Legend Due +). This system was only used for our slightly non-degenerate BD measurements where the pump and probe beam pulses were both generated from separate OPA systems. We have studied the spectral dependence of the NLO properties of ITO including NLR and NLA, which alters the phase and amplitude of the pulse field, respectively. In order to have
reliable results, we have used two measurement techniques, each of them providing us with information related to the NLO properties. We present nonlinear BD measurements, which directly gives us the time and polarization dependence of the NLR, as opposed to pump-probe methodology in transmission and reflection, which requires us to decouple the effect of NLR and NLA from the two measurements. We have also performed Z-scan to compare the nonlinear phase change, $\Delta \varphi$, measured via BD to the values extracted from Z-scan. We use the definition of effective NLR as

$$\Delta n_{\text{eff}} = \frac{\Delta \varphi}{k_0 L},$$

where $k_0$ is the wavevector of the incident beam in free-space and $L$ is the thickness of the material. It is important to note that, while doing NLO measurements, the results gives us the sum of the effect from each layer of the material, \textit{i.e.} ITO film and glass substrate. However, the sum signal is not necessarily the superposition of the nonlinear phase change of each individual layer. This is mainly because ITO film is highly absorptive and thus the irradiance of the pump at the front and back surface of ITO is different.

We performed slightly nondegenerate BD measurements with pump pulses at $\lambda_e = 1300$ nm and probe pulses at $\lambda_p = 1240$ nm generated from separate OPA systems. We attempted to do these measurements to compare our results with the degenerate measurements obtained from the Z-scan technique. To separate the contribution of the substrate to the total BD signal, we performed similar measurements on the glass substrate in order to eliminate the contribution of the substrate and to scale the measurements of the nonlinear phase change of ITO, $\Delta \varphi_{\text{ITO}}$, to a known reference, in this case $\Delta \varphi_{\text{SiO}_2}$ to account for experimental errors. This process is possible mainly because the carrier nonlinearities are dominant in ITO, and thus they do not exhibit any polarization anisotropy.
We performed measurements for two relative polarization combinations and found that $\Delta \varphi_{\text{ITO}}$ is equal for both parallel and perpendicular polarizations, $\Delta \varphi_{\text{ITO},\parallel} = \Delta \varphi_{\text{ITO},\perp}$ at longer delays with a slight difference close to the zero delay between pump and probe pulses. This indicates that there is a polarization-dependent nonlinear process occurring around the zero delay. This response is due to the contribution of the third-order bound electronic nonlinearities of the glass substrate to BD signal, where $\Delta \varphi_{\text{SiO}_2,\parallel} = 3 \Delta \varphi_{\text{SiO}_2,\perp}$. We can, therefore, consider the total parallel and perpendicular BD signals to be described as $\Delta \varphi_{\text{total},\parallel} = \Delta \varphi_{\text{ITO},\parallel} + \Delta \varphi_{\text{SiO}_2,\parallel}$ and $\Delta \varphi_{\text{total},\perp} = \Delta \varphi_{\text{ITO},\perp} + \Delta \varphi_{\text{SiO}_2,\perp}$. Now by subtracting the total parallel and perpendicular signals, we obtain $\Delta = \Delta \varphi_{\text{total},\parallel} - \Delta \varphi_{\text{total},\perp} = 2/3 \Delta \varphi_{\text{SiO}_2,\parallel}$, which is purely related to the substrate. In this way, we are eliminating the contribution of the ITO film and only calculating the scaled value of $\Delta \varphi_{\text{SiO}_2,\parallel}$. We note that $\Delta$ also follows the cross-correlation between excitation and probe pulses and can be used to establish the zero-delay time between respective pulses. By multiplying $\Delta$ with a factor of $3/2$ and subtracting from $\Delta \varphi_{\text{total},\parallel}$, we can accurately extract the signal from the ITO film from $\Delta \varphi_{\text{ITO},\parallel} = \Delta \varphi_{\text{total},\parallel} - 3/2 \Delta = 3/2 \Delta \varphi_{\text{total},\perp} - 1/2 \Delta \varphi_{\text{total},\parallel}$.

This approach is useful particularly when linear absorption and NLA effects significantly change the irradiance of the excitation beam reaching the back surface of the ITO while the film is on the front of the sample. This is an accurate approach to extract the contribution of ITO to the BD signal, but it needs to be noted that a precise calibration is required to confirm that waveplates used in our experiments are efficiently rotating the polarization. We, therefore, typically perform a calibration mechanism before starting our final measurements, in which we find the maximum signal obtained from a glass substrate and rotate the polarization of the probe beam to measure the polarization anisotropy of the signal versus the relative polarization between pump and probe.
pulses. In some cases, the HWP does not efficiently rotate the polarization, particularly when the quality of the HWP is not good or multi-order broadband HWP is used. The error can be also due to the misalignment of the beam passing through the HWP. However, we note that a small error in HWP rotation efficiency does not significantly alter this extraction approach. To quantify this, let us assume that instead of having an exact ratio of 3 between parallel and perpendicular measurements of third-order nonlinearities, the ratio is slightly less. So we consider that \( \Delta \varphi_{\text{SiO}_2,\parallel} = (3 - \alpha) \Delta \varphi_{\text{SiO}_2,\perp} \), where \( \alpha \) characterizes the error. In this case, we obtain \( \Delta \alpha = \Delta \varphi_{\text{total},\parallel} - \Delta \varphi_{\text{total},\perp} = (2 - \alpha)/(3 - \alpha) \Delta \varphi_{\text{SiO}_2,\parallel} \), where \( \Delta \alpha \) is the value of the difference between total parallel and perpendicular signals in presence of the associated error of \( \alpha \). If we consider \( \alpha \approx 0.3 \), indicating a 10% smaller values than what is expected for third-order effects, we obtain \( (2 - \alpha)/(3 - \alpha) \approx 0.63 \) as opposed to \( 2/3 \approx 0.66 \), which adds \( \approx 5\% \) error to our calculations. By applying this extraction approach we obtained the slightly nondegenerate BD signals directly from ITO films, which are shown in Figure 5-9.
Figure 5-9. (Right axis) BD measurements, $\Delta E/E$ (%), versus delay for glass substrate and extracted BD signals for the ITO film at both parallel and perpendicular relative polarizations. (Left axis) Nonlinear phase change directly calculated from our BD measurements.

As we discussed earlier in CHAPTER 3, the normalized BD signal can be written in terms of the NLR or the nonlinear phase change as

$$\frac{\Delta E}{E}(\tau_D) = \frac{2\sqrt{\pi} w_{0,p}}{\sqrt{\pi} e w_e} \Delta \varphi(\tau_D),$$

where $w_{0,p}$ and $w_e$ are the probe size at the focus and pump size at the sample, respectively, and $\tau_D$ is the relative delay between pump and probe pulses. The right axis, $\Delta \varphi$, is calculated directly from Eq. (5-32).

We note that the maximum overlap between two pulses occurs at $\tau_D = 0$, while the probe pulse experiences the maximum nonlinear phase change at a slightly later delay at around $\sim 100$ fs. This delay in the nonlinear response is indicative of the cumulative nature of the nonlinearity in this material. We analysed and fit the experimental signals via a single exponential rise and decay response function convolved with the cross-correlation of the excitation and probe pulses to extract the temporal dynamics of the nonlinear phase change. The decay time of sub-picosecond in such materials is because the dominance of electron-electron over electron-phonon mechanisms.
It is important to note that the signal obtained for parallel and perpendicular polarizations are identical for ITO. This also exhibits the lack of polarization anisotropy of the nonlinear mechanism in ITO, which indicates the free carrier related origin of the nonlinearities.

We also performed Z-scan measurements at 1240 nm to measure the degenerate $\Delta \varphi$. Figure 5-10 presents OA, CA, and CA/OA signals measured by the Z-scan method. We observe an increase in transmission for OA signal, Figure 5-10(a), which indicates that upon an increase in the input irradiance, the transmission of the beam increases. This can be due to either a saturable absorption behavior or a significant change in material index, which subsequently reduces the reflection of light from air/ITO interface. In ITO, since the material index is less than one (index of air), when the material index increases significantly, the index mismatch between the medium that includes the input light (air) and the ITO film drops drastically. This reduces the reflection, and thus the transmission increases. Therefore, as opposed to typical Z-scan measurements, there will a contribution from NLR to the OA signal, which makes extraction of nonlinear parameters complicated. Additionally, because of the large signal obtained in OA measurements, the signal from CA will be greatly distorted, Figure 5-10(b); however, by extracting CA/OA it is still possible to efficiently calculate the nonlinear phase change from $\Delta T_{p-v} \approx 0.406(1 - S)^{0.27}\Delta\varphi_0$, in which $S$ is the transmission of the closed aperture and $\Delta\varphi_0$ is the on-axis nonlinear phase change. We note that this relation is accurate within 2% for $|\Delta\varphi_0| < \pi$ [106].
Figure 5-10. (a) OA, (b) CA, and (c) CA/OA signals obtained from Z-scan measurements on the ITO film.

We also compared our Z-scan results with our degenerate BD measurements by plotting the extracted $\Delta \varphi$ (rad) and $\Delta n_{\text{eff}}$ versus irradiance in Figure 5-11.

Figure 5-11. Irradiance dependence of $\Delta n_{\text{eff}}$ and $\Delta \varphi$ calculated from BD (red) and Z-scan (blue) measurements.

Note that the magnitude of the $\Delta E/E$ signal, and thus $\Delta \varphi$, in ITO is only $\sim 3 \times$ larger than that of the glass substrate at parallel polarization. However, the ITO film is $\sim 3500 \times$ thinner than the substrate, which indicates that the nonlinear phase change occurs in an extremely smaller propagation length compared to glass. This means that the effective NLR is $\sim 10^4$ larger than that of glass and can reach values near unity below the damage threshold of the film. This also agrees with our measurements obtained with the Z-scan technique as reported above. We will discuss this superior performance of ITO in more detail in §5.3.2.
We extended our measurements to investigate non-degenerate nonlinear effects of ITO in proximity to the ENZ region. The next two section studies the effect of the enhanced coupling to longitudinal plasmons, which is due to the enhanced absorption of the longitudinal electric field. We present the physical origin of this enhancement and BD measurements to investigate their contribution to enhancement in NLO effects. Second to characterize the effect of the ENZ wavelength at normal incident and the importance of the reduced group velocity and enhanced light-matter interaction.

5.3.1. Enhanced Absorption Mechanism for Obliquely Incident Light at ENZ

The loss mechanism is the main limitation in plasmonics and metamaterials where typically metals are used for their plasma oscillations. There have been many methods to control the loss in these materials, but one of the most practical approaches is to use highly doped semiconductors in place of metal [262]. This is simply because the Fermi level in highly doped semiconductors is much lower than in metals, and the density of states is an order of magnitude lower than in metals [262]. Also, in highly doped semiconductors, \( \omega_p \) is an order of magnitude lower than in metals, making them useful in the near-IR and mid-IR ranges where losses are not that large compared to in the visible. However, in this work, we utilize the lossy nature of TCOs to an extent to obtain large optical nonlinearities from such materials.

As we mentioned earlier when light interacts with metals, the absorption of the laser energy through photon-electron interactions makes a non-equilibrium phase between electrons and lattice which results in an energy transfer mechanism. The photon absorption arises both from the excitation of single electrons in the solid and also from the excitation of collective oscillations called plasmons. These effects are indeed due to the electron loss mechanism in the material, which can
be attributed to the scattering by the crystal lattice, \textit{i.e.} electron-phonon scattering. Similarly, there are also random collisions with other electrons, \textit{i.e.} electron-electron scattering, which also contributes to the loss mechanism. This is the same principle behind the electron energy loss spectroscopy (EELS) where electrons are excited via a beam of monoenergetic electrons, and the energy of the transmitted and reflected beam is measured to deduce information from the solid [263-265].

In EELS, there are, in general, many types of interactions including phonon excitations, interband transitions, intraband transitions, collective plasmon excitations, and Cherenkov radiation [232, 266]. The collision mechanism has a time constant, \( \tau \), that can also be considered as a mean free time of electrons under any collision that results in a loss mechanism. Similar to EELS, when a photon is absorbed by free-electrons, a non-equilibrium between electrons and phonons (\textit{i.e.} lattice) is formed. In this process, free-electrons absorb the energy and the lattice remains cold, meaning there is no energy transferred to the lattice. In a short timescale, in the order of picoseconds, the energy is transferred to the lattice by electron-lattice (phonon) collisions. This collisional process modifies the response of electrons to an external field and must, in general, be taken into account. We first briefly review the light-matter interaction in such systems.

5.3.1.1. The significance of the plasma frequency in the interaction of light with free-electrons

At normal incidence, the externally applied light is polarized parallel to the interface and electrons oscillate in the transverse plane. At ENZ, the response of free-electrons to a transverse electromagnetic wave is a collective oscillation of many electrons moving in phase over an extended range mainly because the wavelength inside the material diverges, \( \lambda = \lambda_0/n_0 \to \infty \). These oscillations occur in the presence of an externally applied electromagnetic field, and there is no restoring force after the field is removed [66]. This is not the case when the externally applied
electromagnetic field is at oblique incidence, where it includes a longitudinal component of the electric field. When free-electrons are oscillating in a longitudinal direction, the collective oscillation of electrons maintains even after removing the external field. In other words, these oscillations are the normal modes (eigenfunctions) of the equation of motion governing on the cooperative collective oscillation of electrons. To clarify this, consider a system of electrons in a material with thickness \( L \), as shown in Figure 5-12, excited by an electromagnetic wave at oblique incidence and transverse magnetic (TM) polarization.

\[
\vec{P} = -\frac{(NeA\Delta z)L}{AL} = -Ne\Delta z, \quad (5-38)
\]

where \( e \) is the charge of electron. Now, we can write the equation of motion of an electron in the polarized region in terms of \( E = P/\varepsilon_0 \) as
The above equation describes a simple harmonic oscillator with characteristic frequency (eigenvalue) of \( \omega_p^2 = N e / m^* \varepsilon_0 \). This is important because it emphasizes that the oscillations of electrons in the longitudinal direction becomes a natural mode of the system at the plasma frequency even after removing the externally applied electromagnetic field. Considering that the total longitudinal electric field inside the medium, \( E_L \), is sum of the external longitudinal field, \( E_L^{ext} \), and the induced longitudinal field inside the medium, \( E_L^{ind} \), we can write

\[
E_L = E_L^{ext} + E_L^{ind}
\]  

(5-40)

Now, based on the Maxwell’s equations and the macroscopic Gauss’ law \([16, 17]\), we write the total charge density as

\[
\varepsilon_0 \nabla \cdot E_L = \rho_{total} = \rho^{ext} + \rho^{ind}
\]  

(5-41)

This equation only holds for longitudinal electric field, which accumulation of charge in the volume is the source of the field. Note that the source of a transverse wave is the time variation of a current. Thus, for a plane wave, we get

\[
 ik_L \varepsilon_0 E_L(\omega) = \rho^{ext}(\omega) + \rho^{ind}(\omega),
\]

(5-42)

where \( \omega \) is the angular frequency, \( E_L \) is the magnitude of the longitudinal electric field, and \( k_L \) is the magnitude of the longitudinal wavevector. The induced charges in the material, \( \rho^{ind}(\omega) \), is a consequence of the induced charge fluctuations that will produce a field. In a metallic system, like a Drude material which we addressed earlier, this field is generated because of the rapid oscillations of the electron density, called plasma oscillation and also known as Langmuir waves, after
Irving Langmuir [267]. The quasiparticle resulting from the quantum of plasma oscillation is called the plasmon, and the induced charge can also be found based on the induced polarization via

\[ \rho^{\text{ind}}(\omega) = -ik_L P_L(\omega), \quad (5-43) \]

where \( P_L(\omega) = \varepsilon_0(\varepsilon_L(\omega) - 1)E_L(\omega) \). Similarly, for the external field, we can write

\[ \rho^{\text{ext}}(\omega) = \varepsilon_0 i k_L E_L^{\text{ext}}(\omega) \quad (5-44) \]

By substituting Eq. (5-43) and (5-44) in Eq. (5-41), we conclude that

\[ ik_L \varepsilon_0 E_L(\omega) = \varepsilon_0 i k_L E_L^{\text{ext}}(\omega) - i k_L \varepsilon_0 (\varepsilon_L(\omega) - 1)E_L(\omega), \quad (5-45) \]

which results in

\[ E_L^{\text{ext}}(\omega) = \varepsilon_L(\omega) E_L(\omega), \quad (5-46) \]

This equation exhibits that there can be a sustained oscillation in the longitudinal field unless \( \varepsilon_L(\omega) = 0 \). It also explains the extent to which the external field is reduced (i.e. screened) by the charge of the system. These oscillating electromagnetic fields do not propagate and only act as a restoring force. In an isotropic material, \( \varepsilon_L(\omega) = \varepsilon(\omega) \), and thus Eq. (5-45) is identical to the boundary condition governing on the longitudinal displacement field. However, in anisotropic materials the difference needs to be taken into account. For example, it is imperative to consider the longitudinal dielectric constant in hyperbolic metamaterials where loss can be controlled separately in the longitudinal and transverse direction [245].

The conservation law of energy, omitting the magnetic field, for a system of charged particles and electromagnetic fields can be described via the Poynting theorem as [17]

\[ \frac{\partial W}{\partial t} = -\frac{1}{2\varepsilon_0} E \cdot \frac{\partial}{\partial t} D^*, \quad (5-47) \]
where $W$ is the total electromagnetic energy density averaged over time and stored in the material and $D$ is the displacement field. By taking a real part and integrating the above equation, we get the dissipated power, $W_d$, as

$$w_d = -\frac{1}{2\varepsilon_0} Re \left\{ \int_{-\infty}^{+\infty} dt E \cdot \frac{\partial}{\partial t} D^* \right\}$$  \hspace{1cm} (5-48)

Now by writing $E$ and $D$ based on the Fourier components for a monochromatic plane wave, we get

$$w_d = -\frac{1}{2\varepsilon_0} Re \left\{ \int_{-\infty}^{+\infty} dt \int_{-\infty}^{+\infty} d\omega_2 \int_{-\infty}^{+\infty} d\omega_1 i\omega_2 E_{\omega_1} D_{\omega_2}^* e^{i(\omega_1-\omega_2)t} \right\}, \hspace{1cm} (5-49)$$

where $\omega_1$ and $\omega_2$ are integration dummy variables. This integration is non-zero only when $\omega_1 = \omega_2$, and thus

$$w_d = -\frac{1}{2\varepsilon_0} Re \left\{ \int_{-\infty}^{+\infty} i\omega E_{\omega} D_{\omega}^* d\omega \right\}, \hspace{1cm} (5-50)$$

Here, $E_{\omega} = E_L(\omega) = \varepsilon^{-1}(\omega) E_L^{ext}(\omega)$ is the total longitudinal electric field and $D_{\omega} = D_L(\omega) = \varepsilon_0 \varepsilon(\omega) E_L(\omega)$. Therefore,

$$w_d = -\frac{1}{2} Re \left\{ \int_{-\infty}^{+\infty} i\omega \frac{\varepsilon(\omega)}{|\varepsilon(\omega)|^2} |E_L^{ext}|^2 d\omega \right\}, \hspace{1cm} (5-51)$$

Note that by taking the real part operator into the integrand, we get

$$w_d = -\frac{1}{2} \int_{-\infty}^{+\infty} i\omega \frac{\varepsilon''(\omega)}{|\varepsilon(\omega)|^2} |E_L^{ext}|^2 d\omega, \hspace{1cm} (5-52)$$

where we can define the electron energy loss spectrum of the metal as

$$L(\omega) = \frac{\varepsilon''(\omega)}{|\varepsilon(\omega)|^2} = -Im \left\{ \frac{1}{\varepsilon(\omega)} \right\}, \hspace{1cm} (5-53)$$

$122$
Thus,

\[ w_d = -\frac{1}{2} \int_{-\infty}^{\infty} i\omega L(\omega)|E_L^{\text{ext}}|^2 d\omega, \]  

(5-54)

In EELS, \( L(\omega) \) determines the probability of inelastic scattering event. In electromagnetics, this function determines the response of the electrons in solid to an electromagnetic perturbation. It should now be strongly emphasized that the above results have been derived on the assumption that the optical properties of metals near the plasma frequency are essentially determined by the free-electron contributions. The electron energy loss function has a great importance at the plasma frequency. The frequency dependence of \( L(\omega) \) for a Drude material, follows a Lorentzian with a maximum at the plasma frequency, which can be described as below

\[ L(\omega) = \frac{\omega \omega_p^2 \gamma}{(\omega^2 - \omega_p^2) + \gamma^2 \omega^2}, \]  

(5-55)

The spectral dependence of \( L \) of our ITO sample is plotted below.

Figure 5-13. (black) Real and (red) imaginary part of the Drude permittivity for \( \sim320 \) nm ITO versus wavelength. The blue curve is the associated electron energy loss function for this material fitted by a Lorentzian with a maximum at \( \omega_p \) (\( \lambda_{\text{ENZ}} \approx 1240 \) nm) and width of 113 nm.

The above analysis shows that at oblique incidence, a TM polarized light represents an enhanced coupling to longitudinal plasmons for materials with vanishing real part of permittivity. Thus, the
material exhibits and enhanced absorption at ENZ, which gives rise to an enhanced thermalization of the electrons, and thus a larger induced change in the plasma frequency. We note that this phenomenon has been reported and utilized for many different applications from 100 years ago. The first explanation of such characteristics was presented by Fermi in 1923 [268] and expanded later for metals by Fröhlich and Pelzer in 1955 [269]. A similar mechanism was found for polaritonic materials in the IR and mid-IR attributed to the longitudinal phonon polaritons by Berreman in 1963 [270]. This characteristic of metals can explain different phenomena observed in plasmonics and metamaterials such as perfect absorption [271-273], ENZ mode coupling into thin films [274-277], and enhanced harmonic generation and NLO effects from ENZ materials [3, 215, 277, 278].

5.3.1.2. Modeling of Fresnel coefficients in a three-layer material at ENZ

The enhanced coupling to longitudinal plasmons, described by the electron energy loss function, can be described by macroscopic electromagnetic wave parameters, and therefore, we can use wave equations to study this effect. We present a formulation for the linear transfer matrix in APPENDIX A for a three-layer system, where the first medium is air ($n_a = 1$), the second medium is a lossy film (complex index $\eta_f = n_f + i\kappa_f$), and the third medium is a semi-infinite dielectric, e.g. glass ($n_s = 1.5$). The structure that we consider is shown below, where the first interface is at $z = 0$ and the second interface is at $z = \ell$, and the angle of incident is $\theta_i$. 

124
We derived the transmission, \( t(\lambda, \theta_i) \), and reflection, \( r(\lambda, \theta_i) \), coefficients for the field in two cases of TM and TE polarization, and extracted the absorption coefficient, \( A(\lambda, \theta_i) \), from these parameters (See APPENDIX A). It is essential to note that these parameters can be derived for electromagnetic wave intensity, however, in this case interference effects within the film will be omitted. Associated Fresnel coefficients of the field can be described as

\[
\begin{align*}
    r(\lambda, \theta_i) &= \frac{r_{12} + r_{23}e^{i2\delta}}{1 + r_{12}r_{23}e^{i2\delta}}, \\
    t(\lambda, \theta_i) &= \frac{(1 + r_{12})(1 + r_{23})e^{+i\delta}}{1 + r_{12}r_{23}e^{i2\delta}},
\end{align*}
\]

(5-56)  

(5-57)

where \( r_{ij} \) and \( t_{ij} \) (\( i,j = 1,2,3 \)) are the reflection and transmission coefficients of the field at the interface between layers \( i \) and \( j \) for either polarization, respectively, and \( \delta = k_{f,\perp}\ell \), where \( k_{f,\perp} \) is the normal component of the wavevector in layer 2. The reflection and transmission of the light intensity can be derived from

\[
R_{TM,TE}(\lambda, \theta_i) = |r_{TM,TE}|^2.
\]

(5-58)
\[
T_{TE}(\lambda, \theta_i) = |t_{TE}|^2 Re\left\{\frac{k_{\perp}}{k_{\parallel}}\right\} \quad \& \quad T_{TM}(\lambda, \theta_i) = |t_{TM}|^2 Re\left\{\frac{k_{\perp}}{k_{\parallel}} \times \frac{\varepsilon_s}{\varepsilon_a}\right\}
\]

Finally, the absorption can also be calculated as
\[
A(\lambda, \theta_i) = 1 - R(\lambda, \theta_i) - T(\lambda, \theta_i).
\]

We calculated and plotted the normalized reflection, transmission, and absorption of electromagnetic waves in a sample with similar structure to Figure 5-14 where the lossy material is considered to have similar permittivity as ITO, which we measured using ellipsometry.

Figure 5-15. Transmission, reflection, and absorption versus angle of incident and wavelength for TM polarization (top) and TE polarization (bottom). Absorbed power exhibits an enhanced feature at oblique incident angles close to the ENZ wavelength.

We note that the substrate of our sample is a ~1.1 mm thick glass, however, since the thickness of this layer is significantly larger than our wavelengths, we can consider this as a semi-infinite glass substrate in our modeling. Figure 5-15 represents the transmission, reflection, and absorption of the ITO film versus angle of incident and wavelength for TM and TE polarizations. The absorption enhances significantly at oblique incidence \((20^\circ < \theta_i < 50^\circ)\), which can be attributed to the en-
hanced coupling of the electric field to longitudinal plasmons as discussed previously. The enhanced absorption of the TM polarization at oblique incident results in an enhanced excitation mechanism to redistribute the carriers and subsequently to improve NLO effects.

This behavior is not observed for TE polarization where there is no longitudinal electric field. The enhanced absorption at TM polarization is the direct consequence of loss mechanisms in such materials. We, therefore, calculated $A(\lambda, \theta_i)$ for TM polarization and plotted the results for three different cases where we considered $\text{Im}(\varepsilon)$ to be (a) 2× larger, (b) 2× smaller, or (c) 10× smaller values over entire spectra. The results of this calculation is plotted in Figure 5-16.

![Figure 5-16. $A(\lambda, \theta_i)$ for (a) 2× larger, (b) 2× smaller, or (c) 10× smaller values of $\text{Im}(\varepsilon)$ for all wavelengths.](image)

From these results, we can conclude that the magnitude of the absorbed power will be larger for higher loss materials; however, a narrower resonant behavior is expected when loss is smaller. This gives us more insight on how to utilize loss and design more efficient ENZ-based NLO devices based on the desired performance and application.

We note that, it is also essential to consider interference effects in our model. This becomes apparent in Figure 5-17, where we plot $T$, $A$ for various incident angles for TM polarization. For instance, the magnitude of the absorbed power at 50° is under estimated by ~ 15% when interference is ignored.
Figure 5-17. $R$ (black), $T$ (red), and $A$ (blue) calculated for different incident angles for TM polarization while (a) interference effects are not included, i.e. Fresnel coefficients for intensity and (b) interference effects are included, i.e. Fresnel coefficients for field, as discussed in APPENDIX B.

The significant effect of interference in Fresnel parameters can be an indication of the contribution of collective oscillation of electromagnetic waves. In other words, if we consider a two-layer system with a semi-infinite metal as the second medium, this behavior will not be observed. This is also the case for our analysis based on the electron energy loss function in previous section, where we considered a bounded medium comprising of electrons (thickness of $L$) and found the natural characteristic frequency of the equation of motion. This “duality” between oscillation of electrons and oscillation of electromagnetic wave, is the main characteristic and core advantage of ENZ materials, which is also sometimes called the ENZ mode for electrons (or plasmons) and Berreman mode for phonons (or polaritons) [274, 277]. This enhanced absorption at oblique incident for TM polarization can be utilized to increase NLO effects such as NLR and harmonic generation [3, 87, 92, 122, 215, 260]. To control this absorption, it is possible to engineer the doping to obtain an optimal loss where the enhancement in the normal electric field component gives the largest absorption in the film.
5.3.1.3. Beam-Deflection Measurements at Oblique Incidence

We performed nondegenerate BD measurements with the probe beam centered at 775 nm and the pump beam in the range of 1150-1400 nm to cover the ENZ spectral region. For these measurements, we used the fundamental output of our Clark-MXR CPA system as the probe beam. The pump beam is generated from the signal beam of an OPA system and spectrally filtered with 10 nm (FWHM) bandpass filters at each wavelength. These measurements show us how differently the effective mass and subsequently the plasma frequency will be altered for each pump wavelength. We, therefore, fix our probe wavelength at 775 nm to isolate our studies to different carrier excitation conditions. The measurements were performed for two incident angles of $\theta_i = 6^\circ$ (close to normal incident) and $\theta_i = 30^\circ$ to study the effect of enhanced coupling to longitudinal oscillation of plasmons. We rotated the sample to obtain these incident angles, thus both pump and probe beams were illuminated at $\theta_i$. Identical measurements were also performed on a piece of bare substrate and sample (film + substrate) for each condition. For each pump wavelength, the pulse-width was measured via the autocorrelation technique. Since we also measured the BD signal of the substrate, we do not need an absolute characterization of the pulsewidth to find the irradiance, and we can scale our ITO measurements relatively to that of the glass substrate to calibrate the setup and eliminate unwanted errors. However, we measure the pulsewidth to have a rough estimate of the NLR coefficients that we extract from our BD measurements on the substrate. We, therefore, used the absolute irradiance to fit our BD measurements of the substrate and found that values of NLR coefficient, $n_{2SiO_2}$, is very close to literature values [104, 183]. The contribution of the substrate to the NLR is different at each irradiance due to linear loss and the NLA effects of the ITO film (thin film is on the front surface). Thus, we used a similar approach that we used in §5.3. to extract the effect of the substrate. It is important to note that, it is also possible to use
nonlinear transfer matrix methodology to extract nonlinear Fresnel coefficients and accurately determine the contribution of the glass substrate to the BD signal. However, it is much simpler to take advantage of the abilities provided to us with the BD technique and eliminate the contribution of the substrate by only utilizing the polarization-dependence of the measured signals, as discussed in §5.3.

Figure 5-18. BD measurements for $\lambda_e = 1250$ nm and $\lambda_p = 775$ nm on (a) ITO + substrate and (b) bare glass substrate for various input beam irradiances. Irradiance dependence of the measured $\Delta E/E(\%)$ for (c) ITO at two different delays where signal is maximum (A) and at a delay where the contribution of substrate is zero (B), and (d) the bare substrate.

In Figure 5-18, we present BD measurements of the ITO sample and bare substrate for various incident light irradiances at normal incident. We also plot the irradiance dependence of the $\Delta E/E$ to study the scaling of NLR versus irradiance. We also attempted to fit the data by first fitting the magnitude and decay of the signal at very long delays ($\sim 1500$ fs) by considering a single rise and fall exponential function (red solid curve in Figure 5-19(a-c)). This contribution is due to
the electron-phonon scattering mechanisms and the decay time is relatively long for this process (*i.e.* slow decay mechanisms > 1.5 ps). Next, we fitted the signal with faster decay at shorter delays with a similar response function (black solid curve in Figure 5-19(b-c)), and finally an instantaneous function following cross-correlation of pump and probe pulses to fit the signal purely from the glass substrate.

Figure 5-19. The fitting procedure for each mechanism contributing to the BD measurements. The red solid line in (a-c) is the contribution of the electro-phonon mechanisms, the black solid line in (b-c) is the contribution of the electron-electron mechanisms. The sum of the two effects is the blue solid line in (c-d). In (d) we plot the contribution of ITO (blue) and glass (red) and sum of these two effects (black). The steps of this fitting process are shown in Figure 5-19. We note, however, that this fitting procedure and the previous one based on eliminating the substrate give us almost identical results as expected. The second method can be more insightful, particularly, in the case of high irradiances, where an unusual dip is observed at longer delays, see Figure 5-18(a) at $I_e = 98.6 \text{ GW/cm}^2$. This dip is mainly due to the fact that at high irradiances the signal from ITO is saturated; however, the third-order bound electronic signal from the substrate grows linearly with the irradiance, as shown in Figure 5-18(c, d), and apparent in Figure 5-20(b, c).
Figure 5-20. BD measurements and fit for (a) various irradiances and (b,c) two irradiances where an unusual dip is observed at delay time of ~250 fs.

We, therefore, used this procedure to fit our experimental data at the probe wavelength $\lambda_p = 775$ nm and pump wavelength $\lambda_e = 1250$ nm and show our results in Figure 5-20. We separately plotted fits for $I_e = 77.8$ GW/cm$^2$ and $I_e = 98.6$ GW/cm$^2$ to show how the peak of the blue and black solid fit lines are temporally separated.

We also directly calculated $\Delta n_{\text{eff}}(\tau_D) = \Delta \phi(\tau_D)/k_0L$ from our BD measurements, where $k_0$ is the wavevector in free-space at $\lambda_p = 775$ nm and $L \approx 320$ nm. We plotted our results in Figure 5-21, where we plot the extracted values of $\Delta n_{\text{eff}}$ from our fits to the BD signal. We observe that $\Delta n_{\text{eff}}$ reach values much smaller than unity ($\Delta n_{\text{eff}} \approx 0.035$ at the highest irradiance), as opposed to our slightly nondegenerate measurements, which we presented in §5.3. and observed $\Delta n_{\text{eff}}$ close to unity. Additionally, at $\lambda_p = 775$ nm, the value of the index of refraction is $n_0 \approx 1.5$, and thus $\Delta n_{\text{eff}} \ll n_0$. 

132
We note, however, that $\Delta n_{\text{eff}} \sim 0.05$ is still extremely large compared to other typical NLO materials with very large NLR coefficients, such as AlGaAs [279] and chalcogenide glasses such Arsenic triselenide (As$_2$Se$_3$) [280] where $\Delta n_{\text{eff}} \sim 10^{-4} - 10^{-2}$ is attainable for relatively higher irradiances over a longer propagation length. This large NLR is inherently because of carrier-related NLO effects; however, the true advantage of TCOs at ENZ is their sub-picosecond fast response time. Since $\Delta n_{\text{eff}} \ll n_0$, we can consider NLR as a perturbation and expand it versus irradiance. This let us to define $n_{2,\text{eff}} = \Delta n_{\text{eff}}/I_0$ at zero-delay where pump and probe pulses overlap completely and for those irradiances that $\Delta n_{\text{eff}}$ scales linearly with irradiance. We use this condition because it mimics the condition in single-beam techniques such as Z-scan.

Figure 5-22(a-e) represents the BD measurements for $\lambda_p = 775 \, nm$ and $\lambda_e = 1150, 1200, 1250, 1300$, and $1350 \, nm$, respectively. The blue solid line are the fits to the contribution of ITO, the black solid lines are the fit to the contribution of the glass substrate, and the red solid line are the fits to the total signal. We also plotted $n_{2,\text{eff}}$ for different pump wavelengths in Figure 5-22(f).
Figure 5-22. (a-e) BD measurements on a thin piece of ITO deposited on a glass substrate at $\lambda_{\text{probe}} = 775 \text{ nm}$ and $\lambda_{\text{pump}} = 1150 - 1300 \text{ nm}$. (f) The spectral dependence of the extracted effective NLR compared to the absorption spectra calculated at normal incident.

Figure 5-22(f) shows the correlation between the values of $n_{2,\text{eff}}$ at each wavelength and the absorbance of the film at the normal incidence. The discrepancy between the absorption spectra and
the effective NLR is because these two are dependent on each other non-linearly. However, they agree qualitatively because as the absorption increases, more electrons will be thermalized and thus NLR will be larger. We note that this is not an exact theory to fit the values of $n_{2,\text{eff}}$, but it can be very useful to understand the scaling of the NLR.

Additionally, we plotted the absorbance of the sample for both TE and TM polarized incident light at different incident angles in Figure 5-23(a, b). As expected and previously mentioned, we observe a resonant absorptive behavior at ENZ when the TM-polarized light is obliquely illuminating the film. This behavior is due to both the loss in the material and the enhancement of the normal component of the electric field, which we explained by the enhanced coupling to the longitudinal plasmons in detail.

Figure 5-23. Absorption spectra for (a) TE and (b) TM polarized light at various incident angles. (c) The correlation between the calculated absorption spectra and extracted NLR values from the BD measurements at ~ 6° and 30°.

Figure 5-23(c) shows the spectral dependence of the effective NLR extracted from the BD measurements close to normal incidence as well as at ~ 30° for a TM-polarized pump. Same as before, the $n_{2,\text{eff}}$ is the ratio of the induced change in the index, $\Delta n_{\text{eff}}$, at zero-delay to the peak irradiance of the excitation beam. The induced $\Delta n_{\text{eff}}$ reaches the value of $\Delta n_{\text{eff}} = 0.07 \pm 0.01$ for the highest irradiance at 30° close to the ENZ wavelength at $\lambda_p = 1250$ nm. Here, we can conclude that longitudinal electron oscillations result in an absorption peak at ENZ for TM-polarized oblique
incident light. This behavior enhances the redistribution of electrons in the conduction band and consequently increases the NLR. This enhancement is not extremely large; however, it can be utilized for increasing excitation of carriers and subsequently improving the NLO performance [3, 87, 241, 260, 277].

5.3.1.4. Nonlinear transmission and reflection

As we mentioned earlier, the nonlinear enhancement effects of the ITO are so large that can greatly alter the input irradiance. Precise characterization of the irradiance after the first interface between air and ITO is required to find the NLR coefficient. These multiple enhancements make the analysis complicated, particularly because the Fresnel coefficients will also be altered significantly. In these cases, transmission and reflection from the film will be greatly dependent on both NLR and NLA. Thus, in addition to BD measurements, we have performed time-resolved pump-probe measurements collecting the total transmitted and reflected beams at each condition. These measurements, by itself, can be used to find the NLR and NLA using nonlinear transfer matrix method, but the advantage of BD is that it directly gives us the magnitude, sign, and transients of NLR. For instance, at larger irradiances, additional scattering mechanisms contribute to the NLR such as those related to the ballistic electron scattering mechanisms. Below we see BD measurements together with pump-probe measurements at transmission and reflection when pump wavelength is 1150-1350 nm and probe wavelength is 775 nm. The $\Delta n$ is positive, hence the index is increasing from the initial value of 1.522 for probe at 775 nm. Therefore, the reflection is increasing and transmission is decreasing. The complex dynamics of the NLR at irradiances higher than ~ 100 GW/cm$^2$ can be clearly seen in the BD measurements as opposed to the other two measurements. Experimental signals obtained using BD and nonlinear transmission and reflection
at each wavelength are reported in the next three pages. For each wavelength, we also performed BD measurements on a 1 mm thick fused silica to calibrate our experimental setup and accurately determine the irradiance.

Figure 5-24. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1150 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row.

Figure 5-25. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1200 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row.
Figure 5-26. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1250 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row.

Figure 5-27. First row) BD, reflection, and transmission measurements performed on ITO at pump wavelength 1300 nm and probe wavelength of 775 nm and BD measurements on fused silica. Second row) irradiance dependence of signals plotted in 1st row.
In this section, we present nondgenerate BD measurements of ITO where the probe beam is tuned across the ENZ region and the pump beam is kept far away from ENZ fixed at $\lambda_e = 775$ nm. It is important to note that in this case, the condition in which carriers are excited is identical for all probe wavelengths, as opposed to the previous section where we studied different conditions of carrier excitations by fixing the probe wavelength. This is fundamentally different than the previous section, particularly because we only study how ENZ condition affects the propagation of the probe beam. In simple words, upon a given change in $\omega_p$, the Drude dispersion will be shifted; however, this manipulation will be different at each frequency. This can be clarified further by considering a “tiny” change in material susceptibility, $\chi_{0,p}$, at frequency $\omega_p$ of the probe, in which the induced changes in the complex refractive index, $\eta(\omega_p)$, in its general form while considering nonlinearities as a perturbation, is given by
\[ \eta(\omega_p) = n(\omega_p) + i\kappa(\omega_p) = \sqrt{1 + \chi_{0,p} + \Delta\chi_p}, \quad (5-60) \]

where \(\Delta\chi_p\) is the nonlinear change in susceptibility at \(\omega_p\) induced by the pump beam at \(\omega_e\). Considering a tiny \(\Delta\chi_p \ll \chi_{0,p}\), we can expand the above equation to

\[ \eta(\omega_p) \approx \eta_{0,p} \left(1 + \frac{1}{2n_{0,p}^2} \Delta\chi_p\right) = \eta_{0,p} + \frac{1}{2(n + i\kappa)} \Delta\chi_p, \quad (5-61) \]

Therefore, the induced change in the complex index can be described as

\[ \Delta\eta(\omega_p) = \frac{1}{2(n + i\kappa)} \Delta\chi_p, \quad (5-62) \]

This means that nonlinear behavior of the complex index depends upon both the nonlinear susceptibility induced by the pump and the linear parameters of the material at the probe wavelength. For materials with very small refractive index (while loss is ignored), the magnitude of \(\Delta\eta\) enhances dramatically. Therefore, both real part of the index (NLR) and imaginary part (NLA) will experience huge changes in proximity of the ENZ region. We note that at ENZ, \(\text{Re}(\epsilon) = n^2 - \kappa^2 = 0\), therefore, for a relatively low imaginary part, \(\kappa\), the real part of the refractive index is small and thus the enhancement factor in NLR is larger. It is important to note that the behavior of \(\Delta\chi_p\) depends upon the origin of the nonlinear response, which for ITO is dominated by carrier induced nonlinearities. For semiconductors, the dominant mechanism is bound-electronic nonlinearities for sub-bandgap photon energies, as we discussed in detail in previous chapters [14].

This change in complex index is manifested in experimental measurements as induced nonlinear phase-change and nonlinear transmission. In fact, in all nonlinear spectroscopic measurements, the measurable parameters are the irradiance and phase of the interacting beam. Therefore, it is not possible to directly measure \(\Delta n\) and \(\Delta\kappa\); however, for typical small NLO effects, where perturbative analysis is valid, we can extract NLR and NLA directly from measurement of
nonlinear phase-change and nonlinear transmission. We use our BD method which is remarkable in a sense that it can directly measure the nonlinear phase change, \( \Delta \varphi \), and its temporal dynamics and polarization dependence without resorting to sophisticated data analysis such as nonlinear transfer matrix.

5.3.2.1. Beam-Deflection and Cross-phase Modulation Measurements

We used our BD method to measure nondegenerate nonlinearities and to effectively limit the NLO enhancement to circumstances where the probe pulse is tuned across the ENZ, where the nonlinear phase change is enhanced, while avoiding enhancement associated with the pump pulse being at ENZ (i.e. where the excitation of carriers is increased). For these measurements, the pump pulse is centered at 775 nm, generated directly from our Clark-MXR CPA system. The pulsewidth of the pump at the output of the laser is \(~ 150\) fs (FWHM), and it is slightly chirped while propagating through various polarizing elements. The pulsewidth of the 775 nm pump beam at the sample plane is measured to be \(~ 154\) fs (FWHM). We did not spectrally filter this beam, since the bandwidth was initially narrow and around \(~ 6\) nm, which is expected for bandwidth-limited Gaussian pulses. The probe beam was generated from an OPA system and tuned across the ENZ region. For those measurements where the probe wavelength was between 1100 – 1550 nm, we directly used the signal beam, however, for the measurement performed at 950 nm, we used a BBO crystal to generate second harmonic of the idler beam at 1900 nm. We measured the beam size of the probe beam using InGaAs camera and fit the vertical and horizontal cross-sections with Gaussian distribution. For each probe wavelength, we found the focus of the probe beam and placed the sample at that position. We also measured the pump beam size and attempted to control the beam size as much as possible to have identical pump beam size at each probe wavelength and subsequently
perform experiments under identical pump conditions. We also verified that our experiments are identical by separately measuring BD signals from a piece of fused silica substrate and extracted very similar values of $n_{2,el}$ for the fused silica sample in agreement with the previously reported values in the literature [104, 183].

The contribution of the substrate is eliminated with a similar approach as discussed in §5.3, in which we utilize the polarization anisotropy of the bound-electronic nonlinearities of the substrate to effectively subtract its contribution to the BD signal. We note, however, that for this case we can also subtract the contribution of a separately measured bare substrate directly from the BD measurements of the ITO film on the substrate. This is because the pump beam (centered at 775 nm) does not experience linear and nonlinear losses as big as the case where the pump is around the ENZ wavelength. Therefore, the irradiance of the pump beam reaching the back surface of the ITO film (when film is on the front surface of the sample) will not be significantly changed, and the total signal can be considered as the superposition of each individual measurement from ITO and bare substrate. Figure 5-29(a) presents BD measurements for $\lambda_p = 1250$ nm and $\lambda_e = 775$ nm. We calculated $\Delta\varphi$ directly from our BD measurements, $\Delta E/E$, and fit the signal with a single exponential rise and decay convolved with the cross-correlation of the pump and probe pulses.
We also simultaneously measured the optically-induced changes in probe wavelength, as shown in Figure 5-29(b), via cross-phase modulation measurements. For these measurements, we used a fiber coupled optical spectrum analyzer (OSA) to collect the spectrum of the probe beam while tuning the delay between pump and probe pulses. We measured the spectrum at negative delays, before probe pulse reaches the pump pulse. We also measured the spectrum at delays where maximum spectral shift, $\Delta \lambda_p$, or maximum phase-change, $\Delta \varphi$, occurs. We became interested in doing these measurements, since it is possible to directly predict the expected shift in the wavelength from BD measurements while measuring this shift directly from cross-phase modulations. The two measurements together provide us with a remarkable understanding of both temporal and spatial NLO effects altering both nonlinear phase and frequency of the probe beam. Briefly explaining, spatial modification of material index manipulates the wavevector (i.e. self-focusing or self-defocusing), while temporal alteration of the material index induces a shift in optical frequency (i.e. self-phase modulation or cross-phase modulation). The latter is sometimes referred to as the time-refraction effect [281-285], where the spectrum of an optical pulse can change inside
a medium whose refractive index changes with time. We note that the spatial change in index can be understood by considering an interface, in which the refractive index changes from \( n_1 \) to \( n_2 \), which subsequently results in a shift in wavevector from \( k_1 = n_1 \omega / c \) to \( k_2 = n_2 \omega / c \), while frequency is conserved. In highly nonlinear materials, such as ITO, the refractive index changes both temporally and spatially, and thus both wavevector and frequency will be manipulated. These “spatiotemporal” modifications in refractive index can also be described by considering that phase is a function of both spatial coordinates, \( \vec{r} \), and temporal coordinates, \( t \), and described by \( \varphi(\vec{r}, t) \).

Therefore, the instantaneous wavevector and frequency of light can be deduced as

\[
\vec{k} = \frac{\partial}{\partial \vec{r}} \varphi(\vec{r}, t), \tag{5-63}
\]

\[
\omega = -\frac{\partial}{\partial t} \varphi(\vec{r}, t) \tag{5-64}
\]

Eqs. (5-63) and (5-64) indicates that a spatial nonlinear phase-change manipulates wavevector (photon momentum) and optically-induced temporal nonlinear phase-change alters the frequency (photon energy) of the beam.

We, therefore, calculate the expected \( \Delta \omega \) (or \( \Delta \lambda \)) from our BD measurements. The optically-induced wavelength change can be easily determined from \( \Delta \lambda / \lambda = -\Delta \omega / \omega \). We plotted \( \Delta \varphi \) and \( \Delta \lambda \) together, as shown in Figure 5-30. The cross-phase modulation data are plotted at five different delays, where the red data point refers the largest measured red shift and the blue data point refers to the largest measured blue shift. This condition happens at a delay where the magnitude of \( \partial \Delta \varphi / \partial t \) is maximized.
We measure $\Delta \lambda_p \approx 5$ nm at $\tau_D = -100$ fs and $\Delta \lambda_p \approx -2$ nm at $\tau_D = 380$ fs. We note, however, that since the material is very thin, the magnitude of this shift is limited by the group velocity mismatch (GVM) between excitation and probe pulses and the temporal dynamics of the nonlinear phase change [6]. The shift in frequency may be enhanced by using shorter pulsewidths resulting in a faster nonlinear phase accumulation. Our results, shown in Figure 5-30, are in excellent agreement with our predictions extracted from the time-derivative of $\Delta \varphi$ (black solid line). These two simultaneous measurement techniques directly characterize the large nonlinear effects excited by both optically-induced spatial and temporal modulation of optical parameters. These NLO effects are very large, particularly, because $\Delta \varphi \approx 0.2\pi$ rad and $\Delta \lambda \approx 5$ nm is attainable in a thin subwavelength film. This simultaneous tunability of both frequency and phase may be utilized in all-optical signal modulators enabling switching and frequency synthesizing applications at the same time. We attempted to perform similar measurements for probe beams away from ENZ condition to spectrally characterize the enhancement of the nonlinear optical mechanisms in this region. The additional measurements at various probe wavelengths are summarized in Figure 5-31.
Figure 5-31. BD measurements and the extracted $\Delta \varphi$ for each probe wavelengths from 950 – 1550 nm.

The results of BD and $\Delta \varphi$ under identical pumping conditions are plotted in equal scales to present the enhanced behavior of the nonlinear phase change in proximity to the ENZ condition. In order
to study the spectral dependence of the measured $\Delta \varphi$, let us revisit the Drude model describing the dispersion of TCOs. We calculate $\Delta n_{\text{eff}} = \sqrt{\varepsilon + \Delta \varepsilon} - \sqrt{\varepsilon}$ for an arbitrary $\Delta \omega_p$ and normalize it to the maximum value of $\Delta n_{\text{eff}}$ to obtain the predicted spectral dependence. We plot this together with the extracted values of $\Delta \varphi_{\text{ITO}}/\Delta \varphi_{\text{SiO}_2}$ to eliminate unwanted calibration errors and to scale our measurements to a well-known material, in this case the glass substrate, which has an approximately constant NLR over the spectra of our studies. Since we have normalized our nonlinear phase change measurements, the calculated ratio will be independent of the irradiance of the incident beam below the saturation limit of ITO. This is the main reason that we used an arbitrary $\Delta \omega_p$ to calculate the spectral dependence of the normalized effective NLR. We note that it is also possible to fit the relative measurements and extract $\Delta \omega_p$, which is expected to be constant in all our measurements for identical pump conditions. In fact, Figure 5-32 is indicative of the constant $\Delta \omega_p$, which modifies the Drude dispersion differently for each probe frequency.

![Figure 5-32. Spectral dependence of the measured ratio of the nonlinear phase change in ITO to the glass substrate for an excitation at 775 nm and comparison to the spectral dependence of $\Delta n_{\text{eff}}$ obtained from the Drude model.](image)

We also calculated the expected shift in probe wavelength from each BD measurement and compared the results with the experimentally measured cross-phase modulation data that gives us $\Delta \lambda$ at each probe wavelength for different delays between pump and probe pulses. As discussed
earlier, for each individual probe wavelength a set of measurements were performed to characterize $\Delta \lambda$ at different delays between pump and probe pulses. We attempted to perform these measurements to experimentally characterize the effect of the optically-induced temporal modulation of the nonlinear phase change. We, therefore, plotted the measured values of $\Delta \lambda$ at each delay and compared it with the predictions obtained from BD measurements by calculating $\Delta \omega = -\frac{\partial \Delta \varphi}{\partial t}$ and subsequently $\Delta \lambda$. All of these measurements were performed under identical pump irradiances at $I_e = 521 \text{ GW/cm}^2$, where the material nonlinear response was not saturated. We note that at the peak of $\Delta \varphi$, the temporal change of the phase is minimized ($i.e. \frac{\partial \Delta \varphi}{\partial t} = 0$), and no shifts will be observed for the probe wavelength. However, for increasing and decreasing phase of the probe, $\Delta \lambda$ will be non-zero as shown in Figure 5-34. Upon an increase in the nonlinear phase ($i.e. \frac{\partial \Delta \varphi}{\partial t} > 0$), the probe beam experiences a red shift (maximum where $\frac{\partial \Delta \varphi}{\partial t}$ is maximum) and for a decrease in the nonlinear phase ($i.e. \frac{\partial \Delta \varphi}{\partial t} < 0$), the probe beam experiences a blue shift (maximum where $\frac{\partial \Delta \varphi}{\partial t}$ is minimum). Note that the magnitude of this shift depends upon both the rate of change in phase and also the magnitude of the nonlinear phase change. Therefore, for a material with extremely large nonlinear phase change but slow cumulative response, the shift in the probe wavelength will be small. We found that the rate of change in the phase is very close in all of our measurements at different probe wavelengths, but the magnitude of this change is larger closer to the ENZ wavelength, and therefore, $\Delta \lambda$ is enhanced. We note that this shift does not utilize the full benefit of the large nonlinearities of ITO and it can be further enhanced for degenerate effects where pump and probe beams travel with similar group velocities in the medium [6, 227] and, therefore, the interaction time of two pulses will be optimized. We also note that for
shorter pump and probe pulses, the nonlinear phase change occurs in a shorter scale of time resulting in a bigger rate of change, and the two pulses will, therefore, interact more efficiently within the film. Thus, it is expected to see a larger wavelength shift when using shorter pulsewidths. The downside of using a shorter pulsewidth for the probe is losing the spectral resolution of the NLO enhancement. Additional measurements for different probe wavelengths are summarized in Figure 5-34.

In Figure 5-32, the data point at 775 nm is obtained from degenerate Z-scan measurements analogous to pumping the material at 775 nm and probing at 775 nm, which are shown in Figure 5-33. We scaled the magnitude of the spectral dependence of $\Delta n_{\text{eff}}$ to fit the measured values of the nonlinear phase change. Note that this approach is valid since we normalized nonlinear phase measurement of ITO to that of the substrate, hence factors of $k_0$ and irradiance in the nonlinear phase change will be eliminated, and the ratio will only depend on the effective NLR.

Figure 5-33. Degenerate Z-scan measurements at $\lambda_p = 775$ nm, at $I_0 = 368$ GW/cm$^2$. (a) OA, (b) CA, and (c) CA/OA.

Since the photon energy at 775 nm is closer to half of the bandgap of ITO ($E_g \sim 3.2$ eV [237]) compared to near-IR photons, pumping at $\lambda_p = 775$ nm results in a small interband 2PA effect, as shown Figure 5-33. The 2PA of 775 nm pump slightly limits the enhancement mechanism compared to when the pump wavelength is close to ENZ wavelength, where the excitation mechanism is purely due to redistribution of carriers via intraband processes.
Figure 5-34. (Left axis) Values of \( \Delta \varphi \) extracted from BD measurements (blue data and solid line fits), (right axis) experimentally predicted (black solid line) and experimentally measured (data points) magnitude of \( \Delta \lambda \) imposed upon the probe beam via the pump beam at \( I_e = 521 \text{ GW/cm}^2 \).
Our cross-phase modulations measurements (data points in Figure 5-34) are in an excellent agreement with our predictions (black solid line in Figure 5-34) from the BD measurements. We plotted the spectral dependence of the predicted values of $\Delta \lambda$ from our BD measurements and observed an enhancement in optically-induced wavelength change of the probe beam. The maximum achievable wavelength translation (red shift and blue shift) for a constant pump irradiance are the highest near the 1250 nm (where $\text{Re}(\varepsilon) = 0$) than at other wavelengths. For example, at $\lambda = 950$ nm or $\lambda = 1550$ nm the measured maximum magnitude of the redshift is a factor of four smaller than what can be achieved at ENZ.

![Graph showing spectral dependence of optically induced red shift and blue shift](image)

Figure 5-35. Spectral dependence of the optically induced red shift and blue shift of the probe pulse for a 775 nm excitation irradiance at 521 GW/cm$^2$ predicted from BD measurements.

5.3.2.2. The Importance of Group-Velocity and Phase-Velocity

Experimental techniques to characterize NLO effects typically characterize measurable quantities such as intensity and phase of the transmitted or reflected field. Therefore, when the goal is to characterize material parameters, we use these measurable quantities and extract nonlinearly induced changes in permittivity, refractive index, or absorption coefficient from them. For instance, the simplest relation to find NLR or NLA from an experiment, is to use the simple relation of the complex phase for plane waves, $k_0\eta_0 z$, where $\eta_0 = n_0 + i\kappa_0$ is the linear complex index. For a
propagation length of $L$, the complex phase becomes $\exp(i k_0 n_0 L) \cdot \exp(-\kappa_0 L)$, and therefore, the electric field accumulates a phase of $\varphi = k_0 n_0 L$ and its amplitude attenuates with $\exp(-\kappa_0 L) = \exp(-\alpha_0 L/2)$. In presence of nonlinear effects, $n_0$ and $\alpha_0$ will be modified by adding a positive or negative nonlinear term $\Delta n$ and $\Delta \alpha$ where usually $\Delta n \ll n_0$ and $\Delta \alpha \ll \alpha_0$. The origin of these changes were described by Eqs. (2-56) and (2-57).

This description is valid for bulk materials; however, other factors need to be taken into consideration while studying very thin materials such as 2D semiconductors, monolayer materials, or surfaces where quantum confinement significantly affects the effective length of propagation and typically large phase changes are reported [124]. In such cases, since the physical length of propagation is extremely small, the extracted values of $\Delta n$ can reach unusual values of larger than unity. This description can also become problematic in NLO devices such as waveguides and photonic crystals where modal index of refraction needs to be considered [279, 286-291]. In these cases, the measured quantities are related to the effective mode propagating within the NLO device. For instance, if we have a silicon ridge waveguide or a silicon photonic crystal waveguide, nonlinear phase change will alter the refractive index of silicon by $\Delta n_{Si}$, but the measured NLR obtained directly from $\Delta \varphi$ will be $\Delta n_{mode} = \Delta \varphi/k_0 n L$, which is different than $\Delta n_{Si}$. In such cases, if the goal is to find $\Delta n$ of the material, it essential to first obtain the exact relation between the modal behavior and the material refractive index. We note, however, that the important parameter in NLO devices is the modal NLR, which directly influences the output of the device as defined by $\Delta n_{eff} = \Delta \varphi/k_0 n L$. Since modal behavior plays an important role in NLO device characteristics, we need to consider the dispersion of the mode propagating within that medium while we
extract nonlinear parameters. This becomes even more important at frequencies where the dispersion curve does not exhibit a linear behavior similar to the free-space. At these frequencies, the group-velocity, \( v_g \), and phase-velocity, \( v_p \), are usually significantly different, and the phase-sensitivity upon a change in material refractive index enhances dramatically. This is a well-known phenomenon particularly in photonic crystals where the dispersion of the system can be designed on demand \([279, 286-291]\). In this case, since the induced nonlinear phase change actually depends on the induced shift in wavevector by \( \Delta \varphi = \Delta k \cdot L \), we can argue that the phase change will be greatly enhanced when \( \Delta k = \Delta \omega (\partial k / \partial \omega) = \Delta \omega / v_g \) is enhanced at small group velocities for a given shift in frequency. This means that for a given change in frequency, it is expected to obtain a larger \( \Delta k \) and subsequently an enhanced nonlinear phase change. More precisely, for a given change in material index, \( \Delta n \), we can write the induced shift in frequency while considering a fixed \( k \) by

\[
\Delta \omega \approx \Delta n \cdot \frac{\partial}{\partial n} \left( \frac{ck}{n} \right)_{k=\text{const.}} = -\frac{ck}{n^2} \Delta n = -\frac{\omega}{n} \Delta n,
\]

which is also called the time-refraction relation for an index change in temporal domain, where wavevector is conserved \([137, 281, 282]\). Therefore, the nonlinear phase change becomes

\[
\Delta \varphi \approx \Delta k \cdot L = -\frac{\omega}{c} \left( \frac{n_g}{n} L \right) \Delta n = \frac{2\pi}{\lambda_0} \left( \frac{n_g}{n} L \right) \Delta n.
\]

In other words, for a given \( \Delta n \), the thickness to obtain a phase change of \( \Delta \varphi \) is smaller by a factor of \( n/n_g \), which is also called the “slow-down” factor. This means that in such materials, the demanded nonlinear effect is achieved in a shorter propagation length \([279, 286-291]\). This behavior is also referred to as the structural slow light enhancement of the NLO effects \([292]\). This effect can be understood further by considering that upon slowing down the pulse, the interaction time
\( \tau = \frac{L}{v_g} \), increases and pulse stays a longer time in the nonlinear medium. Another approach to explain this phenomenon is considering a longer effective interaction length described by \( L_{\text{eff}} = \frac{n_g}{n} L \), where \( L \) is the physical thickness of the nonlinear medium. We note, however, that interaction length and interaction time are fundamentally related to each other.

A similar concept also applies to TCO materials at their ENZ condition, particularly because they are inherently dispersive. In TCOs, a given change in permittivity translates to a significant change in the nonlinear phase \( \Delta \varphi \). Thus, in these cases, we cannot attribute the enhanced nonlinear phase change only to the small value of the index, and the dispersive nature of ITO needs to be taken into consideration. To address this complication, we relate \( \Delta \varphi \) to the dispersion of the material. Similar to photonic crystals, the scaling factor of \( n_g/n \) also applies to the enhanced phase sensitivity. The advantage of using NLO devices such as silicon photonic crystals is the possibility to design the dispersion on demand to obtain large \( n_g \); however, there is an additional advantage for TCOs where they exhibit large group index while the refractive index is also small, hence \( S = n_g/n \) exhibits a resonant behavior in vicinity of the ENZ region. We calculate the slow-down factor for our ITO sample and plotted the results in Figure 5-36. We can observe that close to the ENZ wavelength (\( \lambda_{\text{ENZ}} \approx 1250 \text{ nm} \)), the slow-down factor reach values higher than ten, indicating that the pulse can accumulate nonlinear phase change with \( \sim 10 \) times better efficiency due to the small group velocity and large phase velocity at this wavelength.
Large slow-down factor in TCOs can dramatically increase the phase sensitivity caused by any given changes in $\varepsilon$. In this case $\Delta \varphi$ experiences an intrinsic enhancement mechanism similar to the case of resonant conditions in traditional nonlinear optical materials. Here, a decrease in the plasma frequency, $\Delta \omega_p$, increases the permittivity by $\Delta \varepsilon$, and subsequently modifies the dispersion curve. It is important to note that $\omega$ and $k$ are related by the dispersion of the material that can be described via $\omega = \omega(k)$. Therefore, we can also consider that nonlinearities effectively change the dispersion relation altering both frequency and wavevector and define $(\omega, k)$ as a "photonic state”, which undergoes a transition in presence of NLO effects. Therefore, for a given change in the index a transition occurs from $(\omega_A, k_A)$ to $(\omega_B, k_B)$, where the induced change in frequency is $\Delta \omega = \omega_B - \omega_A$ and the induced change in wavevector is $\Delta k = k_B - k_A$. We note that in typical NLO interactions, the material is perturbed both in time and space and $\omega$ and $k$ will not be conserved. However, in our measurements the shift in frequency is significantly smaller than that of the wavevector. In this case, the optical transitions will be entirely horizontal as illustrated in Figure 5-37. We theoretically calculated $\Delta k$ for $\Delta \omega_p/\omega_p \approx 5\%$ ($\omega_p$ is the plasma frequency) and plotted the results in Figure 5-37(b). The maximum $\Delta k$ occurs close to the ENZ wavelength where
its magnitude is a factor of two larger than achievable values at 1000 nm and around factor of ten larger than values that can be achieved at 1500 nm.

Figure 5-37. (a) Dispersion of ITO without NLO effects (black) and after an optically-induced change in the plasma frequency $\Delta \omega_p/\omega_p \approx 5\%$ occurs (red). (b) Calculated $\Delta k$ at each frequency while only considering horizontal transitions.

To understand the enhanced phase sensitivity more precisely, when considering a tiny shift in frequency, we can write $\Delta \varphi$ as

$$\Delta \varphi \approx \Delta k \cdot L = \frac{\partial k}{\partial \omega} \Delta \omega \cdot L = \frac{L}{v_g} \Delta \omega$$  \hspace{1cm} (5-67)

In this relation, we can calculate $\Delta \omega$ from perturbation analysis while considering a given small change in the index, $\Delta n$, modifies the frequency. This means that the frequency, $\omega = ck/n$, will change to $\omega + \Delta \omega = ck/(n + \Delta n)$, thus we can find

$$\Delta \omega \approx \frac{ck}{n + \Delta n} - \frac{ck}{n} = -\frac{ck \Delta n}{n(n + \Delta n)} \approx -\frac{ck}{n^2} \Delta n = -\frac{\omega}{n} \Delta n$$  \hspace{1cm} (5-68)

Now substituting this in the previous equation, we obtain

$$\Delta \varphi \approx -\frac{L}{v_g} \frac{\omega}{n} \Delta n = -k_0 \Delta n \left(\frac{\nu_p}{v_g} L\right) = -k_0 \Delta n \left(\frac{\nu_p}{n} L\right)$$  \hspace{1cm} (5-69)
The above equation exhibits that the required length to obtain a given nonlinear phase change, $\Delta \varphi$, for a given change in index, $\Delta n$, is scaled by a factor of $S = v_p/v_g = n_g/n$. This scaling factor needs to be considered in all materials; however, since typically experiments are performed away from resonances, $S$ becomes very close to unity and can be ignored. For example, in GaAs the slow-down factor possesses values of $S = 1.07$ and $S = 1.04$ at 1300 nm and 1550 nm, respectively. This scaling factor is simply the direct consequence of the reduced group-velocity, which is further pronounced in ITO because of its increase phase velocity at wavelengths where refractive index is very small. A similar behavior happens for lossy materials, where linear absorption results in a shorter effective length by a factor of $L_{\text{eff}}/L = (1 - \exp(-\alpha L))/\alpha L$ requiring a longer propagation length to achieve a given phase change or nonlinear absorption behavior [15, 105, 106]. This is an intrinsic enhancement in NLO effect that can be observed in TCOs and is fundamentally different than enhancement mechanisms occurring in structured materials or organic molecular compounds, which are resorting to local enhancement effects or resonant nonlinear phenomena. We note, however, that besides this intrinsic enhancement mechanism, NLR is still very large in TCO materials enabling to achieve near unity values in the effective nonlinear refraction.

In the case of having large shifts for both frequency and wavevector, where ignoring vertical transitions is not possible, for a given change in index, $\Delta n$, the frequency shift, $\Delta \omega$, can be written as below

$$\Delta \omega \approx \Delta \omega_{\text{direct}} + \frac{\partial \omega}{\partial k} \cdot \Delta k,$$

(5-70)

This can also be shown geometrically on the dispersion curve, as illustrated in Figure 5-38. $\Delta \omega_{\text{direct}}$ is the amount of vertical shift at that given frequency of the probe pulse while $k$ is conserved.
Figure 5-38. Dispersion of ITO without NLO effects (black) and after an optically-induced change in the plasma frequency $\Delta \omega_p / \omega_p \approx 5\%$ occurs (red). The inset shows the contributions of $\Delta \omega_{\text{direct}}$ and projection of intraband frequency shift while considering the group velocity to the total shift in frequency.

This shift can be simply approximated as $\Delta \omega_{\text{direct}} / \omega \approx -\Delta n / n$, as mentioned above. Considering $v_g = \partial \omega / \partial k$, Eq. (5-70) can also be described as

$$
\Delta \omega \approx -\frac{\omega}{n} \Delta n + v_g \cdot \Delta k
$$

(5-71)

Knowing that $v_p = c / n$ and $v_g = c / n_g$, we get

$$
\Delta \omega \approx -v_p \frac{\omega}{c} \Delta n + v_g \cdot \Delta k
$$

(5-72)

By keeping $\Delta k$ on one side and taking other terms to the other side and dividing both sides to $k = k_0 n = n \omega / c$, we get
\[
\Delta k \approx \frac{v_p}{v_g} \left( -\frac{\Delta n}{n} + \frac{\Delta \omega}{\omega} \right)
\]

This equation can also be used in the absence of any nonlinear mechanisms, \textit{i.e.} \(\Delta n = 0\), where it results in the simple linear dispersion relation relating group and phase velocity in dispersive materials via

\[
\frac{\omega}{k} \approx \frac{v_p}{v_g} \frac{\Delta \omega}{\Delta k}
\]

It is important to note that to consider a general situation, we need to consider that the optical medium is both inhomogeneous in space (spatial change in index) and non-stationary in time (temporal change in index), and thus we need to consider the nonlinear index in a space-time domain, which we ignored in our derivations to simplify our equations.
CHAPTER 6: CONCLUDING REMARKS

This chapter provides the concluding remarks on the subjects discussed in this dissertation. The two main subjects in this work was characterization of the nonlinear optical processes in semiconductors, particularly GaAs, and studying the enhancement mechanisms of transparent conducting oxides, particularly ITO, at ENZ.

6.1. Characterization of Nonlinear Response of Semiconductors

We presented an overview of the fundamentals of different linear and nonlinear processes in semiconductor materials, focusing on two major mechanisms one related to the bound-electrons, and the other one to the free-carrier induced nonlinear mechanisms. We presented experimental and theoretical developments to measure and analyze NLO effects including 2PA, FCA, electronic Kerr effect, and FCR of direct-gap semiconductors particularly GaAs. The goal of this study was to address the discrepancies in the literature by performing similar measurements on GaAs using multiple methods in different laboratories. The experiments were mainly conducted by BD technique, which allows us to temporally separate the bound-electronic nonlinearities from the carrier-induced effects. We also presented comparison of our results with the ones performed with Z-scan at our institution and at the Georgia Institute of Technology. Moreover, similar measurements were conducted on a sample of Silicon to report the nonlinear response of this popular material in the NLO community. These independent measurements construct a reference for other researchers and scientists in the fields of silicon photonics, integrated nonlinear photonics, optical communication, and many other areas employing nonlinearities of semiconductors. Additionally, we presented 3PA measurements of GaAs and compared our results with our theory based on the Kane 8-band model.
This model gave excellent agreement with our experimentally determined 3PA spectra, and allowed us to make predictions for other semiconductors, i.e. for InSb, InAs, GaSb, InP, GaAs, CdTe, CdSe, ZnTe, CdS, ZnSe, GaN, ZnO, ZnS, AlP, and AlN. We were also able to separate contributions of the various bands to 3PA including an understanding of how spin-orbit coupling changes these contributions, and gives insight into multi-photon absorption in other semiconductor materials and low-dimensional systems such as quantum-wells, quantum dots, nanocrystals, and two-dimensional semiconductors. This theoretical understanding opens possibilities for researchers to use them in designing nonlinear optical devices.

6.2. Enhancement Mechanisms of Optical Nonlinearities at Epsilon-Near-Zero

We presented a fundamental study on the nonlinear optical effects of highly doped semiconductors, such as TCOs, at the spectral range where the photon energy is below their bandgap and in proximity to their ENZ condition. We also discussed the importance of treating the electrons as plasmas and studied the mechanism of loss in such materials. In this work, we presented characterization of slightly nondegenerate NLR using BD and compared our results with the degenerate NLR conducted via Z-scan. We conclude that without rigorously fitting the experimental data, we can calculate the NLR directly from CA/OA and BD data. The results from these two techniques are in complete agreement, and we showed that $\Delta n_{\text{eff}}$ can reach values near unity. We also presented BD measurements while fixing the probe wavelength away from the ENZ region and tuning the pump wavelength across the ENZ region. We performed these measurements to study the effect of exciting the carriers in proximity to the ENZ region. Additionally, we showed that the electron energy loss function increases while longitudinal component of the electric field excites the plasma. This effect was studied in detail using our BD methodology for different polarization
combinations and linear transfer matrix modeling, and we showed that it results in enhanced coupling to the longitudinal plasma oscillations of the material, hence NLR. We also performed BD measurements for a fixed pump away ENZ and tuning the probe beam across the ENZ region. In these measurements, we simultaneously characterized the effect of optically-induced temporal and spatial modulation of material permittivity, by simultaneously measuring $\Delta \varphi$ and $\Delta \lambda$ using BD and cross-phase modulation measurements.

In addition, we discussed the implications of the ENZ region on the group velocity of the pulse. We also presented a theoretical perspective relying on the influence of group velocity and phase velocity on the enhanced sensitivity of the nonlinear phase change to any material perturbations. This explanation unveils the role of slow-light propagation in the NLO response at ENZ. The outcomes of this work can be used for an efficient dispersion engineering close to material resonances and photonic systems towards improving light-matter interaction and paving the way towards a deeper understanding of the physics and complexity of nonlinear effects in such systems.
APPENDIX A: THEORETICAL APPROACH FOR CALCULATING 3PA USING THE KANE 8-BAND MODEL
In this appendix, we provide detail about the theoretical treatment of 3PA in zinc-blende semiconductors. This theory was developed by Dr. Claudiu M. Cirloganu and modified by Dr. Matthew Reichert to add corrections, both previous members of the NLO group [184-186]. In this theory, the Kane 8-band model is used for electronic states and the 3PA coefficients are calculated by employing the third-order perturbation theory.

We use the band structure model developed by Kane [181], because it includes band symmetry intermixing that is ignored in simplified 2-band models. We review this model following the original work by Kane and the summary given in Ref. [293]. We start with the Schrödinger equation described by

\[ H_0 \Psi_n(r) = E_n \Psi(r), \]  

(A-1)

where \( H_0 = p^2/(2m_0) + V(r) \) is the usual Hamiltonian operator. \( \Psi_n \) and \( E_n \) are the wavefunction and energy, respectively, \( p \) is the momentum operator, \( m_0 \) is the electron mass, and \( V(r) \) is the one-electron potential. By substituting the Bloch form, \( H_0 \Psi_n(r) = E_n \Psi(r) \), we get

\[ H_0 u_{nk}(r) + \left( \frac{\hbar}{m} \right) k \cdot p \ u_{nk}(r) = \left( E_n - \frac{\hbar^2 k^2}{2m_0} \right) u_{nk}(r), \]  

(A-2)

At \( k = 0 \) we obtain the simple Schrödinger equation governing on \( u_{n0} \). We, therefore, choose a finite set of orthonormal \( u_{n0} \) and assume that they form a complete set, then expand the \( u_{nk} \) as

\[ u_{nk} = \sum_m c_{mn}(k) u_{m0}(r), \]  

(A-3)
Note that the zone center \((k = 0)\) is the simplest basis that can be used. This is because at this point we have the highest symmetry. By multiplying \(u_{k0}(r)\) from left and integrating over the crystal volume, we can form a matrix equation for the expansion coefficients \(c_{kn}\). We diagonalize the matrix exactly and ignore perturbative contributions from bands outside the basis. The spin-orbit interaction term can be written as below, where we include spin by taking the total unit cell function to be the tensor product of spatial and spin degrees of freedom,

\[
H_{SO} = \frac{\hbar}{4m_0^2c^2}(\nabla V \times p) \cdot \sigma, \tag{A-4}
\]

Here, \(\sigma\) is the vector of Pauli spin matrices that act on the spin components and \(c\) is the speed of light.

A direct result of this analysis is that the unperturbed conduction band wave-functions have a spherical symmetry and are denoted by \(S \uparrow\) and \(S \downarrow\) for the two spin states, while the ones corresponding to the valence bands have a \(p\)-like orbital symmetry denoted here by \(X \uparrow, X \downarrow, Y \uparrow, Y \downarrow, Z \uparrow,\) and \(Z \downarrow\). These functions belong to the symmetry type \(\Gamma_1\) and \(\Gamma_4\), respectively. Using this similarity to the atomic wavefunctions we can recall that the \(s\) states correspond to \(l = 0\) while the \(p\) states correspond to \(l = 1\) and are triply degenerate, with \(l\) being the orbital angular momentum. The three degenerate states can be chosen to be the eigenstates of \(l_z\), the \(z\) projection of \(l\). The magnetic quantum numbers, eigenvalues of \(l_z\), are \(-1, 0\) and \(1\) for the \(p\) states. The eigenfunctions of the spin-orbit Hamiltonian are eigenstates of the total angular momentum \(j\), the sum of orbital momentum and spin, and its \(z\) components. For the \(p\) states (valence) \(j\) can take either the values of \(1 - 1/2 = 1/2\) or \(1 + 1/2 = 3/2\) while for the \(s\) states (conduction) \(j\) is \(1/2\). Thus, the eigenfunctions of \(j\) and \(j_z\) are for the valence states.
\[ |j m_j \rangle \rightarrow \begin{cases} \left| \frac{3}{2}, \frac{3}{2} \right\rangle, \left| \frac{3}{2}, -\frac{3}{2} \right\rangle & \text{Corresponding to the heavy - hole band} \\ \left| \frac{3}{2}, \frac{1}{2} \right\rangle, \left| \frac{3}{2}, -\frac{1}{2} \right\rangle & \text{Corresponding to the light - hole band} \\ \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \left| \frac{1}{2}, -\frac{1}{2} \right\rangle & \text{Corresponding to the split - off band} \end{cases} \]

The spin-orbit interaction removes the triple degeneracy, splitting the \( j = 3/2 \) states from the \( j = 1/2 \) states. Thus, the split-off band is shifted by an energy of \( \Delta \) at \( \mathbf{k} = 0 \).

In the most general form, including the interaction of the valence band and the conduction band via the \( \mathbf{k} \cdot \mathbf{p} \) coupling and the spin-orbit interaction, the electron Hamiltonian for a cell-periodic structure is written as

\[
H = \left( \frac{\mathbf{p}^2}{2m} + \frac{\hbar^2 k^2}{2m} + V + \left( \frac{\hbar}{m} \right) \mathbf{k} \cdot \mathbf{p} + \left( \frac{\hbar}{4m^2 c^2} \right) (\nabla V \times \mathbf{p}) \cdot \sigma \right)
\]  

(A-6)

For the case of \( \mathbf{k} \) oriented along \( z \), the Hamiltonian can be represented as

\[
\begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}
\]

where

\[
H = \begin{bmatrix} E_s & 0 & kP & 0 \\ 0 & E_p - \Delta/3 & \sqrt{2} \Delta/3 & 0 \\ kP & \sqrt{2} \Delta/3 & E_p & 0 \\ 0 & 0 & 0 & E_p + \Delta/3 \end{bmatrix}
\]

(A-7)

in the following basis

\[
| i \downarrow \rangle, \left| \frac{(X - iY)}{\sqrt{2}} \right\rangle, | Z \downarrow \rangle, \left| \frac{(X + iY)}{\sqrt{2}} \right\rangle, | i \uparrow \rangle, \left| \frac{(X + iY)}{\sqrt{2}} \right\rangle, \left| \frac{(X - iY)}{\sqrt{2}} \right\rangle, | Z \uparrow \rangle, \left| \frac{(X - iY)}{\sqrt{2}} \right\rangle
\]

(A-8)

where \( E_p \) and \( E_s \) correspond to the two eigenvalues obtained solving for \( E_{n0} \) in Eq. 6 for the valence and conduction band respectively, \( P \) is the Kane momentum parameter defined by
\[ P = -i \left( \frac{\hbar}{m} \right) \langle S|p_x|Z \rangle = -i \left( \frac{\hbar}{m} \right) \langle S|p_y|Y \rangle = -i \left( \frac{\hbar}{m} \right) \langle S|p_z|X \rangle, \quad (A-9) \]

and

\[
\Delta = \frac{3hi}{4m^2c^2} \left( X \left[ \frac{\partial V}{\partial x} p_y - \frac{\partial V}{\partial y} p_x \right] \right) \quad (A-10)
\]

For an arbitrary orientation of the \( \mathbf{k} \) vector, we can obtain the same form for the Hamiltonian using similar basis functions differentiated through the prime symbol. The primed functions are obtained through rotations from the original functions as given by

\[
\begin{bmatrix}
X' \\
Y' \\
Z'
\end{bmatrix} =
\begin{bmatrix}
cos \theta \cos \phi & cos \theta \sin \phi & -\sin \theta \\
-\sin \phi & cos \phi & 0 \\
\sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} \quad (A-11)
\]

and

\[
\begin{bmatrix}
\uparrow' \\
\downarrow'
\end{bmatrix} =
\begin{bmatrix}
e^{-\phi/2} \cos(\theta/2) & e^{\phi/2} \sin(\theta/2) \\
e^{\phi/2} \sin(\theta/2) & e^{\phi/2} \cos(\theta/2)
\end{bmatrix}
\begin{bmatrix}
\uparrow \\
\downarrow
\end{bmatrix} \quad (A-12)
\]

with \( S' = S \) invariant under this rotation. Here the angles \( \theta \) and \( \phi \) are the usual polar angles of the \( \mathbf{k} \) vector referred to the crystal symmetry axes \( x, y \) and \( z \). The characteristic equation yields four double roots as

\[ E' = 0 \]

\[ E'(E' - E_G)(E' + \Delta) - k^2 P^2 (E' + 2\Delta/3) = 0 \]

Where \( E_k = E'_k + (\hbar^2/2m)k^2 \) is the energy of a state with a wavevector \( \mathbf{k} \). The electronic wave functions are then given by
\[ \Phi_{i\alpha} = a_i [iS \downarrow] \prime + b_i [(X - iY) \uparrow/\sqrt{2}] \prime + c_i [Z \downarrow] \prime \]

\[ \Phi_{i\beta} = a_i [iS \uparrow] \prime + b_i [-(X + iY) \uparrow/\sqrt{2}] \prime + c_i [Z \uparrow] \prime \]

(A-14)

\[ \Phi_{HH\alpha} = [(X + iY) \uparrow/\sqrt{2}] \prime \]

\[ \Phi_{HH\beta} = [(X - iY) \downarrow/\sqrt{2}] \prime \]

where \( \alpha \) and \( \beta \) denote the two degenerate spin states (up and down). The index \( i \) refers to the conduction (c), light-hole (LH) and split-off (SO) bands. The coefficients \( a, b, \) and \( c \) are real and are obtained from

\[ a_i = kP(E'_i + 2\Delta/3)/N \]

\[ b_i = (\sqrt{2}\Delta/3)(E'_i - E_G)/N \]  

(A-15)

\[ c_i = (E'_i - E_G)(E'_i + 2\Delta/3)/N \]

where \( N \) is a normalizing factor such that \( a_i^2 + b_i^2 + c_i^2 = 1 \).

Using this formalism, we can obtain a complete description of the band structure taking into account the effects of non-parabolicity and non-zone center wave functions. We write the electron-radiation interaction (perturbation) Hamiltonian in SI units as

\[ H_{int} = \frac{e}{i\omega m_0} \left( \frac{l}{2e_0 n_0 c} \right)^{1/2} \hat{a} \cdot \hat{p} \]  

(A-16)
where $I$ is the irradiance of the incident beam, $n_0$ is the material’s index of refraction and $\hat{a}$ is a unit vector parallel to the direction of the incident electric field. Using third-order perturbation theory we can express the 3PA rate as

$$W_3 = \frac{2\pi}{h} \sum_{v,c} \left| \sum_{i,j} \langle \psi_c | H_{int} | \psi_j \rangle \langle \psi_j | H_{int} | \psi_i \rangle \langle \psi_i | H_{int} | \psi_v \rangle \right|^2 \left( E_{j\nu}(k) - 2\hbar\omega \right) \left( E_{i\nu}(k) - \hbar\omega \right)$$

(A-17)

The index $v$ corresponds to a valence band (initial state), $c$ corresponds to the conduction band (final state) including spin, while $i$ and $j$ are intermediate states and can be any of the four bands considered, while the $\Psi$’s are the respective Bloch wavefunctions. This means that the summation is performed over all valence to conduction band routes considering all possible intermediate states (same as the four bands in our model). From this, the 3PA coefficient $\alpha_3$ can be obtained using the relation $\alpha_3(\omega) = 3\hbar\omega W_3 I^{-3}$ [14, 176]. We performed our calculations over the entire wavelength range where 3PA is possible while 2PA is not possible (i.e. $E_g/3 \leq \hbar c/\lambda < E_g/2$). For each wavelength, we consider the transitions conserving the total energy and calculate the contributions to the 3PA coefficient arising from transitions from all three valence bands. This way, contributions from each of the three valence bands can be independently analyzed as well as their total effect which is the simple sum. In practice, the above equation takes a more complicated form which is shown below, which is evaluated numerically by using only a few parameter as inputs such as bandgap, split-off energy, Kane parameter and wavelength dependent refractive index, and is evaluated at the point in k-space $k = k_r$ where $E_{cv}(k_r) = 3\hbar\omega$. 

169
\[ \alpha_3 = 3\omega \frac{(2\pi)^5}{(n_0 c)^3} \left( \frac{eP}{\hbar \omega} \right)^6 \sum_{c,\nu} \int_0^{2\pi} \left( \left| \sum_{i,j} M_{cj}(k_r, \theta) M_{ji}(k_r, \theta) M_{iv}(k_r, \theta) (E_{i\nu}(k_r) - \hbar \omega)(E_{j\nu}(k_r) - 2\hbar \omega) \right| \right)^2 \kappa_r^2 \sin \theta \, d\theta \] 

where

\[ M_{ij}(k_r, \theta) = 3 \frac{\hbar}{m_0 P} \hat{\alpha} \cdot \langle \psi_i | \hat{\vec{p}} | \psi_j \rangle, \] 

are the scaled momentum matrix elements [34].
APPENDIX B: DERIVATION OF FRESNEL COEFFICIENTS FOR A THREE-LAYER SYSTEM
Here we present the formulation for the Fresnel coefficients in a linear case, considering a three-layer system, where the first medium is air, the second medium is a lossy material, and the third medium is a semi-infinite dielectric, e.g. glass. The structure that we are studying is comprising of three layers where the first interface is at $z = 0$ and the second interface is at $z = d$. The angle of incident is $\theta_i$ and the dielectric constants are $\varepsilon_1, \varepsilon_2 = \varepsilon'_2 + i\varepsilon''_2$, and $\varepsilon_3$, respectively.

In order to derive the relations for transmission, reflection, and absorption, there are two approaches. The simple one is to not consider field interference effects and simply derive the coefficients for the intensity of light. The second approach is to derive it for the electric field.

**Approach 1:**

Considering Fresnel reflection coefficients for the intensity to be $R_{12}$ for the interface between first and second medium and $R_{23}$ for the interface between second and third medium, and the absorption coefficient of the lossy medium to be $\alpha$, we can calculate the total reflection and transmission as

$$ R = R_{12} + (1 - R_{12})^2 e^{-2\alpha d} R_{23} + (1 - R_{12})^2 e^{-4\alpha d} R_{12} R_{23}^2 $$

$$ + (1 - R_{12})^2 e^{-6\alpha d} R_{12}^2 R_{23}^3 + \cdots $$

$$ T = (1 - R_{12}) e^{-\alpha d} (1 - R_{23}) + (1 - R_{12}) e^{-3\alpha d} R_{12} R_{23} (1 - R_{23}) $$

$$ + (1 - R_{12}) e^{-5\alpha d} R_{12}^2 R_{23}^2 (1 - R_{23}) + \cdots $$

Thus, $T$ and $R$ can be simplified further and resulting in

$$ R = R_{12} + (1 - R_{12})^2 e^{-2\alpha d} R_{23} (1 + e^{-2\alpha d} R_{12} R_{23} + \cdots ) $$

$$ = R_{12} + \frac{(1 - R_{12})^2 e^{-2\alpha d} R_{23}}{1 - e^{-2\alpha d} R_{12} R_{23}} $$

$$ T = (1 - R_{12}) e^{-\alpha d} (1 - R_{23}) (1 + e^{-2\alpha d} R_{12} R_{23} + \cdots ) = \frac{(1 - R_{12}) e^{-\alpha d} (1 - R_{23})}{1 - e^{-2\alpha d} R_{12} R_{23}} $$

172
Note that, in this approach multiple reflections are considered; however, the interference between fields is ignored.

**Approach 2:**

Let us now consider the transverse electric (TE) polarization, where the electric field is perpendicular to the plane of interface and the magnetic field is in plane. We can write the electric field for each medium as below

\[
\vec{E}_i = \hat{y}E_i e^{i(k_{i,norm}z + k_{i,tan}x)}
\]

\[
\vec{E}_r = \hat{y}E_r e^{i(k_{r,norm}z + k_{r,tan}x)}
\]

\[
\vec{E}_f = \hat{y}E_f e^{i(k_{f,norm}z + k_{f,tan}x)} \tag{B-5}
\]

\[
\vec{E}_r = \hat{y}E_b e^{i(k_{b,norm}z + k_{b,tan}x)}
\]

\[
\vec{E}_t = \hat{y}E_t e^{i(k_{t,norm}(z-d) + k_{t,tan}x)}
\]

The continuity relations of the Maxwell’s equation, considering there is no surface current or charge will result

\[
\begin{align*}
@ z = 0: \ &(\hat{z} \times (\vec{E}_i + \vec{E}_r - \vec{E}_f - \vec{E}_r)) = 0 \tag{B-6} \\
@ z = d: \ &\hat{z} \times (\vec{E}_f + \vec{E}_r - \vec{E}_t) = 0
\end{align*}
\]

Therefore,

\[
E_i e^{ik_{i,tan}x} + E_r e^{ik_{r,tan}x} = E_f e^{ik_{f,tan}x} + E_b e^{ik_{b,tan}x}
\]

\[
E_f e^{i(k_{f,norm}d + k_{f,tan}x)} + E_b e^{i(k_{b,norm}d + k_{b,tan}x)} = E_t e^{i(k_{t,norm}(z-d) + k_{t,tan}x)} \tag{B-7}
\]

These two equations hold for any given \(x\), thus we conclude that \(k_{i,tan} = k_{r,tan} = k_{f,tan} = k_{b,tan} = k_{t,tan} = k_0 \sqrt{\varepsilon_1} \sin \theta_i\). Therefore,

\[
E_i + E_r = E_f + E_b
\]
\[ E_f e^{i k_{f,\text{norm}} d} + E_b e^{i k_{b,\text{norm}} d} = E_t \]  

where \( k_{f,\text{norm}} = -k_{b,\text{norm}} = k_0 \sqrt{\varepsilon_2 - \varepsilon_1 \sin \theta_i^2} = \delta / d \). Similarly, \( k_{t,\text{norm}} = k_0 \sqrt{\varepsilon_3 - \varepsilon_1 \sin \theta_i^2} \). Thus, if we define

\[
\begin{align*}
    r_{TE} &= E_r / E_i \\
    t_{TE} &= E_t / E_i
\end{align*}
\]

We have

\[
E_f + E_b = E_i (1 + r_{TE})
\]

\[
E_f e^{i \delta} + E_b e^{-i \delta} = E_i t_{TE}
\]

Similarly, for the magnetic field, \( \mathbf{\hat{H}} = -\frac{k_{\text{norm}} E}{\omega \mu_0} \), therefore,

\[
\begin{align*}
    -k_{i,\text{norm}} E_i + k_{i,\text{norm}} E_r &= -k_{f,\text{norm}} E_f + k_{f,\text{norm}} E_b \\
    -k_{f,\text{norm}} E_f e^{i k_{f,\text{norm}} d} + k_{f,\text{norm}} E_b e^{i k_{b,\text{norm}} d} &= -k_{t,\text{norm}} E_t
\end{align*}
\]

Therefore,

\[
\begin{align*}
    -k_{i,\text{norm}} (E_i - E_r) &= -k_{f,\text{norm}} (E_f - E_b) \\
    -k_{f,\text{norm}} (E_f e^{i \delta} - E_b e^{-i \delta}) &= -k_{t,\text{norm}}
\end{align*}
\]

Which will result in

\[
\begin{align*}
    E_f - E_b &= E_i (1 - r_{TE}) k_{i,\text{norm}} / k_{f,\text{norm}} \\
    E_f e^{i \delta} - E_b e^{-i \delta} &= E_i t_{TE} k_{t,\text{norm}} / k_{f,\text{norm}}
\end{align*}
\]

Dividing the second series to the first series

\[
\begin{align*}
    \frac{E_f - E_b}{E_f + E_b} &= \frac{(1 - r_{TE}) k_{i,\text{norm}}}{(1 + r_{TE}) k_{f,\text{norm}}}
\end{align*}
\]
\[
\frac{E_f e^{i\delta} - E_b e^{-i\delta}}{E_f e^{i\delta} + E_b e^{-i\delta}} = \frac{k_{t,\text{norm}}}{k_{f,\text{norm}}}
\]

Where with a simple algebra, we can conclude from the second equation that

\[
\frac{E_b e^{-i2\delta}}{E_f} = \frac{k_{f,\text{norm}} - k_{t,\text{norm}}}{k_{f,\text{norm}} + k_{t,\text{norm}}} = r_{ft} \tag{B-15}
\]

Substituting in the first equation, and \(\gamma_{if} = \frac{k_{\text{inorm}} - k_{f,\text{norm}}}{k_{\text{inorm}} + k_{f,\text{norm}}}\), we will get

\[
\frac{1 - r_{ft} e^{-i2\delta}}{1 + r_{ft} e^{-i2\delta}} = \frac{(1 - r_{TE})}{(1 + r_{TE})} \frac{1 + \gamma_{if}}{1 - \gamma_{if}} \tag{B-16}
\]

Thus, we conclude that

\[
\gamma_{TE} = \frac{\gamma_{if} + r_{ft} e^{i2\delta}}{1 + \gamma_{if} r_{ft} e^{i2\delta}} \tag{B-17}
\]

where \(\gamma_{if} = \frac{k_{\text{inorm}} - k_{f,\text{norm}}}{k_{\text{inorm}} + k_{f,\text{norm}}}\), \(r_{ft} = \frac{k_{f,\text{norm}} - k_{t,\text{norm}}}{k_{f,\text{norm}} + k_{t,\text{norm}}}\) and \(\delta = k_{f,\text{norm}} d\), and

\[
t_{TE} = \frac{(1 + r_{ft})(1 + r_{if}) e^{i\delta}}{1 + \gamma_{if} r_{ft} e^{i2\delta}} \tag{B-18}
\]

Similarly for a transverse magnetic (TM) field, we can define \(\gamma_{if} = (k_{\text{inorm}}/\varepsilon_1 - k_{f,\text{norm}}/\varepsilon_2)/(k_{\text{inorm}}/\varepsilon_1 + k_{f,\text{norm}}/\varepsilon_2)\) and \(r_{ft} = (k_{f,\text{norm}}/\varepsilon_2 - k_{t,\text{norm}}/\varepsilon_3)/(k_{f,\text{norm}}/\varepsilon_2 + k_{t,\text{norm}}/\varepsilon_3)\). Then, similar to the TE case

\[
\gamma_{TM} = \frac{\gamma_{if} + r_{ft} e^{i2\delta}}{1 + \gamma_{if} r_{ft} e^{i2\delta}} \tag{B-19}
\]

\[
t_{TM} = \frac{(1 + r_{ft})(1 + r_{if}) e^{i\delta}}{1 + \gamma_{if} r_{ft} e^{i2\delta}}
\]

Finally, we can derive the transmission and reflection of the light intensity as

\[
R_{TE} = |\gamma_{TE}|^2 \quad \text{and} \quad R_{TM} = |\gamma_{TM}|^2
\]
Finally, absorbed power can be found via $A = 1 - R - T$.


84. S. R. Marder, "Organic nonlinear optical materials: where we have been and where we are going," Chemical communications, 131-134 (2006).


120. S. Benis, N. Munera, E. W. Van Stryland, and D. J. Hagan, "Large time-dependent nonlinear phase shift of transparent conducting oxides at epsilon-near-zero (Conference Presentation)," in Nonlinear Frequency Generation and Conversion: Materials and Devices XIX, (International Society for Optics and Photonics, 2020), 112640N.


123. S. Benis, P. Zhao, D. J. Hagan, and E. W. Van Stryland, "Measurement of the ultrafast dynamics of nonlinear refraction and absorption of highly doped semiconductors at epsilon-near-zero (Conference Presentation)," in *Ultrafast Bandgap Photonics III*, (International Society for Optics and Photonics, 2018), 106380X.


125. D. J. Hagan, S. A. Benis, and E. W. Van Stryland, "Large, ultrafast induced index changes in ITO (Conference Presentation)," in *Metamaterials, Metadevices, and Metasystems 2019*, (International Society for Optics and Photonics, 2019), 110800E.


230. J.-Y. Ou, J.-K. So, G. Adamo, A. Sulaev, L. Wang, and N. I. Zheludev, "Ultraviolet and visible range plasmonics in the topological insulator Bi 1.5 Sb 0.5 Te 1.8 Se 1.2," Nature communications 5, 1-7 (2014).


