

University of Southampton Research Repository ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

Optoelectronics Research Centre

**Characterising Nonlinear Waveguides by
Scanning Near-field Optical Microscopy**

by

Tipsuda Chaipiboonwong

Thesis for the degree of Doctor of Philosophy

September 2008

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
OPTOELECTRONICS RESEARCH CENTRE

Doctor of Philosophy

by Tipsuda Chaipiboonwong

Scanning near-field optical microscopy (SNOM) has been applied to investigate the dispersion and nonlinear phenomena in a multimode Ta_2O_5 rectangular waveguide. Unlike the conventional approach of observing only the output spectra, the SNOM technique can collect the localised spectra from the evanescent field at various locations of the waveguide. This provides the visualisation of pulse evolution prior to the final development as the output light. The SNOM-acquired spectra consist of unique features which have not been observed before in previous nonlinear pulse propagation researches. These distinctive characteristics are attributed to the localised nature of the data and the multimode nonlinear pulse propagation. In order to understand the underlying physics of the experimental data, a numerical model simulating this SNOM visualisation has been developed. The simulation was based on the nonlinear Schrödinger equation, adapted for multimode pulses, and performed by the split-step Fourier algorithm.

The spectra exhibit very fine features which can be attributed to the interference of various modes with different phase modulation owing to dispersion and nonlinear effects. Accordingly, the complexity of the spectral features increase with the propagation distance and the number of contributing modes. The multimode spectra rapidly broaden at the beginning stage of the propagation, owing to the supplementary intermodal phase modulation. Unlike the single-mode case, in which the spectral broadening caused by the self-phase modulation continuously develops along the propagation distance, the broadening process in the multimode pulse is decelerated at the later distance. This is owing to the separation of the higher-order modes and consequently the influence of the cross-phase modulation on the spectral broadening is reduced.

The SNOM technique can also provide the observation of high resolution evolution of the pulse spectra. Both spectral variations along the length of the waveguide and across the waveguide are observable. Such a variation over the wavelength scale is caused by the interference of modes with different phases complexly formed by the dispersion and nonlinear effects.

Contents

Table of Contents	ii
List of Figures	v
List of Tables	xi
List of Symbols	xii
List of Abbreviations	xiii
Declaration of Authorship	xv
Acknowledgements	xvi
1 Introduction	1
1.1 General review of SNOM	1
1.2 Application of SNOM to investigate continuum generation light sources .	3
1.3 Thesis overview	5
1.3.1 Chapter 1, Introduction	5
1.3.2 Chapter 2, Scanning near-field optical microscopy theory and review	5
1.3.3 Chapter 3, Construction of a SNOM system	5
1.3.4 Chapter 4, Waveguide Characteristics	6
1.3.5 Chapter 5, Numerical modeling of nonlinear pulse propagation . .	6
1.3.6 Chapter 6, Visualisation of nonlinear pulse propagation by SNOM	7
1.3.7 Chapter 7, Modal dispersion analysis and other aspects relating to the accuracy of the simulation	7
1.3.8 Chapter 8, Conclusions and suggested further work	8
1.4 References	8
2 Scanning near-field optical microscopy theory and review	13
2.1 Chapter introduction	13
2.2 Diffraction-limited resolution	14
2.3 Angular spectrum of diffracted field	14
2.4 Scanning near-field imaging system	16
2.4.1 Operation modes	16
2.4.2 Near-field probe	18
2.4.2.1 Probe shape	18
2.4.2.2 Metal coated or uncoated probes	20
2.4.3 Height controlling system	20

2.5	Interaction of probe and near-field	22
2.5.1	Various approaches for analysing the interaction	22
2.5.2	Ultrashort light with SNOM probes	24
2.6	Chapter conclusion	25
2.7	References	26
3	Construction of a SNOM system	32
3.1	Chapter introduction	32
3.2	Near-field probe fabrication	33
3.3	Probe oscillation	36
3.4	Probe height control	39
3.5	Scanning system and data acquisition	41
3.6	Vibration and thermal effects	43
3.7	Chapter conclusion	44
3.8	References	44
4	Waveguide characteristics	47
4.1	Chapter introduction	47
4.2	Planar waveguides	48
4.2.1	Field distribution of planar waveguides	48
4.2.2	Propagation modes of planar waveguides	54
4.3	Rectangular waveguides	57
4.4	Nonlinear pulse propagation	65
4.5	Chapter conclusion	74
4.6	References	75
5	Numerical modeling of nonlinear pulse propagation	77
5.1	Chapter introduction	77
5.2	Split-step Fourier Method	78
5.2.1	Basic algorithm	78
5.2.2	Numerical errors and 2 nd -order approximation	80
5.3	Determination of propagation modes and dispersion parameters	82
5.4	Mode transverse field distribution	95
5.5	NLS equation for multimode pulses	98
5.6	Enhancement of numerical simulation	103
5.6.1	The effect of step size	103
5.6.2	Higher-order approximation	107
5.6.3	Adaptive step size	109
5.7	Chapter conclusion	113
5.8	References	114
6	Visualisation of nonlinear pulse propagation by SNOM	117
6.1	Chapter introduction	117
6.2	SNOM setup for detecting continuum spectra	118
6.2.1	Continuum light sources	119
6.2.2	Collecting continuum spectra	119
6.3	Spectral evolution along the waveguide	123
6.3.1	Experimental data	123

6.3.2	Simulation of multimode pulse propagation along the waveguide length	125
6.4	Spectral growth along the distance of propagation	129
6.4.1	Experimental data	129
6.4.2	Spectral growth of the multimode simulation data	129
6.5	High resolution spectral evolution	135
6.5.1	Experimental data	135
6.5.2	Simulation result	135
6.6	Spectral variation across the waveguide	139
6.6.1	Experimental data	139
6.6.2	Simulation result	141
6.7	Chapter conclusion	143
6.8	References	144
7	Modal dispersion analysis and other aspects relating to the accuracy of the simulation	146
7.1	Chapter introduction	146
7.2	Measurement of group velocity difference	147
7.2.1	Spectral interference	147
7.2.2	Experimental result	150
7.2.3	Simulation result	156
7.2.4	The effect of nonlinearity on the measurement	159
7.3	The effect of third-order dispersion	161
7.4	The effect of wavelength-dependent transverse field distribution	163
7.5	Chapter conclusion	168
7.6	References	169
8	Conclusions and suggested further work	170
8.1	Conclusions	170
8.2	Suggested further work	172
8.3	References	173
A	Publications	175
B	Operation parameters of P-2000 Micropipette Puller	177
B.1	References	179

List of Figures

2.1	Schematic diagrams of illumination and collection modes	17
2.2	Meniscus etching method	19
2.3	Metal coating on the side walls of a tapered fibre tip	20
2.4	A quartz tuning fork with near-field fibre probe attached	21
2.5	An example of frequency response of an oscillating system. The energy or amplitude of oscillation is plotted against various driving frequencies	22
3.1	Schematic diagram of the SNOM system used during this research	32
3.2	Diagram of a Sutter P-2000 micropipette puller [1]	33
3.3	A fibre tip fabricated by the Sutter P-2000 puller. (a) A microscope image displaying the taper length of the tip ~ 1 cm (b) An SEM closeup image of the tip diameter which is in the range of 100 nm [3]	35
3.4	Probe dithering diagram	37
3.5	Frequency response of the fibre tip/tuning fork system, measured when there was no sample. (a) measured X output ($R \cos \Phi$), Y output ($R \sin \Phi$) and calculated amplitude (R) from the LIA as a function of frequency (b) calculated phase (Φ) as a function of frequency	38
3.6	Schematic diagram of the feedback system	39
3.7	Three consecutive line scans on the same area of a grating shows the data repeatability of the system. There was no vibration isolation while the data was being acquired and the time for all scans is around 12 minutes. Note that the data has been processed by removing an inclined plane inherent in the data, caused by the sample stage not being entirely flat.	42
3.8	SEM image and schematic diagram of TGX01 test sample [18]. The grating is the periodic pattern of trapezoidal ridges with the period 3 μm . Each ridge is 0.9 μm high and 1.1 μm wide (the top side).	42
3.9	A false-color topographical image of TGX01. The height data from scanning is plane removed by MATLAB scripts. The pitch of pattern in the image ranges 3.15-3.21 μm whereas the height is around 0.8 μm	43
4.1	Total internal reflection in a one-dimensional waveguide	48
4.2	Basic structure of a rectangular waveguide	58
4.3	Various types of rectangular waveguides	59
4.4	Two orthogonal planar waveguides for the analysis of a rectangular waveguide by the EIM. There are two choices of utilising the method: (a) determining the propagation modes from the planar waveguide along the x -direction with the effective index of the core given by the propagation constant previously determined from the planar waveguide along the y -direction or (b) vice versa	59

4.5	Schematic diagram of the effective index method with the introduction of the correcting parameter (ζ) in the dielectric constant of the side cladding area. The analysis starts with determining modes in the planar waveguide with the refractive index profile along the y direction in (a), and is then followed by the planar waveguide along the x direction in (b)	60
4.6	Equivalent waveguide by the effective index method	62
5.1	1 st -order algorithm for modeling pulse propagation	80
5.2	2 nd -order algorithm for modeling pulse propagation	81
5.3	Waveguide geometry and the detail of polarisation notation	82
5.4	Calculated intensity distributions for modes with various mode indices m and n . Note the white lines indicating the waveguide profile	83
5.5	Refractive index dispersion of Ta ₂ O ₅ and SiO ₂ calculated from Equation 5.10 and Equation 5.11 respectively, for the wavelength range 400-1200 nm	83
5.6	Mode determination along the y -direction. The left-hand (LH) and right-hand (RH) sides of the characteristic equations in Equation 4.52a and Equation 4.52c.	84
5.7	Mode determination along the x -direction	85
5.8	The variation of propagation constant β for some of the low-order modes calculated by the EIM over the wavelength 400 nm - 950 nm	88
5.9	Percentage difference between the propagation constant, β , calculated by EIM of the mode calculation programme written in this research and FDM of the commercial software OympIOS for (a) symmetric modes and (b) antisymmetric modes	89
5.10	The dispersion of the propagation constant over the wavelength range 400 - 950 nm. On the left column ((a), (c) and (e)) are for the symmetric modes whereas the asymmetric modes are shown on the right ((b), (d) and (f)). The dispersion of β_1 , β_2 β_3 are shown in the first row ((a),(b)), second row ((c),(d)) and last row ((e),(f)) respectively.	90
5.10	The dispersion of the propagation constant over the wavelength range 400 - 950 nm. On the left column ((a), (c) and (e)) are for the symmetric modes whereas the asymmetric modes are shown on the right ((b), (d) and (f)). The dispersion of β_1 , β_2 β_3 are shown in the first row ((a),(b)), second row ((c),(d)) and last row ((e),(f)) respectively. (cont)	91
5.11	Evaluation of the values of the dispersion parameters at a particular wavelength	91
5.12	The effect of wavelength data size on determining (a) the dispersion constant and (b)-(d) its first three orders of derivatives at the wavelength 800 nm for the E_{00}^y mode. The effect is quantified as the percentage difference between the value determined from utilising different wavelength data sizes ranging from 2^5 to 2^{12} and that derived from the maximum data size of 2^{12} . Two methods of determining the derivatives of the dispersion constants can also be compared in the plots: red curves are for the non-direct method whereas green curves are for the direct method. . .	94
5.13	Normalised field distribution of various propagation modes: (a) E_{00}^y (b) E_{10}^y (c) E_{01}^y and (d) E_{01}^x . The waveguide area is also displayed as a dashed rectangle.	96

5.14	Intensity integration $\iint_{-\infty}^{+\infty} U(x, y) ^2 dx dy$ against the parameter ρ which is the ratio between the relative range of x (or y) and the number of data points, as defined in Equation 5.15: (a) symmetric E^y -polarisation modes (b) asymmetric E^y -polarisation modes (c) symmetric E^x -polarisation modes (d) asymmetric E^x -polarisation modes	97
5.15	Deviation of $\iint_{-\infty}^{+\infty} U(x, y) ^2 dx dy$, calculated by utilising various numbers of data points, from that generated by the maximum data size of $2^{12} \times 2^{12}$ (a) symmetric E^y -polarisation modes (b) asymmetric E^y -polarisation modes (c) symmetric E^x -polarisation modes (d) asymmetric E^x -polarisation modes	98
5.16	Transverse field distribution of various modes along the diameter of the SNOM probe, which is placed at the central position of the waveguide and 20 nm above its top surface	101
5.17	Relative error from single-mode (E_{00}^y) simulations calculated by using different values of step size	104
5.18	(a) Relative error from multimode simulations calculated by using different values of step size (b) Details of contributing modes and the relative intensity ratio	105
5.19	Spectra and time profiles through the probe simulated from different step sizes, shown in the inset box. (a),(b) Single-mode simulation (c),(d) 3-mode simulation (e),(f) 5-mode simulation. Note that the deviation of the spectra in (a) and time profiles in (b), (d) and (f) for various step sizes is hardly noticeable and the curves are quite superimposed	106
5.20	Comparison of the efficiency of 2 nd -order and 3 rd -order algorithms	108
5.21	The spectral error accumulated along the propagation distance, calculated from the simulation results utilising the step size 2 μm (red), 5 μm (green) and 10 μm (blue) with (a) single-mode propagation and (b) 5-mode propagation	110
5.22	(a) The step size in the multimode simulation, adjusted along the distance of propagation in order to maintain the tolerance error of 1.2728×10^{-7} . The red curve is for 3-mode simulation whereas the green curve is for 5-mode simulation. (b) Details of contributing modes and their intensity ratio	112
5.23	The spectral relative error of the error-controlled simulation compared with the results from the constant step size of 1 μm for 3-mode (red) and 5-mode (green) pulse propagations. In contrast to Figure 5.21, the error stays constant for all over the distance of propagation.	113
6.1	Continuum generation chip (CGC)	119
6.2	Experimental setup for visualising continuum spectra developing inside the waveguide	120
6.3	(a) Variation of penetration depth over the study wavelength range (b) Intensity variation caused by the wavelength dependency of the penetration depth when the probe height is 20 nm above the surface	121
6.4	Topography of (a) Waveguide A (b) Waveguide B	122
6.5	Continuum evolution along the waveguide at different input pulse energies: 0.8 nJ (—), 1.5 nJ (—) and 2.1 nJ (—) for (a) Waveguide A and (b) Waveguide B respectively [9]	124

6.6	The spectrum of the input pulse, where the red curve is from the measured spectra and the green curve is from the simulation	126
6.7	(a) Simulation of the spectral evolution of multimode pulse along the length of the waveguide. The simulated spectra are observed by the probe which is placed above the waveguide by 20 nm and shift from the central axis of the waveguide by 200 nm (b) Details of contributing modes and their relative intensity ratio [13]	128
6.8	Comparison of the width of the measured spectra along the length of the waveguide, which is illustrated in Figure 6.5(b), with that from the single mode simulation [9]	129
6.9	(a) RMS pulse width broadening factor along the length of the waveguide of the simulated multimode spectra shown in Figure 6.7, except for an additional 3-mode curve (violet) which is the mixing of E_{00}^y , E_{10}^y and E_{20}^y with intensity ratio 0.4:0.3:0.3 (b) Details of (a) at the first 0.5 mm distance of propagation	131
6.10	Same as Figure 6.9(a) but the phase velocity difference of modes is neglected [13]	132
6.11	(a) RMS pulse width broadening factor of individual modes contributing to the three-mode pulse (green curve) in Figure 6.10 (b) Same as (a) but for three-mode pulse (violet curve) in Figure 6.10	134
6.12	High-resolution spectral mapping along the waveguide over the distance of 2 μm . The spectra were collected when the probe was at the position of $z = 3$ mm and moved toward the end face by the steps of 100 nm [9]	135
6.13	Simulated high-resolution spectra through the probe when the probe moves along the length of the waveguide by the distance of 2 μm , starting from the position of $z = 4$ mm. The simulation is performed with various multimode pulses in the waveguide: (a) $E_{00}^y:E_{10}^y:E_{20}^y$ (0.4:0.3:0.3) (b) $E_{01}^y:E_{11}^y:E_{21}^y$ (0.4:0.3:0.3) (c) $E_{00}^y:E_{10}^y:E_{01}^y$ (0.4:0.3:0.3) (d) Same as (c) but the phase velocity difference is neglected	136
6.14	Same as Figure 6.13(c) but for individual contributing modes (a) E_{00}^y (b) E_{10}^y and (c) E_{01}^y	138
6.15	High-resolution spectral mappings across the waveguide over the distance of 12 μm and their corresponding topographical image. The spectra were collected by the probe translating across the waveguide by the steps of 500 nm at various positions from the entrance face: (a) $z = 2.2$ mm (b) $z = 3$ mm (c) $z = 4$ mm and (d) $z = 5$ mm	139
6.15	High-resolution spectral mappings across the waveguide over the distance of 12 μm and their corresponding topographical image. The spectra were collected by the probe translating across the waveguide by the steps of 500 nm at various positions from the entrance face: (a) $z = 2.2$ mm (b) $z = 3$ mm (c) $z = 4$ mm and (d) $z = 5$ mm (cont)	140
6.16	(a) Simulated spectra through the probe across the centre of the waveguide for the propagation of three-mode pulses. The position of the probe is $z = 4$ mm. The similar spectra of the contributing modes are also displayed: (b) E_{00}^y (c) E_{10}^y (d) E_{20}^y	142

7.1	Spectral interference caused by 2-mode pulse propagation with different group velocities. The time profiles at various locations along the length of the waveguide are shown on the left whereas their corresponding spectra are on the right. The propagation distances are (a)-(b) 0 mm (c)-(d) 2 mm and (e)-(f) 4 mm. The number of fringes in the spectra increases with the propagation distance.	148
7.2	Spectral variation over the distance of 200 μm along the length of the waveguide after the propagation of 5 mm (a) without and (b) with the phase velocity difference	149
7.3	(a) Spectral intensity variation along the length of the waveguide at the wavelength of 800 nm, marked by the white line in Figure 7.2(b) (b) The Fourier transform of (a)	149
7.4	Measured spectral data at various locations along the length of the waveguide (solid violet curves) in comparison with the interpolated data (dashed orange curves) in the new evenly-spaced frequency scale	151
7.5	Fourier transform of spectra along the waveguide at different input energies of 0.8 nJ (red curves), 1.5 nJ (green curves) and 2.1 nJ (blue curves). The separations of higher-ordered modes can be observed with the dashed lines (a), (b) and (c) marking the temporal positions of three different modes at various locations along the waveguide. The details of temporal separations and distances of these three modes will be later shown in Figure 7.6 [3].	153
7.6	The separation in time of the mode (a), (b) and (c) in Figure 7.5 at various propagation distances	154
7.7	Simulated time profile at various propagation distances for the spectrum of the 7-mode pulse in Figure 6.7(a), Chapter 6 [4]	156
7.8	Relative integrated intensity collected by the SNOM probe in the simulation. The dashed curve displays the envelope of the variation when the phase velocity difference term is removed.	157
7.9	Time separation from the fundamental mode of various higher-ordered modes in Figure 7.7 against the distance of propagation	158
7.10	Comparison between the simulations of a 7-mode pulse propagation with the same contributing modes and relative intensity ratio as that in Figure 6.7, Chapter 6. Two cases of the simulations are that one includes both dispersion and nonlinear effects (solid red curves) and the other excludes the nonlinear term (dashed green curves): (a) simulated spectra observed by the probe which is placed at the position of $z = 5.5$ mm and (b) their corresponding time profiles.	160
7.11	The simulated spectra observed by the probe which is placed at the position of $z = 5.5$ mm for (a) 3 modes (b) 5 modes. The result when the dispersion consists only of the second-order term is shown by the solid red curve whereas the dashed green curve is for including the third-order term. The intensity ratio is the same as those in Figure 6.7, Chapter 6. Both curves are totally superimposed and their deviation cannot be observed.	162
7.12	Variation of transverse field over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{01}^y (blue), and E_{11}^y (orange) at different distances from the centre of the waveguide along the line $x = 200$ nm parallel to the height dimension	163

7.13	Variation of transverse field over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{01}^y (blue), and E_{11}^y (orange) at different distances along the $y = 0.27 \mu\text{m}$ parallel to the horizontal dimension of the waveguide	164
7.14	Normalised transverse field $U(x, y)$ over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{20}^y (blue), E_{30}^y (orange) and E_{40}^y (violet) at the position of the SNOM probe $(x, y) = (0.2 \mu\text{m}, 0.27 \mu\text{m})$	165
7.15	Comparison of 5-mode mixing simulated spectra between the case of constant (solid red curve) and wavelength-dependent (dashed green curve) transverse field. The simulated spectra are observed by a probe which is placed at the position of (a) $z = 1 \text{ mm}$ and (b) $z = 5 \text{ mm}$	166
7.16	The root mean squared deviation (RMSD), calculated by Equation 7.9, between the spectra simulated when $U(x, y)$ is constant and wavelength-dependent. The deviations are from the spectra of single mode (red curve), three-mode (green curve) and five-mode pulses (blue curve). The detail of contributing modes and the relative intensity ratio is the same as already described in Figure 6.7, Chapter 6.	167

List of Tables

3.1	Puller parameters (arbitrary units) for pulling fibres as near-field probes. The used value is an integer unit with its corresponding physical value shown in the last column. More details of these parameters can be found in Appendix B	34
5.1	Effective index and side cladding index of the waveguide with the index profile along the x direction. The values are determined from the one-dimensional mode calculation along the y -axis.	85
5.2	Propagation constant β for modes determined from Figure 5.7 at the wavelength 800 nm (a) E^y -polarisation modes (b) E^x -polarisation modes	86
5.2	Propagation constant β for modes determined from Figure 5.7 at the wavelength 800 nm (a) E^y -polarisation modes (b) E^x -polarisation modes (cont)	87
5.3	The propagation constant and its derivatives at the wavelength 800 nm for (a) E^y -polarisation modes (b) E^x -polarisation modes	92
5.4	Nonlinear parameter $\gamma^{(p)(q)}$ for (a) E^x -polarisation modes, and (b) E^y -polarisation modes. Each value in a cell is determined from Equation 5.17 by applying the transverse field of the modes in the corresponding row and column for $U^{(p)}(x, y)$ and $U^{(q)}(x, y)$	100
6.1	Penetration depth for the evanescent field at the interface of air and Ta ₂ O ₅ of the waveguide	121
6.2	Label and average width of waveguides from topographical data. The SNOM probe scanned across the waveguide at various distances from the front face of the waveguide and the precision in the table is calculated from the standard deviation of the measured data.	122
7.1	Group index difference (Δn_g) of some higher-ordered E^y modes whose separation from the main pulse can be observed in Figure 7.5	154
7.2	Group index difference (Δn_g) of some higher-ordered E^y modes calculated from the dispersion curves that are determined by the effective index method	155
7.3	Group index difference Δn_g of E_{20}^y , E_{30}^y , E_{40}^y , E_{50}^y and E_{01}^y whose separation from the main pulse can be observed in Figure 7.7	158
B.1	Corresponding scan length for the unit of FILAMENT [1]	178

List of Symbols

ω	angular frequency
λ	wavelength
f	frequency
c	speed of light in vacuum
n	refractive index
n_{eff}	effective index
Q	quality factor
μ	magnetic permeability
ϵ	electric permittivity
$\chi^{(r)}$	r^{th} -order electric susceptibility
n_2	nonlinear refractive index coefficient
$U(x, y)$	Transverse field profile
$V(z, t)$	Slowly varying amplitude function
A_{eff}	Effective mode area

List of Abbreviations

AFM	Atomic Force Microscopy
CO ₂	Carbon dioxide
CGC	Continuum Generation Chip
DVM	Digital Volt Meter
EIM	Effective Index Method
FDTD	Finite Difference Time Domain
FDM	Finite Difference Method
FIT	Finite Integration Technique
FWHM	Full Width at Half Maximum
GeO ₂	Germanium dioxide
GVD	Group Velocity Dispersion
HF	Hydrofluoric Acid
LIA	Lock-in Amplifier
MMP	Multiple Multipole Method
NA	Numerical Aperture
NLS	Nonlinear Schrödinger (equation)
OCT	Optical Coherence Tomography
ODE	Ordinary Differential Equation
OPD	Optical Path Difference
PID	Proportional-Integral-Derivative unit
PSD	Phase-sensitive Detector
PSTM	Photon Scanning Tunneling Microscopy
RMSD	Root Mean Squared Deviation
SEM	Scanning Electron Microscopy
SiO ₂	Silicon dioxide
SNOM	Scanning Near-field Optical
SSFM	Split-step Fourier Method
SPM	Self-phase Modulation
SPP	Surface Plasmon Polaritons
STM	Scanning Tunneling Microscopy
SVEA	Slowly Varying Envelope Approximation
Ta ₂ O ₅	Tantalum pentaoxide

TE	Transverse Electric
TM	Transverse Magnetic
XPM	Cross-phase Modulation

Declaration of Authorship

I, **TIPSUDA CHAIPIBOONWONG**, declare that the thesis entitled **Characterising Nonlinear Waveguides by Scanning Near-field Optical Microscopy** and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published (see the list of publications in Appendix A)

Signed:

Date:

Acknowledgements

This research could not have been completed without the support and assistance of many people. First of all, I would like to acknowledge the financial support of the Thai Government for my PhD study Scholarship.

I would like to thank Dr. Bill Brocklesby for his supervision, good advice and suggestions for the project.

I am indebted to Dr. John Mills for his supervision, his energetic assistance and for being a friend. John also helped me in collecting the SNOM experimental data in this research. Thanks also for his patience in keeping on correcting my English in all my writings during my PhD study.

My special and grateful thanks go to Dr. Peter Horak for valuable discussions and guidance in the numerical simulation work of this research.

Thanks to Dr. James Gates so much for his advice and assistance in constructing the SNOM system at the beginning of my PhD.

I would like to express my gratitude to Dr. Eleanor Tarbox for her overseeing and helping me to get rid of obstacles along this PhD road. Thank her also for the proofreading of the thesis and providing very useful comments.

I also wish to thank many ORC staff who provided assistance in various aspects during my PhD study: Simon Butler, Chris Nash, and Tim McIntyre for their technical support in the laboratory; Dave Oliver and Kevin Summer for their assistance in all computer and network problems; Eve Smith for her active help in administrative matters.

My appreciation also goes to many friends in UK and Thailand who give me true friendship. They are always beside me to cheer me up and make me laugh whenever I feel depressed.

Finally, I thank my family for their love, understanding, caring and tireless encouragement.

To my beloved Papa

ฉันได้อยู่ดูโลกให้เสียใจนะ พรุ่งนี้ฉันจะ... ไม่รู้ ... ฉันตาย
- ร. ๓. ๓. ๓. ๓๓๓๓๓๓ ๓๓๓๓๓๓

Chapter 1

Introduction

In this research, the characteristics of the spectral broadening of a nonlinear multimode waveguide were studied by the technique of scanning near-field optical microscopy (SNOM). The spectral data was collected experimentally with a SNOM probe along the propagation distance of the waveguide. In contrast to conventional studies, which detect only the output spectra from the end face of the waveguide, this novel approach enables the visualisation of spectral development at various propagation distances of the optical pulses. By comparing the experimental results qualitatively with the simulated spectra (generated from a numerical model developed in this research for nonlinear propagation of multimode pulses), the interplay between the dispersion and nonlinear effects, contributing in the spectral development, can be better understood. For instance, it was found that the additional cross-phase modulation, due to the multimode nature of the waveguide, can provide the rapid spectral broadening at the shorter length scale of the propagation distance, than in the case of single mode, and before the spectral growth is decelerated by the mode separation. Such a knowledge can be useful for the design of continuum generation light sources, as the proper dimension of the waveguide or the refractive index of waveguide material can be optimised in order to take the benefit of the cross-phase modulation.

This chapter provides a preliminary overview of SNOM, including its concept and technology development in Section 1.1. Then, in the following Section 1.2 the application of SNOM as a waveguide characterising tool, as in the case of this research, will be concisely reviewed. Finally the outline of the thesis and each chapter's brief content are given in Section 1.3.

1.1 General review of SNOM

The resolution of conventional optical systems is based on detecting propagating waves scattered from objects, and therefore is limited by diffraction. Firstly formulated by

Abbe and rearranged later on by Rayleigh [1–5], the far-field resolution cannot go far beyond approximately half the wavelength, which falls between 200 - 300 nanometres for the visible region. The resolution can be enhanced by using short wavelengths such as UV or X-ray, applying higher index material in oil-immersion microscopy or excluding out-of-focus light to increase image contrast in confocal microscopy. Unfortunately all these improvements do not make a significant impact on the resolution limit.

Conversely another component of scattered light, the near-field, is not governed by the Abbe limit and hence provides the finest details of the object. The field adheres to the complex propagation wave number and confines itself within the surface of the object. As a result, the field cannot be detected by any conventional imaging optics.

The idea of near-field imaging was firstly proposed by Synge [6] in 1928. He suggested improving images of biological samples by illuminating them with light passing through a tiny pinhole, of diameter around 10 nm, on an opaque screen. The screen must be in the region that is not further away from the surface than the aperture size. Since the scheme of illuminating is localised, the screen must scan across the whole interesting area. The transmitted light through the sample is collected by a remote detector. This arrangement of detecting the near-field is called *illumination mode*. Likewise the task can be achieved in *collection mode*. A subwavelength aperture can be utilised to collect scattered light from an object illuminated by a far-field light source. By reciprocity, both schemes are equivalent. Nevertheless, the excited area on the sample in collection mode is larger than that in illumination mode.

Synge's proposal was technically challenging at that time, especially the ability of bringing down and maintaining the aperture in the near-field region. Therefore the idea was neglected until the advent of scanning tunneling microscopy (STM) which introduced the technology of nanopositioning. The first two designs of scanning near-field optical microscopy (SNOM) were independently invented by IBM Zurich Research Laboratory [7, 8] and Cornell University [9–12]. In both works, sharply pointed transparent rods were utilised as near-field probes. The probes were metal coated except the very end in order to form apertures. They were moved by piezoelectric translators in both the scanning plane and probe height plane. The other type of near-field probe is the so-called apertureless probe. In this case the probe is not required to be transparent. Therefore a very fine metal tip can be utilised. This allows the size of the probe to be smaller, which improves spatial resolution. The tip behaves as a scattering centre which perturbs the field and results in field enhancement [13–15].

Another crucial requirement for the SNOM system is the maintaining of the probe in the near-field region. Unlike STM, which makes use of the detected signal from electron tunneling for the probe height control, SNOM cannot always exploit the optical signal because the relationship between the optical field and the gap distance is not linear and is quite complicated. Therefore the distance-dependent effect of shear force is applied

instead. The shear force, whose exact nature is still ambiguous, consists of several force types such as van der Waals, capillary, or frictional forces. The shear force retards the oscillation of the probe by the amount dependent on the gap distance between the surface and the probe. Hence, controlling the probe height is achieved by detecting the vibration characteristics of the probe and its change. The amplitude of oscillation can be monitored by optical means, such as interferometric techniques, or by detecting the reflection of light incident on the probe [16, 17]. Another approach, proposed by Karrai et al. [18], is based on electronic signals. The probe is attached to one prong of a piezoelectric tuning fork which is dithered laterally. The electrical signal from the fork, associated with the amplitude of oscillation, is fed into the electronic feedback system. The amplitude signal is compared with the dithering signal and the deviation is compensated by the correcting signal fed back to the height-translating unit. The electronic technique prevents the interference of stray light into the imaging system and provides compactness and a less tedious setup. The topographical resolution of SNOM is not at the same level as other high-resolution topography-based microscopy techniques such as scanning electron microscopy (SEM) or atomic force microscopy (AFM). The SNOM technique is non-destructive and able to work under ambient conditions.

1.2 Application of SNOM to investigate continuum generation light sources

The application of SNOM to gain optical information about light propagating inside a waveguide can be found in many research studies. The evanescent field collected by the SNOM probe can provide the modal intensity profile inside a planar waveguide [19, 20]. With the addition of heterodyne interferometric detection, the phase evolution of the optical pulse can be detected, which leads to information on the phase and group velocity of propagating modes [21, 22]. The technique has also been applied to the more complex structure of photonic crystal waveguides [23–26]. In this research, the SNOM technique was exploited to gain insight into the spectral development within a nonlinear waveguide which is designed as a continuum generation source. Additionally, the corresponding Fourier-transformed temporal information has provided the relative group velocity of various propagating modes.

The supercontinuum generation light source is a useful tool for optical coherence tomography (OCT) [27], optical frequency metrology [28, 29], and femtosecond-pulse phase stabilisation [30]. The phenomenon can be observed when pico- or femto-second light pulses are launched into a nonlinear medium. Various nonlinear effects such as self- or cross-phase modulation, four-wave mixing or stimulated Raman scattering contribute to the spectral broadening of the light pulse. The first observation was performed by Alfano et al. in crystal and glass filaments [31] and later in different types of medium by various research groups [32–38]. However, the most common form of the light source

is probably an optical fibre which can be a photonic crystal type or tapered standard fibre [39–41].

The investigation of the phenomena is normally performed by observing the output spectra from the optical sources [42–44]. However, the information acquired by such an approach is an accumulative outcome and does not provide the visualisation of spectral evolution inside the devices. Alternatively, by the application of SNOM collecting the localised spectra, the visualisation of spectral evolution inside the waveguide becomes attainable.

Numerical modeling has been evoked in order to understand the interplay of various effects in the spectral broadening process seen in the experimental observations. The propagation of pulses through a nonlinear dispersive medium can be mathematically depicted in the form of a nonlinear partial differential equation, namely the nonlinear Schrödinger equation (NLS), which originated from Maxwell’s equations [45]. To reduce computational tediousness, the slowly varying envelope approximation (SVEA) is normally applied and only one spatial dependence of the field envelope along the propagation direction is retained. This assumption is valid as long as the spectral bandwidth $\Delta\omega$ is much less than the central frequency ω_0 [46]. The equation can be extended to include the dispersion effect with higher-order terms and many nonlinear phenomena such as self-phase modulation, stimulated Raman scattering, and four-wave mixing [47–51].

Recently there has been interest in the nonlinearity in waveguides on silicon-based chips owing to the future prospect for the application in optical integrated systems [52–56]. The dimensions of the waveguides are in micron or submicron scales and the wide expansion of the spectra can be achieved over the length scale of few centimetres, in contrast to metre length scale in the case of optical fibres. A strong intensity-dependent nonlinearity is obtainable owing to tight confinement of the propagating light inside the waveguide and the higher value of nonlinear refractive index of the waveguide in materials such as Ta_2O_5 or silicon, which have the value of n_2 that is one or two orders of magnitude higher than the silica glass of optical fibres [53, 55, 57].

In this study, Ta_2O_5 waveguides, fabricated on a silicon chip, were investigated experimentally by the SNOM probe and numerically by a numerical model developed for multimode pulse propagation. Such a waveguide was fabricated as a continuum generation light source which can operate at low power [58] and provides a potential device to be applied in optical integrated circuits. However, owing to the physical structure of the waveguide, the pulse propagation is in the multimode regime and the optical phenomena inside the waveguide become more complicated. The study of the spatial variation of the spectra can reveal the effect of the multimode nature contributing in the spectral development, as will be described in detail in later chapters.

The brief overview of the thesis contents will be given in the next section.

1.3 Thesis overview

The layout of the thesis generally consists of three parts. The first part, Chapter 2 and Chapter 4, is the theoretical framework of the research associated with SNOM and the waveguide characteristics. The next part, Chapter 3 and Chapter 5, is the research methodology including the construction of SNOM and the numerical model utilised in the research. Finally both experimental and numerical results are presented and discussed in Chapter 6 and Chapter 7.

The contents of the chapters are briefly given in the following sections.

1.3.1 Chapter 1, Introduction

This chapter provides a general overview related to scanning near-field optical microscopy and its application as a visualising tool to observe optical phenomena in a nonlinear waveguide.

1.3.2 Chapter 2, Scanning near-field optical microscopy theory and review

In this chapter, the principles of SNOM will be described. The resolution of a conventional imaging system, which is bound by diffraction, will be described. With the analysis of the angular spectrum, it can be shown that the scattered light containing the high spatial frequency components is confined within the surface of the object whereas a far-field imaging system can detect only low spatial frequency information. Therefore, the basic scheme for a SNOM system consists of a subwavelength optical probe and a control system which retains the probe in the near-field proximity. The various types of probes, optical data collecting schemes and the interaction of probe and samples will also be discussed in the chapter.

1.3.3 Chapter 3, Construction of a SNOM system

The technical features of a SNOM system, constructed in the research, are presented. Basically the system consists of three integral parts: the subwavelength optical probe, the probe-dithering system and the probe height control. The probe is fabricated from an optical fibre that is tapered by pulling while it is melted by a heating laser. The operation of controlling the gap distance between the probe and the sample is based upon the interaction of the distance-dependent shear-force. The probe is attached to a tuning fork driven by a signal generator and its deviation of oscillation amplitude is monitored by a feedback system. The evanescent light is collected by a probe and

directed to a spectrometer in order to be analysed. Finally, the issues of vibration and thermal effects are also discussed.

1.3.4 Chapter 4, Waveguide Characteristics

Theoretical characteristics of a waveguide will be focused on in this chapter. Light confined by a step index profile waveguide propagates with discrete propagation constants that are specified by the polarisation and the properties of the waveguide medium. For a one-dimensional index profile, the characteristic equation of the propagation modes can be formulated by utilising ray optics and the modal field is determined by Maxwell's equations and the boundary conditions. When the index profile becomes two-dimensional, the effective index method (EIM) can be applied to approximate the propagation modes under the assumption that the two-variable function of the modal field is separable. Consequently the mode analysis in a rectangular waveguide is simplified to a one-dimensional problem, i.e. two orthogonal planar waveguides with refractive index distribution corresponding to the two-dimensional index profile. The propagation modes in the rectangular waveguide are determined from one planar waveguide with the refractive index of the core specified by the propagation constants determined from the other planar waveguide. Moreover, the perturbation theory can also be adapted to enhance the accuracy of the method.

The interaction of the electric field of the propagating light with the medium results in the induced polarisation density which can be linearly and nonlinearly proportional to the field. The former leads to the dispersion effect whereas the latter contributes to optical nonlinear phenomena. The interplay of all the effects can be formulated as a nonlinear partial differential equation, namely the nonlinear Schrödinger (NLS) equation, which originates from the Maxwell's equations.

1.3.5 Chapter 5, Numerical modeling of nonlinear pulse propagation

This chapter addresses the technical aspect of the numerical model utilised in the research. The split-step Fourier (SSF) method, widely adopted to numerically solve a nonlinear partial differential equation, will be applied to model the pulse propagation in a nonlinear waveguide. Its algorithm is described together with a discussion on its numerical error. Two main parameters are required for the numerical models: the dispersion attribute of the waveguide and the nonlinear parameters. The first can be mathematically determined from the dispersion curve, given by the wavelength dependence of the propagation constants, whereas the nonlinear parameter is specified by the nonlinear refractive index of the waveguide material and the transverse field distribution. The EIM is adapted to determine both the propagation constants and the field distribution of various modes.

The propagation of multimode pulses is modeled by the set of NLS equations which are coupled via the nonlinear part of the equations. The coupling nonlinear parameter is proportional to the overlapping transverse field integration of contributing modes. Finally, the error of the numerical model is investigated with suggestions to enhance the accuracy.

1.3.6 Chapter 6, Visualisation of nonlinear pulse propagation by SNOM

Both the spectral data collected by the SNOM probe and generated by the numerical model are presented in parallel in this chapter. The continuum light source investigated in the research is in the form of a silicon-based chip on which various sets of rib waveguides are fabricated. The evolution of the spectra observed by the probe along the distance of propagation exhibits the spectral broadening, including features of complicated fine fringes. The phenomenon is confirmed by the simulation results which show that such a spectral complexity is introduced by the mode inference, and therefore the degree of complexity increases with the number of contributing modes. The spectral broadening of the multimode pulses is enhanced by the intermodal phase modulation at the very beginning of the pulse propagation, which occurs before the mode separation, caused by different group velocities, comes into play. The SNOM technique also enables the observation of spectral variation on the subwavelength scale. It was revealed by the simulation result that such a variation is the result of phase velocity differences of contributing modes.

The spectral variation across the waveguide is unprecedentedly observed by the SNOM. This variation is attributed to the superposition of propagation modes involving various phase and transverse mode field distributions, as confirmed by the simulated spectra.

1.3.7 Chapter 7, Modal dispersion analysis and other aspects relating to the accuracy of the simulation

The phenomenon of mode separation is presented in this chapter. The temporal evolution of the pulse along the propagation distance can be acquired from the Fourier transform of the corresponding spectral data. The relative temporal separation of various contributing modes along the distance of propagation is seen as a consequence. This provides an alternative method to obtain information on group velocity differences of the waveguide's modes. The simulation data also displays similar mode separation and elucidates the spectral interference effect.

Other aspects related to the accuracy of the simulation are also presented in the chapter. The deviation of the simulation results, with the inclusion of the third-order dispersion and the wavelength-dependence of the transverse field, is investigated.

1.3.8 Chapter 8, Conclusions and suggested further work

The final chapter reviews the work of the research and its results. Suggestions for future work are also provided.

1.4 References

- [1] E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9:413–420, 1873.
- [2] J. W. Strutt (Lord Rayleigh). On the manufacture and theory of diffraction-gratings. *Philosophical Magazine*, 47:81–93; 193–205, 1874.
- [3] J. W. Strutt (Lord Rayleigh). Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*, 8:261–274; 403–411; 477–486;, 1879.
- [4] J. W. Strutt (Lord Rayleigh). Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*, 9:40–55, 1880.
- [5] J. W. Strutt (Lord Rayleigh). On the theory of optical instruments, with special reference to the microscope. *Philosophical Magazine*, 42:167–195, 1896.
- [6] E. H. Synge. A suggested method for extending microscopic resolution into the ultra-microscopic region. *Philosophical Magazine*, 6:356–362, 1928.
- [7] D. W. Pohl, W. Denk, and M. Lanz. Optical stethoscopy: image recording with resolution $\lambda/20$. *Applied Physics Letters*, 44(7):651–653, 1984.
- [8] U. Dürig, D. W. Pohl, and F. Rohner. Near-field optical-scanning microscopy. *Journal of Applied Physics*, 59(10):3318–3327, 1986.
- [9] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray. Development of a 500 Å spatial resolution light microscope. I. light is efficiently transmitted through $\lambda/16$ diameter apertures. *Ultramicroscopy*, 13(3):227–231, 1984.
- [10] E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer. Near field scanning optical microscopy (NSOM): Development and biophysical applications. *Biophysical Journal*, 49:269–279, 1986.
- [11] E. Betzig, A. Harootunian, A. Lewis, and M. Isaacson. Near-field diffraction by a slit: implications for superresolution microscopy. *Applied Optics*, 25(12):1890–900, 1986.
- [12] A. Harootunian, E. Betzig, M. Isaacson, and A. Lewis. Super-resolution fluorescence near-field scanning optical microscopy. *Applied Physics Letters*, 49(11):674–676, 1986.

- [13] R. Bachelot, P. Gleyzes, and A. C. Boccara. Reflection-mode scanning near-field optical microscopy using an apertureless metallic tip. *Applied Optics*, 36(10):2160–2170, 1997.
- [14] L. Novotny, E. J. Sanchez, and X. S. Xie. Near-field optical imaging using metal tips illuminated by higher-order Hermite-Gaussian beams. *Ultramicroscopy*, 71(1-4):21–29, 1998.
- [15] H. Furukawa and S. Kawata. Local field enhancement with an apertureless near-field-microscope probe. *Optics Communications*, 148(4-6):221–224, 1998.
- [16] E. Betzig, J. S. Weiner, and P. L. Finn. Combined shear force and near-field scanning optical microscopy. *Applied Physics Letters*, 60(20):2484–2486, 1992.
- [17] R. Toledo-Crow, P. C. Yang, Y. Chen, and M. Vaez-Iravani. Near-field differential scanning optical microscope with atomic force regulation. *Applied Physics Letters*, 60(24):2957–2959, 1992.
- [18] K. Karrai and R. D. Grober. Piezoelectric tip-sample distance control for near field optical microscopes. *Applied Physics Letters*, 66(14):1842–1844, 1995.
- [19] D. P. Tsai, H. E. Jackson, R. C. Reddick, S. H. Sharp, and R. J. Warmack. Photon scanning tunneling microscope study of optical waveguides. *Applied Physics Letters*, 56(16):1515–1517, 1990.
- [20] A. G. Choo, H. E. Jackson, U. Thiel, G. N. De Brabander, and J. T. Boyd. Near field measurements of optical channel waveguides and directional couplers. *Applied Physics Letters*, 65(8):947–949, 1994.
- [21] M. L. M. Balistreri, J. P. Korterik, L. Knipers, and N. F. van Hulst. Local observations of phase singularities in optical fields in waveguide structures. *Physical Review Letters*, 85(2):294–297, 2000.
- [22] M. L. M. Balistreri, H. Gersen, J. P. Korterik, L. Kuipers, and N. F. van Hulst. Tracking femtosecond laser pulses in space and time. *Science*, 294(5544):1080–1082, 2001.
- [23] E. B. McDaniel, J. W. P. Hsu, L. S. Goldner, R. J. Tonucci, E. L. Shirley, and G. W. Bryant. Local characterization of transmission properties of a two-dimensional photonic crystal. *Physical Review B (Condensed Matter)*, 55(16):10878–10882, 1997.
- [24] P. L. Phillips, J. C. Knight, B. J. Mangan, P. St. J. Russell, M. D. B. Charlton, and G. J. Parker. Near-field optical microscopy of thin photonic crystal films. *Journal of Applied Physics*, 85(9):6337–6342, 1999.
- [25] E. Flück, M. Hammer, A. M. Otter, J. P. Korterik, L. Kuipers, and N. F. van Hulst. Amplitude and phase evolution of optical fields inside periodic photonic structures. *Journal of Lightwave Technology*, 21(5):1384–1393, 2001.

- [26] A. L. Campillo, J. W. P. Hsu, C. A. White, and A. Rosenberg. Mapping the optical intensity distribution in photonic crystals using a near-field scanning optical microscope. *Journal of Applied Physics*, 89(5):2801–2807, 2001.
- [27] I. Hartl, X. D. Li, C. Chudoba, R. K. Hganta, T. H. Ko, J. G. Fujimoto, J. K. Ranka, and R. S. Windeler. Ultrahigh-resolution optical coherence tomography using continuum generation in an air-silica microstructure optical fiber. *Optics Letters*, 26(9):608–610, 2001.
- [28] S. A. Diddams, D. J. Jones, J. Ye, S. T. Cundiff, and J. L. Hall. Direct link between microwave and optical frequencies with a 300 THz femtosecond laser comb. *Physical Review Letters*, 84(22):5102–5105, 2000.
- [29] R. Holzwarth, T. Udem, T. W. Hänsch, J. C. Knight, W. J. Wadsworth, and P. St. J. Russell. Optical frequency synthesizer for precision spectroscopy. *Physical Review Letters*, 85(11):2264–2267, 2000.
- [30] D. J. Jones, S. A. Diddams, J. K. Ranka, A. Stentz, R. S. Windeler, J. L. Hall, and S. T. Cundiff. Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis. *Science*, 288(5466):635–639, 2000.
- [31] R. R. Alfano and S. L. Shapiro. Observation of self-phase modulation and small-scale filaments in crystals and glasses. *Physical Review Letters*, 24(11):592–594, 1970.
- [32] W. Werncke, A. Lau, M. Pfeiffer, K. Lenz, H.-J. Weigmann, and C. D. Thuy. An anomalous frequency broadening in water. *Optics Communications*, 4(5):413–415, 1972.
- [33] W. Yu, R. R. Alfano, C. L. Sam, and R. J. Seymour. Spectral broadening of picosecond 1.06 μ pulse in KBr. *Optics Communications*, 14(3):344–347, 1975.
- [34] W. L. Smith, P. Liu, and N. Bloembergen. Superbroadening in H₂O and D₂O by self-focused picosecond pulses from a YAlG:Nd laser. *Physical Review A (General Physics)*, 15(6):2396–2403, 1977.
- [35] R. L. Fork, C. V. Shank, C. Hirlimann, R. Yen, and W. J. Tomlinson. Femtosecond white-light continuum pulses. *Optics Letters*, 8(1):1–3, 1983.
- [36] P. B. Corkum, C. Rolland, and T. Srinivasan-Rao. Supercontinuum generation in gases. *Physical Review Letters*, 57(18):2268–2271, 1986.
- [37] V. François, F. A. Ilkov, and S. L. Chin. Experimental study of the supercontinuum spectral width evolution in CO₂ gas. *Optics Communications*, 99(3-4):241–246, 1993.
- [38] A. Brodeur and S. L. S Chin. Band-gap dependence of the ultrafast white-light continuum. *Physical Review Letters*, 80(20):4406–4409, 1998.

- [39] J. K. Ranka, R. S. Windeler, and A. J. Stentz. Visible continuum generation in air silica microstructure optical fibers with anomalous dispersion at 800 nm. *Optics Letters*, 25(1):25–27, 2000.
- [40] T. A. Birks, W. J. Wadsworth, and P. St. J. Russell. Supercontinuum generation in tapered fibers. *Optics Letters*, 25(19):1415–1417, 2000.
- [41] W. J. Wadsworth, A. Ortigosa-Blanch, J. C. Knight, T. A. Birks, T.-P. M. Man, and P. St. J. Russell. Supercontinuum generation in photonic crystal fibers and optical fiber tapers: a novel light source. *Journal of the Optical Society of America B (Optical Physics)*, 19(9):2148–2155, 2002.
- [42] J. M. Dudley, L. Provino, N. Grossard, H. Maillotte, R. S. Windeler, B. J. Eggleton, and S. Coen. Supercontinuum generation in air-silica microstructured fibers with nanosecond and femtosecond pulse pumping. *Journal of the Optical Society of America B (Optical Physics)*, 19(4):765–771, 2002.
- [43] Q. Cao, X. Gu, E. Zeek, M. Kimmel, R. Trebino, J. Dudley, and R. S. Windeler. Measurement of the intensity and phase of supercontinuum from an 8-mm-long microstructure fiber. *Applied Physics B: Lasers and Optics*, 77(2-3):239–244, 2003.
- [44] T. Hori, N. Nishizawa, T. Goto, and M. Yoshida. Experimental and numerical analysis of widely broadened supercontinuum generation in highly nonlinear dispersion-shifted fiber with a femtosecond pulse. *Journal of the Optical Society of America B (Optical Physics)*, 21(11):1969–1980, 2004.
- [45] K. J. Blow and D. Wood. Theoretical description of transient stimulated Raman scattering in optical fibers. *IEEE Journal of Quantum Electronics*, 25(12):2665–2673, 1989.
- [46] G. P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, CA, USA, 3rd edition, 2001.
- [47] S. Coen, A. H. L. Chau, R. Leonhardt, J. D. Harvey, J. C. Knight, W. J. Wadsworth, and P. St. J. Russell. White-light supercontinuum generation with 60-ps pump pulses in a photonic crystal fiber. *Optics Letters*, 26(17):1356–1358, 2001.
- [48] A. V. Husakou and J. Herrmann. Supercontinuum generation of higher-order solitons by fission in photonic crystal fibers. *Physical Review Letters*, 87(20):203901/1–203901/4, 2001.
- [49] A. L. Gaeta. Nonlinear propagation and continuum generation in microstructured optical fibers. *Optics Letters*, 27(11):924–926, 2002.
- [50] J.M. Dudley and S. Coen. Numerical simulations and coherence properties of supercontinuum generation in photonic crystal and tapered optical fibers. *IEEE Journal of Selected Topics in Quantum Electronics*, 8(3):651–659, 2002.

- [51] F. Biancalana, D. V. Skryabin, and P. St. J. Russell. Four-wave mixing instabilities in photonic-crystal and tapered fibers. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 68(4):46603/1–46603/8, 2003.
- [52] H. K. Tsang, C. S. Wong, T. K. Liang, I. E. Day, S. W. Roberts, A. Harpin, J. Drake, and M. Asghari. Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength. *Applied Physics Letters*, 80(3):416–418, 2002.
- [53] C.-Y. Tai, J. S. Wilkinson, N. M. B. Perney, M. C. Netti, F. Cattaneo, C. E. Finlayson, and J. J. Baumberg. Determination of nonlinear refractive index in a Ta_2O_5 rib waveguide using self-phase modulation. *Optics Express*, 12(21):5110–5116, 2004.
- [54] O. Boyraz, T. Indukuri, and B. Jalali. Self-phase-modulation induced spectral broadening in silicon waveguides. *Optics Express*, 12(5):829–834, 2004.
- [55] G. W. Rieger, K. S. Virk, and J. F. Young. Nonlinear propagation of ultrafast 1.5 μm pulses in high-index-contrast silicon-on-insulator waveguides. *Applied Physics Letters*, 84(6):900–902, 2004.
- [56] E. Dulkeith, Y. A. Vlasov, X. Chen, N. C. Panoiu, and R. M. Osgood, Jr. Self-phase-modulation in submicron silicon-on-insulator photonic wires. *Optics Express*, 14(12):5524–5534, 2006.
- [57] C. S. Wong, T. K. Liang, M. W. K. Mak, H. K. Tsang, I. E. Day, A. Harpin, J. Drake, and M. Asghari. Measurement of nonlinear optical properties of silicon waveguide at 1.55 μm wavelength. In *Conference on Lasers and Electro-Optics/Quantum Electronics and Lasers Science Conference (CLEO/QELS 2001)*, page 178, Baltimore, MD, USA, 2001. Optical Society of America, Washington, DC, USA.
- [58] C. M. Netti, M. E. Zoorob, S. Roberts, M. D. B. Charlton, G. J. Parker, J. J. Baumberg, J. R. Lincoln, M. Lederer, and D. Kopf. Low-noise self-phase modulation continuum generation in high index tapered planar waveguide at 1040 nm. In *Coherence Domain Optical Methods and Optical Coherence Tomography in Biomedicine IX*, volume 5690, pages 222–227, San Jose, CA, USA, 2005. International Society for Optical Engineering.

Chapter 2

Scanning near-field optical microscopy theory and review

2.1 Chapter introduction

In this chapter, the theoretical aspects of scanning near-field optical microscopy (SNOM) will be presented. In Section 2.2, the resolution limit of the conventional imaging system is described. The limit is bound by diffraction and, approximately, the limit cannot be below half a wavelength. Consequently, in order to achieve better resolution, a novel imaging criteria based on the detection of the near-field, instead of the far-field, is necessary. The analysis of scattered light from a finite object, in Section 2.3, illustrates that the fine details of the object's features, which is equivalent to the high spatial frequency components, are inherent in the non-propagating wave. Such a wave confines itself at the surface of the object and is not collected by a far-field optical detector.

The system for detecting the evanescent field will be explained in Section 2.4. Basically the system consists of two main components: the SNOM probe and the height control system. Firstly it requires a subwavelength probe to collect the localised light information from the object. Various types of SNOM probes are available. However, the widely utilised probe design is a tapered fibre with or without metal coating. The tapering process can be performed by chemical etching or heating and pulling techniques. The probe is required to be positioned into the near-field region, tens of nanometres from the surface of the waveguide, in order to detect the evanescent field and transform it into the propagating field to be processed by an imaging system. As a result, the height control system is imperative. Several techniques have been applied in SNOM research but the most adopted technique is probably based on detecting the shear force. Since the shear force interaction is distance dependent, a feedback system measuring the strength of the interaction can succeed in controlling the probe height.

The last section, Section 2.5, explains the interaction of the probe and the near-field which contributes to the intensity collected by the probe. Approximated theoretical models, of scattering from a small sphere and diffraction of light through a small hole, have been adopted to represent the light collecting mechanisms of the uncoated and metal coated probes respectively. However, more rigorous analysis should include the coupling between the probe and the sample, especially in the case of metal coated probes or samples where the coupling becomes strong and cannot be neglected. Some researchers have studied the propagation of ultrashort light pulses through a SNOM probe in which the pulses can be distorted in both time and frequency domains.

2.2 Diffraction-limited resolution

The resolution of far-field optics is limited by diffraction. The limitation was firstly described by Abbe [1] and later rearranged by Rayleigh as the Rayleigh criterion [2–5], which can be expressed as

$$r \geq \frac{1.22\lambda}{2n \sin \theta} \quad (2.1)$$

where r is the separation distance between the two points, λ is the wavelength, n is the refractive index of the medium in which the light from the object propagates, and θ is the half angle of light cone accepted by the imaging system.

From Equation 2.1, the observable separation between two points can be approximated to be around half a wavelength, for observation in air, although in practice the resolution of a conventional imaging system is less than that due to the quality limitation of optical devices such as aberrations. The resolution can be improved by decreasing the operation wavelength (electron or ultraviolet microscopy) or increasing the numerical aperture $n \sin \theta$ (increasing the size of lens or the oil-immersion objective) [6, 7]. However, detection of the near-field, which confines itself within the surface of the sample, can provide much higher resolution which is not bound by the diffraction limitation.

2.3 Angular spectrum of diffracted field

It has been shown that scattered light from a finite object contains both propagating and non-propagating fields [8]. Only the former can be collected by a remote imaging system to give an image containing low spatial frequencies of the object. The latter, which relates to the finer details of the object, exponentially decays within a few tens of nanometres from the surface of the scattering object. Therefore the resolution beyond

the diffraction limit can be acquired if the non-propagating field is detected. The different nature of both fields can be clearly described by analysing the angular spectrum of the scattered field [9].

Consider a space-limited object at the plane $z = 0$. If the transmittance or reflectance of the object is described by $f(x, y, 0)$, then the spatial spectrum of the object is given by the Fourier transform of $f(x, y, 0)$ as

$$F(u, v) = \iint_{-\infty}^{+\infty} f(x, y, 0) \exp[-j2\pi(ux + vy)] dx dy \quad (2.2)$$

where u and v are the spatial frequencies of the object along the x and y axes, respectively. If a unit plane wave is incident on the object, the transmitted or reflected field $U(x, y, 0)$ will be equal to $f(x, y, 0)$. Therefore, $U(x, y, 0)$ can be expressed as the inverse 2D Fourier transform of Equation 2.2.

$$\begin{aligned} U(x, y, 0) &= \iint_{-\infty}^{+\infty} F(u, v) \exp\left[j\frac{2\pi}{\lambda}((u\lambda)x + (v\lambda)y)\right] du dv \\ &= \iint_{-\infty}^{+\infty} F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) \exp[jk(\alpha x + \beta y)] du dv \end{aligned} \quad (2.3)$$

where

$$\alpha = \lambda u \quad (2.4a)$$

$$\beta = \lambda v \quad (2.4b)$$

Equation 2.3 physically means that the field $U(x, y, 0)$ is actually the superposition of plane waves propagating with direction cosine (α, β, γ) with $\gamma = (1 - \alpha^2 - \beta^2)^{1/2}$. The term $F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)$ gives the angular spectrum of the field $U(x, y, 0)$. Then if this field propagates by a distance z , the field at the new plane can be written as

$$U(x, y, z) = \iint_{-\infty}^{+\infty} F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) \exp[jk(\alpha x + \beta y)] du dv \quad (2.5)$$

where the angular spectra $F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right)$ are related to $F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)$ by the Helmholtz equation [10]

$$\nabla^2 U + k^2 U = 0 \quad (2.6)$$

Substituting Equation 2.5 into Equation 2.6 gives

$$\frac{d^2}{dz^2} F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) + k^2 [1 - \alpha^2 - \beta^2] F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) = 0 \quad (2.7)$$

The solution of the differential equation Equation 2.7 is given by

$$F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) = F\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) \exp\left(-jk(1 - \alpha^2 - \beta^2)^{1/2} z\right) \quad (2.8)$$

Equation 2.8 represents the propagating field along the z axis as long as $(1 - \alpha^2 - \beta^2)^{1/2}$ is real. That is,

$$\alpha^2 + \beta^2 < 1 \quad \text{or} \quad u^2 + v^2 < \frac{1}{\lambda^2} \quad (2.9)$$

Therefore, the evanescent field, which has rapid attenuation along the z axis and confines itself within the xy plane, is related to the condition

$$\alpha^2 + \beta^2 > 1 \quad \text{or} \quad u^2 + v^2 > \frac{1}{\lambda^2} \quad (2.10)$$

The condition in Equation 2.9 specifies that only lower spatial frequencies of the scattered light can propagate and are detected by a remote imaging system. However, high spatial frequencies are inherent in the non-propagating field, whose wavenumber is complex, as shown in Equation 2.10. This non-propagating field exponentially decays within a few tens of nanometres from the surface of the scattering object. The rapidness of decay depends on the finesse of the object's features. Higher spatial frequency components will be attained if a detector is brought closer to the surface.

2.4 Scanning near-field imaging system

The basic elements of a scanning near-field imaging system are a nanoprobe and a height control unit to keep the probe in the near-field region. Both, together with SNOM operation modes, will be described in more detail in this section.

2.4.1 Operation modes

SNOM can be operated in various configurations. The choice depends on the types of samples and optical phenomena to be observed. For transparent samples, SNOM can be operated in the illumination or collection modes, as shown in Figure 2.1. In

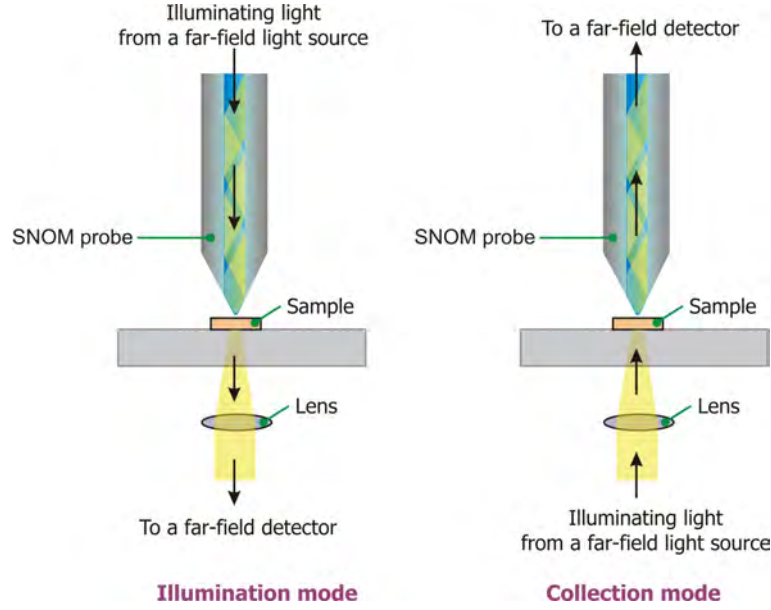


FIGURE 2.1: Schematic diagrams of illumination and collection modes

the illumination mode, light is coupled to the probe and the sample is illuminated by the output light from the end of the tip. Then, transmitted light from the sample is passed through a far-field detector. On the contrary, in the case of the collection mode, the direction of light is reversed and the locations of the source and the detector are interchanged. That is, the sample is illuminated by a far-field source and the transmitted light is detected by the near-field probe. Owing to the reciprocity, both schemes are theoretically equivalent and the resolution, which depends on the probe size, is not expected to be different. Nevertheless, both the difference in the size of the illuminated area and the type of fields (near-field for the illumination mode and far-field for the collection mode) can excite the effect of surface plasmons differently [11, 12]. The SNOM images acquired experimentally from both operation modes, however, display no significant difference [12].

On the other hand, for an opaque sample, the reflected light can be detected instead. Fisher et al. applied a subwavelength pinhole as an aperture for the scattered light from the sample, which was illuminated by the diffracted light from the pinhole itself [13, 14]. Alternatively the illumination can be performed by the light beam from the far-field light source and then the reflected light is collected by a fibre probe [15], or the function of the probe is reversed as an illuminator and the reflected light is collected by a remote detector [16, 17]. However, in this case, Cline et al. suggested that a shadowing effect could be caused by the different nature of reflection within a cavity formed between the aluminium coating of the probe and the sample, before the light gets to the detector [18]. On the other hand, the same fibre tip can both illuminate the sample and also collect the reflected light [9]. One drawback of the reflection configuration is the interference in the detected optical signal with the illuminating light. Additionally the reflected light from the probe, or coupling lens, can reduce the signal-to-noise ratio, and therefore this

lowers the resolution and image quality. The approach can be improved by analysing the light signal with cross-polarised detection [19, 20] in which a polariser is used with its axis rotated 90° from the polarisation of the incident light. The discrimination is fulfilled owing to the conversion of the polarisation of the reflected light from the sample's surface.

Another configuration is based on the principle of frustrated internal reflection. By placing a dielectric probe within the skin depth, the evanescent field can be transformed into a propagating wave which is collected by the probe itself and sent to a remote detector [21–28]. Since the detection is associated with the phenomenon of the optical tunneling, the configuration is normally referred to as photon scanning tunneling microscopy (PSTM).

2.4.2 Near-field probe

2.4.2.1 Probe shape

The nanoaperture can play a role as a near-field collector or emitter. In the case of the collection mode, a sample is illuminated by a far-field light source. By optical tunneling, the scattered evanescent field is converted to the propagating field which passes through the aperture of the near-field probe and is collected further away by a remote detector. Alternatively the sample can be illuminated by the evanescent field exiting from the aperture and the scattered light is collected by far-field optics. Both regimes are theoretically equivalent but practically the excited area of the sample in collection mode is much larger than that in illumination mode.

Several designs of apertures have been available. The most basic shape is a pinhole, with diameters ranging from 50 nm to 1 μm , on a glass substrate coated by a metal film [13, 29]. The disadvantage of such an aperture is that due to the large dimension of the screen it can easily make contact with the sample's surface. Another more favourable candidate is a quartz rod that is etched by hydrofluoric acid (HF) to make a sharp point and then coated with a thin gold film [30]. The aperture is created at the end face by pressing the tip on a microscope slide until the gold film gets thinner and becomes transparent. This type of near-field probe can give good results but the fabricating process is not reproducible.

Firstly introduced by Betzig et al., the most widely-used probe is a tapered fibre [31]. The tapering process can be done by chemical etching or heating and pulling. In the primitive etching technique, called meniscus etching [32], a bare fibre is submerged into HF that is covered by an organic protective layer such as toluene. The layer forms a meniscus etching area which determines the taper shape, as shown in Figure 2.2. Another tapering technique applies a tungsten coil or focused CO_2 laser beam to melt a fiber

which is later pulled by a tensile force. The latter technique is adopted in this research to fabricate a SNOM probe. Several parameters, such as the power of the heating source, the velocity of the pulling and the strength of the force, contribute towards defining the taper shape and its quality. Rigorous experimental and theoretical studies can be found in the literature [33, 34].

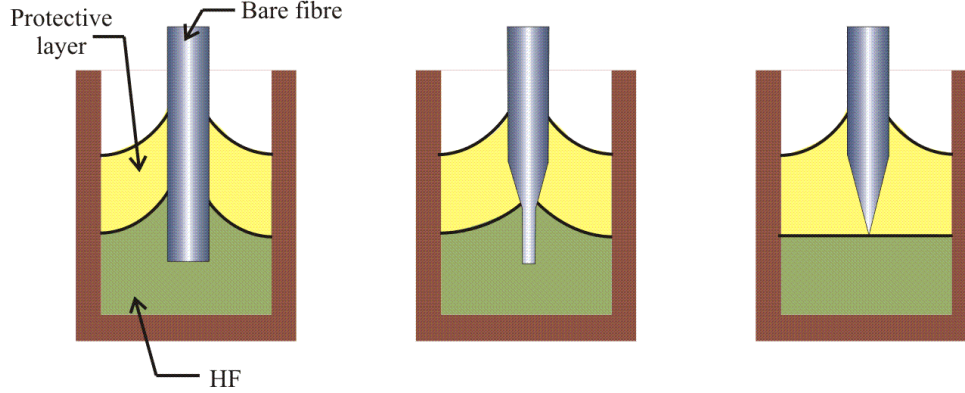


FIGURE 2.2: Meniscus etching method

Both techniques are quite competitive. Transmission efficiency is a crucial quality for near-field probes. That is because the field amplitude of the evanescent field detected by the probe has a tremendously small signal-to-noise ratio. A small change in the amount of light throughput can reduce the achievable resolution. The efficiency depends on the tip shape parameters: the cone angle and the aperture size. The etching technique can produce a tip with a greater cone angle and has less affect on the core diameter. Novotny et al. theoretically showed that such a shape provides better optical throughput [35]. A transmission efficiency of 10^{-3} has been reported for tips fabricated by the etching method [36] whereas 10^{-5} - 10^{-6} is recorded for the heating and pulling technique [37]. Note that the transmission efficiency is defined as the ratio of the output optical power at the end of the tip, measured by a far-field detector such as a photomultiplier tube (PMT), to the launched power coupled into the probe [38]. However the shape and surface roughness of the tip are highly susceptible to the etching environment such as temperature or acid concentration. This drawback can be reduced by the selective etching technique in which the polymer cladding of the fibre is not peeled off but left to form the enclosure [39]. Because the etching takes place inside the polymer cladding, the fluctuation of environmental conditions is diminished. Nevertheless the hazardous etchant and long processing times are still disadvantages of the etching method.

There are still more varieties of probe shape utilised among SNOM researchers, such as a nanoparticle, a dielectric tip, a metal tip, or an AFM tip [40–45]. This kind of apertureless probe behaves like a scattering centre perturbing the global field and the outcome is a scattered signal that is detected by remote optics [11].

2.4.2.2 Metal coated or uncoated probes

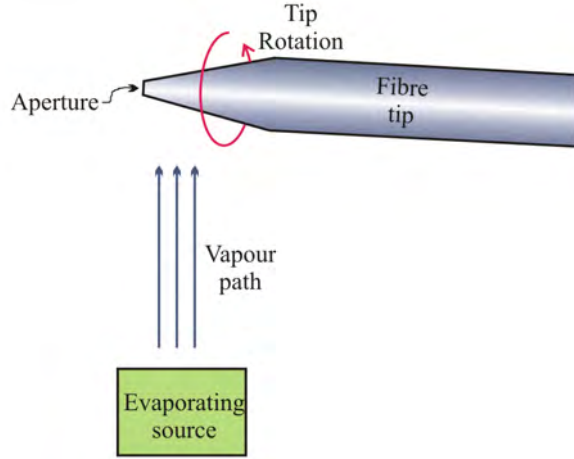


FIGURE 2.3: Metal coating on the side walls of a tapered fibre tip

There are two categories of probes which are utilised in the SNOM research community: uncoated or metal-coated probes. The former, which is the type of probe used in the experiments described in this thesis, behaves as the scattering point whereas the latter acts as the diffracting aperture.

After the tapering process, the probe is ready to use directly or can be further passed through the metal-coating process in which the side wall of the taper region is coated by a metal film. Aluminium is the most common coating material owing to its small skin depth and its ability to attach well to the surface of glass and silica. The process is arranged in such a way that the tip is rotated during vapour deposition and the very end is aligned away from the vapour path, as shown in Figure 2.3 [46]. This provides a smooth coating on the side walls without covering the extremity of the tip. Both types of probes dissimilarly interact with the sample and therefore can give different optical information. The details of the interaction between the probe and sample will be later discussed in Section 2.5.

2.4.3 Height controlling system

In order to keep the probe in the near-field region and avoid a collision between probe and surface, a controlling system for the probe height is required. The system monitors the gap distance and adjusts the height corresponding to the topography of the surface. Though several gap-dependent phenomena, such as electron or optical tunneling, can be applied for height control, the most effective technique is the shear-force control method. Several phenomena have been proposed to participate in the shear force, including van der Waals and capillary forces [47], nonlinear bending forces [48], and friction forces [49]. Some literature also demonstrates the dependence of shear force on humidity, which

causes contamination of water layers on the surface [50, 51]. In reality, the combination of various mechanisms influences the overall shear force, where the dominant contribution depends upon working conditions and types of surfaces.

To visualise the effect of the shear force, the probe is driven to oscillate at its resonant frequency. When the tip is brought down into the shear force range (a few tens of nanometres), the oscillation is damped and the amplitude is decreased by the amount depending on force strength and distance from the surface. Previously the amplitude variation has been inspected by optical means such as differential interferometry [47] or measuring the oscillation of the transmitted intensity from the probe [52]. These techniques are quite reliable but risk the interference of stray light with the near-field optical signal. Besides, the space requirement of the height-detecting system curtails the compactness of the whole system.

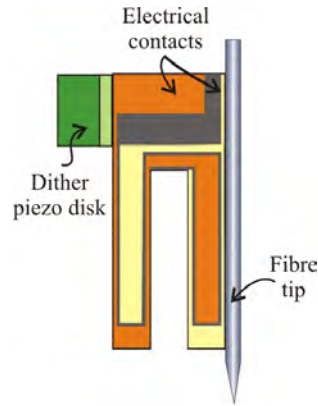


FIGURE 2.4: A quartz tuning fork with near-field fibre probe attached

Another technique, suggested by Karrai et al., is more widely employed among SNOM research communities [53, 54]. A fibre probe is attached to a piezoelectric tuning fork which is laterally dithered by a piezoelectric disk at the resonant frequency of the fork, as shown in Figure 2.4. The amplitude of oscillation is transformed into an electrical potential by the piezoelectric effect. Then the signal is fed into the feedback system to compare it with the dithering signal and the probe height is subsequently adjusted to maintain the desired gap distance.

The oscillating tip/tuning fork can be viewed as a damped oscillation system whose characteristic can be described by its frequency response, as shown in Figure 2.5 which depicts the energy or amplitude of oscillation of the system responding to various driving frequencies. The peak of the oscillation occurs when the driving frequency is equal to the resonant frequency of the system (f_0) and the response falls down when the driving frequency departs from the resonant frequency. The concentration of the energy of the oscillation around the resonance frequency can be measured in terms of the quality or Q -factor, which is defined as

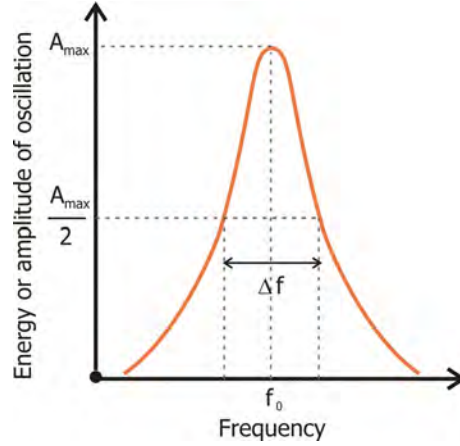


FIGURE 2.5: An example of frequency response of an oscillating system. The energy or amplitude of oscillation is plotted against various driving frequencies

$$Q = \frac{f_0}{\Delta f} \quad (2.11)$$

where f_0 and Δf are the resonant frequency and the bandwidth of the frequency response of the system measured at the half value of the maximum energy or amplitude as depicted in Figure 2.5. Normally a bare tuning fork has the resonance frequency at 32.768 kHz and the Q -factor ranges from 64000 in vacuum to 7500 in air [53]. However, both parameters change when a tip is attached and the Q -factor will drop down to hundreds or thousands.

The Q -factor corresponds to the sensitivity of the shear-force detection. The high value reflects the rapid drop of the amplitude of oscillation when the system is disturbed by a damping force. The system with a Q -factor of around a thousand can be responsive to drag forces of the order of a piconewton [53, 54]. However the Q -factor is relative to the relaxation time which affects the response time of the feedback system. Thus the high sensitivity comes at the expense of a slower system.

2.5 Interaction of probe and near-field

2.5.1 Various approaches for analysing the interaction

The theoretical study of the interaction of light in the near-field region with the probe and sample can be treated with two approaches: non-global and global [11]. The first treats the tip as a passive collector and its presence does not disturb the interaction of the optical field and the sample, whereas in the latter, the modification of the field distribution owing to the presence of both tip and sample is studied.

In the non-global scheme, several approaches have been utilised to describe the optical

intensity collected by the probe. The total energy can be calculated from integrating the flux of the Poynting vector along the tip surface. In the case of the dielectric probe functioning as a light collector, the extremity of the tip can be modeled as a sphere scattering the near-field through the probe. For a small radius of the sphere, the scattering of the field can be approximated as that of only a dielectric dipole, \mathbf{d} , induced by the incident field, \mathbf{E} , [55]

$$\mathbf{d} = \alpha \mathbf{E} \quad (2.12)$$

where α is the polarisability of the sphere depending on its radius, R , and dielectric constant, ε_t .

$$\alpha = \frac{(\varepsilon_t - 1)}{(\varepsilon_t + 2)} R^3 \quad (2.13)$$

The intensity collected by the electric dipole can be written as [55]

$$I(x, y, z) = \frac{\omega^4}{96c^3} \left[(|d_x|^2 + |d_y|^2) (16 - 15 \cos \delta - \cos 3\delta) + |d_z|^2 (16 - 18 \cos \delta + 2 \cos 3\delta) \right] \quad (2.14)$$

where d_x , d_y and d_z are the x -, y - and z -components of \mathbf{d} , ω is the angular frequency, c is the speed of light and δ is the collection angle, and the coordinates x , y , and z are defined such that x and y are on the cross-section plane whereas z is along the length of the probe.

Unlike the case when the incident field is the propagating field, Barchiesi et al. showed that the electrical Mie coefficients (the weight of the interaction) for multipolar terms do not decrease monotonically and the magnetic terms are also enhanced [56]. Therefore a more accurate study was performed by taking the scattering of the sphere as the Mie's scattering of a multipolar system. However the model of dipolar scattering is accurate enough as long as $R < 200$ nm [55].

On the other hand, the approach of light diffraction through a subwavelength aperture should be applied for an aperture probe such as a tapered fibre with metal coating. Diffraction through a small hole had been studied theoretically by Bethe [57] and the model later on corrected by Bouwkamp [58]. Therefore the model was generally referred to the Bethe-Bouwkamp model. In contrast to the case of a dielectric tip, the detected intensity for this kind of probe also depends on the magnetic field. For a circular aperture with radius R , the intensity can be written as [55]

$$\begin{aligned}
I(x, y, z) = & \frac{\omega^4}{96c^3} \left[\left(|\mu_x|^2 + |\mu_y|^2 + |m_x|^2 + |m_y|^2 \right) (16 - 15 \cos \delta - \cos 3\delta) \right. \\
& + \left(|\mu_z|^2 + |m_z|^2 \right) (16 - 18 \cos \delta + 2 \cos 3\delta) \\
& \left. + 6 (\cos 2\delta - 1) (m_x d_y^* + d_y m_x^* - m_y d_x^* - d_x m_y^*) \right] \quad (2.15)
\end{aligned}$$

where $\mu_{x,y,z}$ and $m_{x,y,z}$ are the components of the effective electric and magnetic dipole moments which are induced by the electric field $\mathbf{E}(x, y, z)$ and magnetic field $\mathbf{B}(x, y, z)$ respectively. The diffracted intensity does not only depend on the sum of intensity from each kind of dipole but also on the interference from both, as can be seen in the last term on the right-hand side of Equation 2.15. For the polarisation of electric field perpendicular to the aperture plane and therefore the magnetic field lines on the xy -plane, both dipole moments can be determined by [55]

$$\mu_x = \mu_y = 0 \quad (2.16a)$$

$$\mu_z = -\frac{2R^3}{3\pi} E_z(x, y, z) \quad (2.16b)$$

$$m_x = \frac{4R^3}{3\pi} B_x(x, y, z) \quad (2.16c)$$

$$m_y = \frac{4R^3}{3\pi} B_y(x, y, z) \quad (2.16d)$$

$$m_z = 0 \quad (2.16e)$$

The coupling between the tip and the sample depends on the distance between the probe and the sample, the polarisability of the tip, and the amplitude surface reflection factor for the sample in air $(\varepsilon_s - 1) / (\varepsilon_s + 1)$, where ε_s is the dielectric constant of the sample [55, 59]. For the case of both dielectric probe and sample, the coupling is weak even if the gap distance is 8 times shorter than the radius of the tip [60]. However, the coupling becomes strong when the tip or sample is metal or both are metal. Moreover, the polarisation of incident light also affects the detected intensity. In such a case, numerical techniques in the global scheme such as finite difference time domain (FDTD) [61–63], multiple multipole method (MMP) [64] or the differential method [60] have been utilised to analyse the interaction of the tip and sample in the near-field region.

2.5.2 Ultrashort light with SNOM probes

There have been several experimental and theoretical studies of ultrashort light propagation through both metal-coated and uncoated tapered fibre probes [65–69]. The

temporal profile of a femtosecond light pulse passing through an aluminium coated probe was experimentally measured by Kawashima et al. and the output is found to be slightly distorted [66]. However, a numerical simulation utilising the FDTD method demonstrates the propagation of a 10 fs Gaussian pulse through a trapezoidal-shaped fibre probe placed in front of a 100 nm thick GaAs layer [67]. The result demonstrated that the temporal shape of the pulse for the *s*-polarisation (perpendicular to the plane of incidence) pulse through a probe coated with perfectly conducting metal, alters noticeably. Additionally, the pulse spectral width was narrowed by 40% and the spectra blue shifted by more than 15 nm. In the same study, on the contrary, no such temporal and spectral deviation is observable in the output from the case of *p*-polarisation (parallel to the plane of incidence) or any kinds of polarisation for an uncoated probe. Similar deviations in temporal and spectral profile were also observed in another numerical study applying the finite integration technique (FIT), although the changes were much less than in Muller's study [69]. Furthermore the phase and group velocity of the pulses also change along the central axis of a metal coated probe, and even the negative values of those velocities are present at some locations inside and outside the probe. The origins for such characteristics, as suggested by Pack et al., are probably due to the multiple reflections at the metal coating and the propagation of surface plasmon polaritons (SPP), which are excited at the interface of metal and dielectric material [69].

2.6 Chapter conclusion

The theoretical basis of SNOM has been described in this chapter. In a conventional far-field imaging system, the resolution is limited by the diffraction and results in the resolving power being bound by $\sim \lambda/2$. By a simple angular spectrum analysis of scattered light from a finite object, the propagating field detected by the far-field detector contains only low spatial frequency components of the scattered field whereas high spatial frequency components belong to the evanescent field, confined above the surface of the object. As a result, detecting this non-propagating field will give subwavelength, localised optical information about the object.

Basically, the near-field detection system consists of the subwavelength probe and the height control system that is maintaining the probe within very close proximity to the surface of the sample. Several types of probe have been used in SNOM research. However, the tapered fibre is probably the most widely adopted. The tapered probe can be used directly, or otherwise its side walls can be coated with a metal layer to form an aperture at the extremity of the tip. Many distance-dependent mechanisms can be applied for the height detection system and one of these is the shear force technique in which the interaction of shear force on the probe oscillation is monitored.

The characteristics of the field collected by the probe depends on the interaction of the

near-field, the probe and the sample. The light collection mechanisms of the uncoated probe can be explained approximately by a scattering model whereas that of a metal coated probe is based on the theory of diffraction through a small aperture. For the case of a dielectric probe and a dielectric sample, the coupling between the probe and the sample can be negligible. However, the coupling becomes strong when either probe or sample becomes a metal or when both are. In such a case, and where ultrafast light is utilised, several research studies have shown some changes in the temporal and spectral profiles in the field collected from the input light pulse, and the strong dependence on the polarisation.

Since the SNOM system in this research is applied to visualise the spectra of ultrashort pulses propagating in a waveguide, an uncoated probe type is appropriately chosen as the SNOM probe to minimise any artifacts introduced in the spectral data. Further details of probe and other components of the SNOM system will be fully given in the next chapter.

2.7 References

- [1] E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9:413–420, 1873.
- [2] J. W. Strutt (Lord Rayleigh). On the manufacture and theory of diffraction-gratings. *Philosophical Magazine*, 47:81–93; 193–205, 1874.
- [3] J. W. Strutt (Lord Rayleigh). Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*, 8:261–274; 403–411; 477–486;, 1879.
- [4] J. W. Strutt (Lord Rayleigh). Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine*, 9:40–55, 1880.
- [5] J. W. Strutt (Lord Rayleigh). On the theory of optical instruments, with special reference to the microscope. *Philosophical Magazine*, 42:167–195, 1896.
- [6] E. Hecht. *Optics*. Addison Wesley, 4th edition, 2002.
- [7] F. L. Pedrotti and L. S. Pedrotti. *Introduction to Optics*. Prentice-Hall, Inc., 2nd edition, 1993.
- [8] E. Wolf and M. Nieto-Vesperinas. Analyticity of the angular spectrum amplitude of scattered fields and some of its consequences. *Journal of Optical Society of America*, 2(6):886–889, 1985.
- [9] D. Courjon, J.-M. Vigoureux, M. Spajer, K. Sarayedine, and S. Leblanc. External and internal reflection near-field microscopy: experiments and results. *Applied Optics*, 29(26):3734 – 3740, 1990.

- [10] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, Inc., USA.
- [11] D. Courjon. *Near-Field Microscopy and Near-Field Optics*. Imperial College Press., London.
- [12] K. Imura and H. Okamoto. Reciprocity in scanning near-field optical microscopy: illumination and collection modes of transmission measurements. *Optics Letters*, 31(10):1474–1476, 2006.
- [13] U. Ch. Fischer. Submicrometer aperture in a thin metal film as a probe of its microenvironment through enhanced light scattering and fluorescence. *Journal of the Optical Society of America B (Optical Physics)*, 3(10):1239–1244, 1986.
- [14] U. Ch. Fischer, U. T. Dürig, and D. W. Pohl. Near-field optical scanning microscopy in reflection. *Applied Physics Letters*, 52(4):249–251, 1988.
- [15] N. Umeda, Y. Hayashi, K. Nagai, and A. Takayanagi. Scanning Wiener-fringe microscope with an optical fiber tip. *Applied Optics*, 31(22):4515–4518, 1992.
- [16] T. D. Harris, R. D. Grober, J. K. Trautman, and E. Betzig. Super-resolution imaging spectroscopy. *Applied Spectroscopy*, 48(1):14A–21A, 1994.
- [17] C. Durkan and I. V. Shvets. 40 nm resolution in reflection-mode SNOM with $\lambda = 685$ nm. *Ultramicroscopy*, 61(1-4):227–231, 1995.
- [18] J. A. Cline and M. Isaacson. Probe-sample interactions in reflection near-field scanning optical microscopy. *Applied Optics*, 34(22):4869–4876, 1995.
- [19] S.I. Bozhevolnyi, O. Keller, and M. Xiao. Control of the tip-surface distance in near-field optical microscopy. *Applied Optics*, 32(25):4863–4868, 1993.
- [20] S. I. Bozhevolnyi, M. Xiao, and O. Keller. External-reflection near-field optical microscope with cross-polarized detection. *Applied Optics*, 33(5):876–880, 1994.
- [21] D. Courjon, K. Sarayedine, and M. Spajer. Scanning tunneling optical microscopy. *Optics Communications*, 71(1-2):23–28, 1989.
- [22] R. C. Reddick, R. J. Warmack, and T. L. Ferrell. New form of scanning optical microscopy. *Physical Review B (Condensed Matter)*, 39(1):767–770, 1989.
- [23] M. A. Paesler, P. J. Moyer, C. J. Jahncke, C. E. Johnson, R. C. Reddick, R. J. Warmack, and T. A. Ferrell. Analytical photon scanning tunneling microscopy [surface topography and stresses]. *Physical Review B (Condensed Matter)*, 42(10):6750–6755, 1990.
- [24] C. Girard and D. Courjon. Model for scanning tunneling optical microscopy: a microscopic self-consistent approach. *Physical Review B (Condensed Matter)*, 92(15):9340–9349, 1990.

- [25] P. J. Moyer, C. L. Jahncke, M. A. Paesler, R. C. Reddick, and R. J. Warmack. Spectroscopy in the evanescent field with an analytical photon scanning tunneling microscope. *Physics Letters A*, 145(6-7):343–347, 1990.
- [26] R. C. Reddick, R. J. Warmack, D. W. Chilcott, S. L. Sharp, and T. L. Ferrell. Photon scanning tunneling microscopy. *Review of Scientific Instruments*, 61(12):3669–3677, 1990.
- [27] J. M. Guerra. Photon tunneling microscopy. *Applied Optics*, 29(26):3741–3752, 1990.
- [28] T. L. Ferrell, J. P. Goundonnet, R. C. Reddick, S. L. Sharp, and R. J. Warmack. The photon scanning tunneling microscope. *Journal of Vacuum Science & Technology B (Microelectronics Processing and Phenomena)*, 9(2):525–530, 1991.
- [29] U. Ch. Fischer. Optical characteristics of 0.1 μm circular apertures in a metal film as light sources for scanning ultramicroscopy. *Journal of Vacuum Science & Technology B (Microelectronics Processing and Phenomena)*, 3(1):386–390, 1985.
- [30] D. W. Pohl, W. Denk, and M. Lanz. Optical stethoscopy: image recording with resolution $\lambda/20$. *Applied Physics Letters*, 44(7):651–653, 1984.
- [31] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak. Breaking the diffraction barrier: optical microscopy of a nanometric scale. *Science*, 251(5000):1468–1470, 1991.
- [32] D. R. Turner. United States Patent 4469554: Etch procedure for optical fibers.
- [33] G. A. Valaskovic, M. Holton, and G. H. Morrison. Parameter control, characterization, and optimization in the fabrication of optical fiber near-field probes. *Applied Optics*, 34(7):1215–1218, 1995.
- [34] R. L. Williamson and M. J. Miles. Melt-drawn scanning near-field optical microscopy probe profiles. *Journal of Applied Physics*, 80(9):4804–4812, 1996.
- [35] L. Novotny, D. W. Pohl, and B. Hecht. Scanning near-field optical probe with ultrasmall spot size. *Optics Letters*, 20(9):970–972, 1995.
- [36] D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi. Pulsed laser-induced desorption and optical imaging on a nanometer scale with scanning near-field microscopy using chemically etched fiber tips. *Applied Physics Letters*, 68(18):2491–2492, 1996.
- [37] Y. D. Suh and R. Zenobi. Improved probes for scanning near-field optical microscopy. *Advanced Materials*, 12(15):1139–1142, 2000.
- [38] Near-field optics : Principles and applications. In X. Zhu and M. Ohtsu, editors, *The Second Asia-Pacific Workshop on Near Field Optics*, Beijing, 1999. World Scientific.

- [39] R. Stöckle, C. Fokas, V. Deckert, R. Zenobi, B. Sick, B. Hecht, and U. P. Wild. High-quality near-field optical probes by tube etching. *Applied Physics Letters*, 75(2):160–162, 1999.
- [40] M. Specht, J. D. Pedarnig, W. M. Heckl, and T. W. Hansch. Scanning plasmon near-field microscope. *Physical Review Letters*, 68(4):476–479, 1992.
- [41] Y. Inouye and S. Kawata. Near-field scanning optical microscope with a metallic probe tip. *Optics Letters*, 19(3):159–161, 1994.
- [42] F. Zenhausern, M.P. O’Boyle, and H. K. Wickramasinghe. Apertureless near-field optical microscope. *Applied Physics Letters*, 65(13):1623–1625, 1994.
- [43] P. Gleyzes, A. C. Boccara, and R. Bachelot. Near field optical microscopy using a metallic vibrating tip. *Ultramicroscopy*, 57(2-3):318–322, 1995.
- [44] R. Bachelot, P. Gleyzes, and A. C. Boccara. Near-field optical microscope based on local perturbation of a diffraction spot. *Optics Letters*, 20(18):1924–1926, 1995.
- [45] R. Bachelot, P. Gleyzes, and A. C. Boccara. Apertureless near field optical microscopy by local perturbation of a diffraction spot. *Ultramicroscopy*, 61(1-4):111–116, 1995.
- [46] D. Courjon and C. Bainier. Near field microscopy and near field optics. *Reports on progress in physics*, 57(10):989–1028, 1994.
- [47] R. Toledo-Crow, P. C. Yang, Y. Chen, and M. Vaez-Iravani. Near-field differential scanning optical microscope with atomic force regulation. *Applied Physics Letters*, 60(24):2957–2959, 1992.
- [48] M. J. Gregor, P. G. Blome, J. Schöfer, and R. G. Ulbrich. Probe-surface interaction in near-field optical microscopy: the nonlinear bending force mechanism. *Applied Physics Letters*, 68(3):307–309, 1996.
- [49] D. A. Lapshin, E. E. Kobylkin, and V. S. Letokhov. Shear force distance control in near-field optical microscopy: experimental evidence of the frictional probe-sample interaction. *Ultramicroscopy*, 83(1-2):17–23, 2000.
- [50] R. Brunner, O. Marti, and O. Holtricher. Influence of environmental conditions on shear-force distance control in near-field optical microscopy. *Journal of Applied Physics*, 86(12):7100–7106, 1999.
- [51] P. K. Wei and W. S. Fann. The effect of humidity on probe-sample interactions in near-field scanning optical microscopy. *Journal of Applied Physics*, 87(5):2561–2564, 2000.
- [52] E. Betzig, J. S. Weiner, and P. L. Finn. Combined shear force and near-field scanning optical microscopy. *Applied Physics Letters*, 60(20):2484–2486, 1992.

- [53] K. Karrai and R. D. Grober. Piezoelectric tip-sample distance control for near field optical microscopes. *Applied Physics Letters*, 66(14):1842–1844, 1995.
- [54] K. Karrai and R. D. Grober. Piezo-electric tuning fork tip-sample distance control for near field optical microscopes. *Ultramicroscopy*, 61(1-4):197–205, 1995.
- [55] D. Van Labeke and D. Barchiesi. Probes for scanning tunneling optical microscopy: a theoretical comparison. *Journal of the Optical Society of America A (Optics and Image Science)*, 10(10):2193–2201, 1993.
- [56] D. Barchiesi and D. Van Labeke. Application of Mie scattering of evanescent waves to scanning tunnelling optical microscopy theory. *Journal of Modern Optics*, 40(7):1239–1254, 1993.
- [57] H. A. Bethe. Theory of diffraction by small holes. *Physical Review*, 66(7-8):163–182, 1944.
- [58] C. J. Bouwkamp. On Bethe’s theory of diffraction by small holes. *Philips Research Reports*, 5(5):321–332, 1950.
- [59] L. Novotny and S. J. Stranick. Near-field optical microscopy and spectroscopy with pointed probes. *Annual Review of Physical Chemistry*, 57:303–331, 2006.
- [60] J. C. Weeber, F. de Fornel, and J. P. Goudonnet. Numerical study of the tip-sample interaction in the photon scanning tunneling microscope. *Optics Communications*, 126(4-6):285–292, 1996.
- [61] D. A. Christensen. Analysis of near field tip patterns including object interaction using finite-difference time-domain calculations. *Ultramicroscopy*, 57(2-3):189–195, 1995.
- [62] R. X. Bian, R. C. Dunn, X. S. Xie, and P. T. Leung. Single molecule emission characteristics in near-field microscopy. *Physical Review Letters*, 75(26):4772–4775, 1995.
- [63] H. Furukawa and S. Kawata. Analysis of image formation in a near-field scanning optical microscope: effects of multiple scattering. *Optics Communications*, 132(1-2):170–178, 1996.
- [64] L. Novotny, D. W. Pohl, and P. Regli. Light propagation through nanometer-sized structures: the two-dimensional-aperture scanning near-field optical microscope. *Journal of the Optical Society of America A (Optics and Image Science)*, 11(6):1768–79, 1994.
- [65] A. Lewis, U. Ben-Ami, N. Kuck, G. Fish, D. Diamant, L. Lubovsky, K. Lieberman, S. Katz, A. Saar, and M. Roth. NSOM the fourth dimension: integrating nanometric spatial and femtosecond time resolution. *Scanning: Journal of Scanning Microscopies*, 17(1):3–13, 1995.

- [66] H. Kawashima, M. Furuki, and T. Tani. Interferometric measurement of femtosecond optical pulses emitted from a fibre probe. *Journal of Microscopy*, 194(.2-3): 516–518, 1999.
- [67] R. Müller and C. Lienau. Propagation of femtosecond optical pulses through uncoated and metal-coated near-field fiber probes. *Applied Physics Letters*, 76(23): 3367–3369, 2000.
- [68] R. Müller and C. Lienau. Three-dimensional analysis of light propagation through uncoated near-field fibre probes. *Journal of Microscopy*, 202(2):339–346, 2001.
- [69] A. Pack, M. Hietschold, and R. Wannemacher. Propagation of femtosecond light pulses through near-field optical aperture probes. *Ultramicroscopy*, 92(3-4):251–264, 2002.

Chapter 3

Construction of a SNOM system

3.1 Chapter introduction

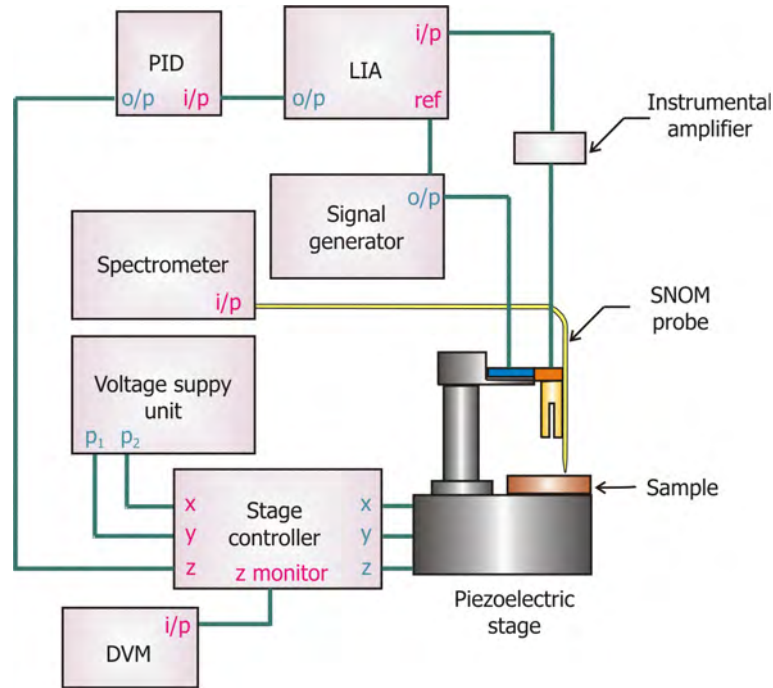


FIGURE 3.1: Schematic diagram of the SNOM system used during this research

In this chapter, the construction of a SNOM system applied in this research to visualise the spectral evolution inside a nonlinear waveguide will be presented. The SNOM system in this research is arranged as shown in Figure 3.1. The signal generator feeds the oscillating voltage to the piezoelectric disc which sits beside and drives the tuning fork. The mechanical oscillation motion of the tuning fork is transformed to an electrical signal (through the piezoelectric effect) which is measured by a lock-in amplifier (LIA) and compared with the original dithering signal. The error signal is then fed into the Proportional-Integral-Derivative (PID) unit which provides a signal to correct

the height of the probe. The probe is translated laterally along the sample plane (x and y coordinates) and vertically (z coordinate) by the stage controller. Note that the definition of the coordinates for the translation of the SNOM probe is different from the coordinate of the waveguide geometry which was described in Chapter 4. The spectral data is collected by the spectrometer. The topographical data is read by the digital voltmeter (DVM) connected to the z (height) channel of the stage controller. Each part of the system will be explained in detail in the following sections.

Firstly the fabrication of the SNOM probe is described in Section 3.2. Basically a fibre was heated by a CO₂ laser and then pulled to taper the melting section of the fibre. All these processes were performed by a commercial micropipette puller, whose instrumental parameters are required to be adjusted for producing a near-field probe. The probe needs to be brought into, and restrained within, the near-field region while it detects the evanescent spectra of the waveguide. The task is accomplished by the application of the feedback system controlling the distance-dependent interaction of the shear force with the probe oscillation. The probe dithering setup and the characteristics of the probe oscillation are explained in Section 3.3. When the probe is in close proximity to the waveguide surface, the oscillation of the probe is damped by the shear force where the amplitude of the oscillation is inversely proportional to the gap distance between the probe and the surface. As a result, the probe height is controlled by a feedback system which detects changes in the amplitude of the oscillating electrical signal and correspondingly adjusts the probe height. The details of the height control system are described in Section 3.4.

Lastly the scanning system controlling the lateral movement of the probe and the data acquisition will be described in Section 3.5, with a brief discussion relating to the vibration and thermal effects in Section 3.6.

3.2 Near-field probe fabrication

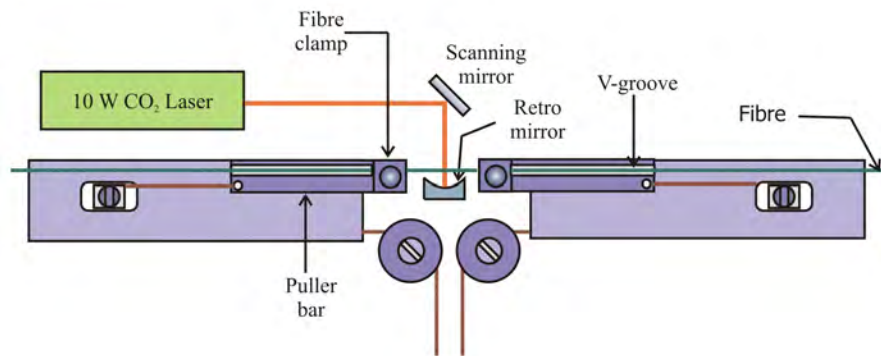


FIGURE 3.2: Diagram of a Sutter P-2000 micropipette puller [1]

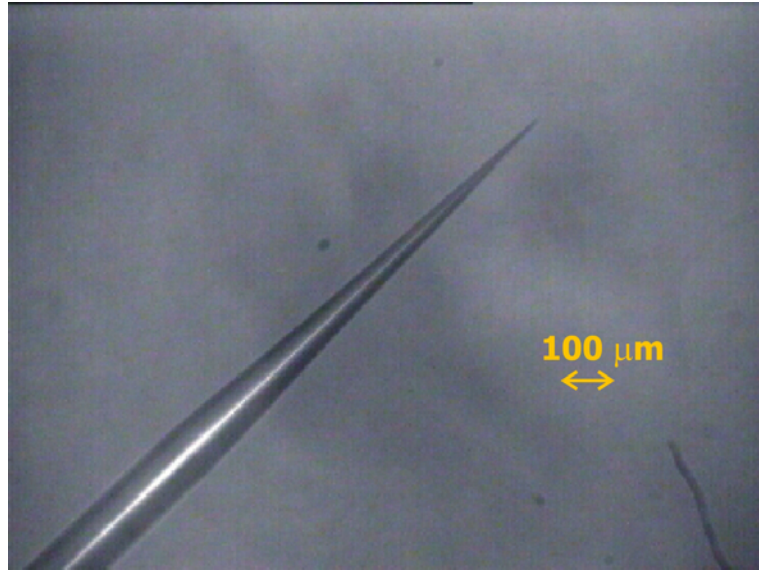
Parameter	Description	Used Value	Physical Value
HEAT	Power of CO ₂ laser	330	Laser power of ~ 2.5 W
FILAMENT	Longitudinal scan length of the laser beam	0	Scan length of 1 mm
VELOCITY	Velocity that the fibre carriage must reach before the hard pull	20	N/A
DELAY	Timing for the start of the hard pull	126	The hard pull starts 2 ms before the laser turns off
PULL	Force of the hard pull	150	N/A

TABLE 3.1: Puller parameters (arbitrary units) for pulling fibres as near-field probes. The used value is an integer unit with its corresponding physical value shown in the last column. More details of these parameters can be found in Appendix B

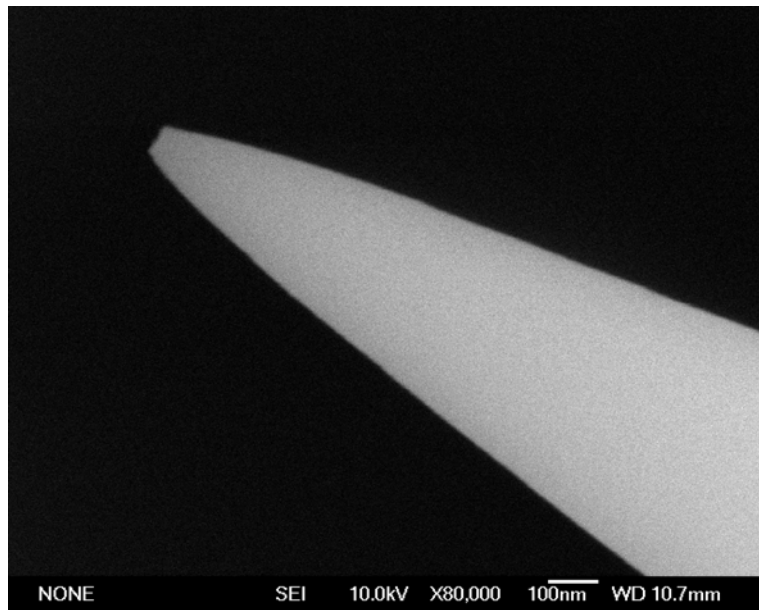
In this research, the fibre tip is produced by the heating and pulling method using a commercial micropipette puller (Sutter P-2000). The benefits of this method are that it does not require hazardous chemical etchants and it can provide a reliable tip within a fraction of second.

A diagram of the puller is shown in Figure 3.2 and the controlling parameters of the puller, whose values are integer units, are shown in Table 3.1. The fabrication starts with the heating process in which the laser light from the CO₂ laser, whose power is specified by the parameter **HEAT**, repetitively scans along a small section of the fibre. The process is achieved by a scanning mirror at the back of the fibre. The length of the scanning is specified by the parameter **FILAMENT**. The uniform heating, all over the surface of the fibre, is fulfilled by the retro mirror at the front which collects the reflected light from the scanning mirror and focuses it on the front part of the fibre. After the fibre melts, the carriage starts to pull apart with a velocity, that depends on the viscosity and therefore on the temperature of the heated area of the fibre. By monitoring this velocity, the hard pull can be applied to break the fibre at the right temperature and this is controlled by the parameter **VELOCITY**. The force of the hard pull and the time gap between the shut down of the laser and the application of the hard pull are specified by the parameters **DELAY** and **PULL** respectively. A greater force in the pull, and a longer time delay, results in a longer taper length of the tip and a smaller tip diameter. The values recommended, by the puller manufacturing company, for an optical fibre can be taken as guidance [1] to fabricate a fine tip with a diameter in the region of 100 μm and a taper length of ~ 1 cm. However, the parameters are suggested to be adjusted empirically to achieve the desired tip shape for various types of fibres. The value of the parameters, used in the probe fabrication of this research, is shown in the third column of Table 3.1 with the corresponding physical values described in the last column. The fibre utilised in this research is a single mode fibre with a cut-off of 665 nm and core

diameter of $4\text{ }\mu\text{m}$ with a cladding diameter of $125\text{ }\mu\text{m}$. The VELOCITY and PULL are not directly related to the velocity of the carriage and the force of the pull. The first specifies the voltage of the velocity transducer whereas the latter controls the current of the pull solenoid. The specific values of the velocity and the pull force are not available from the instrument company since they vary from one puller to another [2].



(a)



(b)

FIGURE 3.3: A fibre tip fabricated by the Sutter P-2000 puller. (a) A microscope image displaying the taper length of the tip $\sim 1\text{ cm}$ (b) An SEM closeup image of the tip diameter which is in the range of 100 nm [3]

Before loading the fibre on the puller, the polymer coating of a 3cm section at approximately 15 cm from one end of the fibre is removed. The peeled area is cleaned with methanol and then loaded on the puller's carriage. The fibre is placed into the V groove making sure that the stripped area is in the laser light path, as shown in Figure 3.2. Before clamping, the alignment of the fibre should be straight along the groove. This can guarantee a good quality tip because it provides symmetric heating and a uniform pull. The quality of the pulling can be checked by recording the time taken by the process. By experience, the processing time of 0.16 ± 0.1 seconds usually produces a good tip. Each pulling procedure produces two identical tips, one from each side of the ruptured fibre. The longer part is used as a near-field probe whereas the shorter one is for quality inspection under a microscope. Examples of tips pulled by the Sutter P-2000 puller can be seen in Figure 3.3.

In experiments requiring the optical image of a nanostructured sample, the tapered fibre probe is normally required to be metal coated to achieve the required high resolution. However, in this research the sample is a planar waveguide of the order of micron scales, and the probe is applied to collect only high resolution spectral information. As has been pointed out in Chapter 2, the metal coating could possibly affect the collected spectra and temporal profile. Moreover, the collected field is also polarisation dependent. Since the interaction of light inside the waveguide is the main objective of this research and the probe is expected to provide similar spectral information to what is happening inside the waveguide, such spectral alterations should be avoided and therefore the uncoated probe is adopted as a SNOM probe in this research. Indeed, since we are only concerned with scattering an evanescent field, our resolution will depend mainly on the size of our probe tip, which is around 100 nm.

3.3 Probe oscillation

The probe height control technique applied in this research is based on Karrai et al.[4, 5]. Basically the gap distance between the probe and the surface is monitored via the interaction of the shear force on the resonance of the oscillating probe. The variation in signal is processed and the outcome fed back to the piezoelectric stage, to adjust the height of the probe accordingly.

The oscillation of the probe is performed by attaching the probe to a piezoelectric tuning fork, as shown in Figure 3.4. The fibre probe is glued to one side of the tuning fork by cyanoacrylate (Superglue) and allowed to overhang by around 2-3 millimetres. The fork is actually a timing crystal whose outer tube was removed. It is laterally driven by the signal generator which supplies an oscillating voltage (around 1 volt peak-to-peak) to the piezoelectric disc attached nearby to the fork, which is in mechanical contact with the fork. The dithering signal is also fed to the reference input of the LIA (Stanford SR530

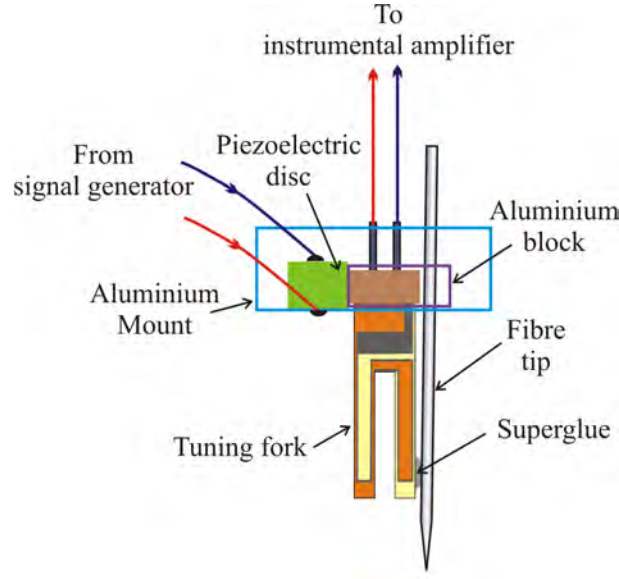


FIGURE 3.4: Probe dithering diagram

dual channel lock-In amplifier). The oscillating amplitude of the fork and also the probe is transformed to the voltage signal, pre-amplified by the instrumental amplifier, and then fed to the LIA. For a typical oscillating system of probe/tuning fork, the amplitude is of the order of 10^{-1} nm - 1 nm [4, 6–8].

The LIA is a phase-sensitive detector (PSD) which can measure small AC signals even when those are buried in noise with a much larger amplitude. The detector isolates the signal, which has the same frequency as the reference, and rejects the noise at other frequencies. The scheme is carried out by multiplexing the signal with the reference, which is the driving signal from the function generator in this case. The result consists of the sum and difference frequency components. If the frequencies of signal and reference are exactly the same, after passing the low-pass filter, only the DC component, whose amplitude is related to the magnitude of the signal (R) and the phase difference (Φ) between the signal and the reference, can be detected. The separation between R and Φ can be achieved by two PSDs. One multiplexes the signal with the reference and the other with its $\pi/2$ phase shift which gives out two DC signals $X = R \sin \Phi$ and $Y = R \cos \Phi$ after passing the low-pass filter. The amplitude and phase of the signal are derived by

$$R = \sqrt{X^2 + Y^2} \quad (3.1)$$

$$\Phi = \arctan \left(\frac{Y}{X} \right) \quad (3.2)$$

The characteristics of the oscillating tuning fork attached to a fibre tip at various frequencies can be seen in Figure 3.5. The response was measured when there was no sample. The X and Y curves are derived directly from the outputs of the LIA whereas

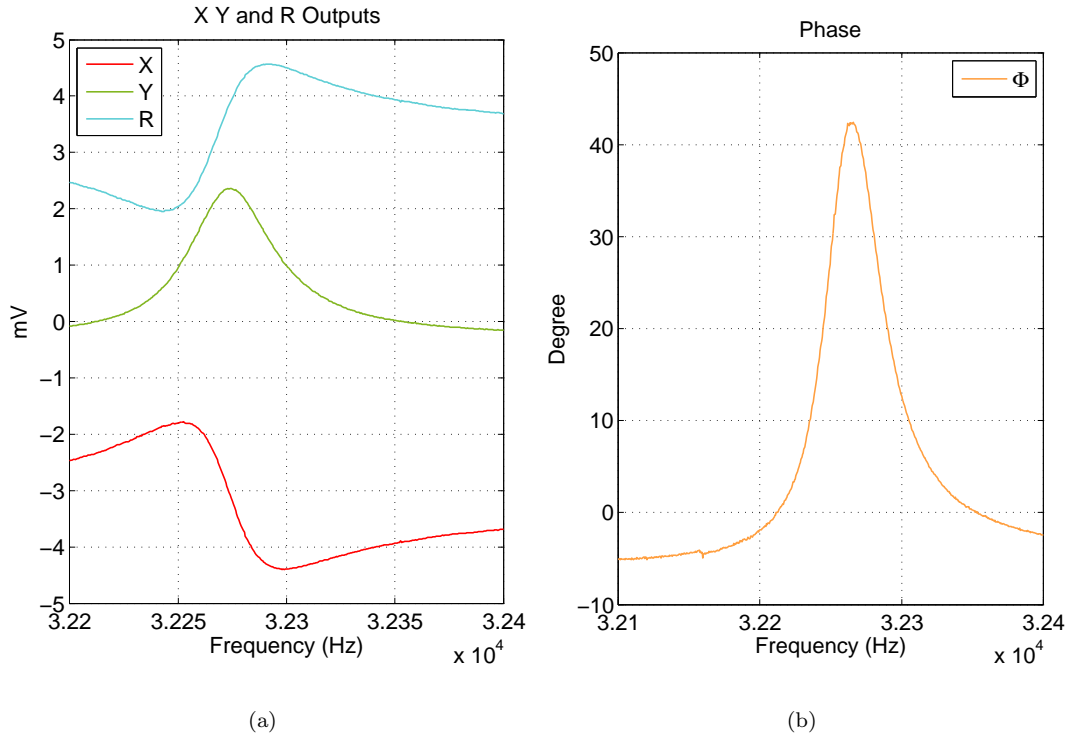


FIGURE 3.5: Frequency response of the fibre tip/tuning fork system, measured when there was no sample. (a) measured X output ($R \cos \Phi$), Y output ($R \sin \Phi$) and calculated amplitude (R) from the LIA as a function of frequency (b) calculated phase (Φ) as a function of frequency

R and Φ curves are obtained from the calculation utilising Equation 3.1 and Equation 3.2. A commercial tuning fork generally has the resonance frequency at 32.768 kHz. However, when the probe is attached and therefore the overall mass of the oscillating system increases, the resonance frequency drops down. As seen in Figure 3.5, the maximum amplitude is present at 32.274 kHz. The asymmetry of the R curve can be attributed to both parasitic capacitance of electrical contacts and connecting wires and the asymmetric coupling of the oscillations of the prongs since one is attached to the probe and the other is not [8].

From Figure 3.5(a), the Q -factor calculated from the Y curve is 686.89 which can be too small for fragile samples. However in this research the sample is a planar waveguide with a smooth surface and therefore the high sensitivity of the probe height control is not crucial. On the other hand, a high Q -factor can retard the response of the feedback system [9] since the settling time of the oscillating system is around 5 times its time constant (τ) which is proportional to the Q -factor as [4, 5]

$$\tau = \frac{\sqrt{3}}{\pi \Delta f} = \frac{\sqrt{3}Q}{\pi f_0} \quad (3.3)$$

The previous value of Q -factor gives the settling time around 60 ms. However, all parameters change from one system to another, depending on the position of the probe on the tine of the tuning fork and the amount of glue applied to attach both. Moreover, the resonant frequency and also the Q -factor varied within a day after gluing the probe, due to stiffening of the glue.

3.4 Probe height control

As the probe is brought down to the surface, the interaction of the shear force with the tip changes both the resonant frequency and the Q -factor. Different phenomena contribute to these variations. As suggested by Wei et al., interactions related to frictional force change the Q -factor whereas the resonant frequency is altered by elastic forces [10]. The overall outcome is a decrease in the amplitude of oscillation. The range of shear-force interaction for which it was measured in previous research with the same instrumental setup is found to be around 10 nm - 20 nm. [3, 11, 12]. The interaction range provides only roughly an approximation of the probe-surface distance.

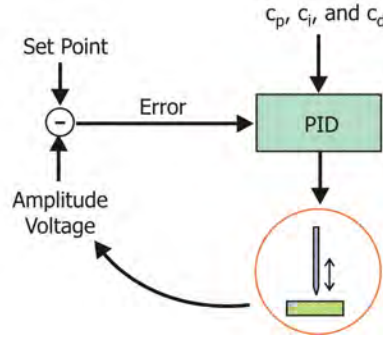


FIGURE 3.6: Schematic diagram of the feedback system

The gap distance of the surface and the probe is maintained by a feedback system, whose schematic diagram is shown in Figure 3.6. As suggested by an experimental study of probe-surface interaction [13], the physical contact of the probe and the surface occurs at the beginning of the damped oscillation. Therefore, in order to avoid damage to the tip, the setpoint of the feedback system should be at the position more than 95% of the undamped amplitude of the oscillation [3]. The value is set by the Y offset of the LIA. The Y output is directly derived from the phase-sensitive detector of the LIA whereas the R output comes from an extra calculating electronic circuit inside the LIA whose digitisation may introduce more noise in the feedback system [11, 14]. Since the SNOM probe in this research operates in the constant height mode, owing to the plain surface of the sample, the phase is constant and the Y output is directly proportional to the R output and used for the feedback system. The deviation of the detected amplitude voltage from the setpoint value is an error signal which is then relayed to the Proportional-Integral-Derivative (PID) unit. The output signal from the PID unit

is sent back to the z (height) channel of the piezoelectric stage controller to correct the probe height, resulting in the correction of amplitude voltage toward the setpoint.

As its name suggests, the PID unit consists of three functions: proportional, integral and derivative components. The output of each subunit, which has the error signal (e) as its input, can be described by Equation 3.4a-Equation 3.4c, and the overall feedback signal (F) is the summation of all, as shown in Equation 3.4d,

$$P = c_p e \quad (3.4a)$$

$$I = c_p c_i \int e dt \quad (3.4b)$$

$$D = c_p c_d \frac{de}{dt} \quad (3.4c)$$

$$F = P + I + D = c_p e + c_p c_i \int e dt + c_p c_d \frac{de}{dt} \quad (3.4d)$$

where P , I and D are proportional, integral and derivative components, respectively. The constants c_p , c_i and c_d are adjusted to optimise the feedback system.

The proportional element corrects the steady-state error. To get back to the set point as promptly as possible, the gain should be high but the rapid response impairs the stability of the system. If only a proportional component contributes to the feedback system, an offset, a deviation from the setpoint, will persist. This can be minimised by the integral unit whose output signal compensates the accumulation of errors during the period of time. Additionally the unit, a lowpass system, is highly responsive to slowly varying noise such as thermal expansion of the stage and probe mount. Therefore it is effectively diminished from the feedback system. The derivative unit has the controlling signal proportional to the rate of error change, so it decreases the time response of the system. This unit should be used only when the scan speed is very low and high-frequency disturbance exists.

The PID unit utilised in the research is a SIM960 Analog PID Controller (Stanford Research Systems). As suggested by the instrument's manual [15], the tuning is manually performed by adjusting the proportional gain first with the integral unit still not operating. The gain should start at a small value and then be increased by factors of two until the oscillation of the system is observed, that is, the fluctuation of the Y -output of

the LIA around the Y offset. After that, the gain should be reduced by a small amount to stabilise the system. Then it should be followed by setting the integral gain, which can be performed in the same way as has been already done for the proportional gain.

3.5 Scanning system and data acquisition

The translation of the probe in the x and y directions is performed by a Melles Griot 3-axis piezoelectric stage with additional stepper motor drivers for coarse translation. The piezoelectric translation range is a maximum of 20 μm with 5 nm resolution and the maximum motor-driven translation range is 4 mm with 25 nm resolution [16].

LabVIEW programmes have been written for scanning control and topographical data acquisition. The programmes communicate with all the instruments via a GPIB interface. The topographical data is gained from a Keithley-196 digital multimeter which reads the voltage from the z channel of the stage controller (corresponding to the height of the probe) because the direct communication with the Melles Griot stage controller via its GPIB port is relatively slow. For the same reason, scanning is controlled indirectly via a Stanford SR400 photon counter's auxiliary voltage outputs, displayed as Port1 (p1) and Port2 (p2) of the voltage supply unit in Figure 3.1. The auxiliary voltage outputs of the photon counter, which supplies ± 10 V, are connected to the x and y external control channels of the stage controller. The scanning of the stage is executed by stepping the output voltage of the photon counter. At each location on the scanning plane, the coordinates of the location, height and spectral data are read and recorded into text files which are later processed by MATLAB scripts.

The spectra collected by the SNOM probe are measured by a high resolution spectrometer, HR2000+ Ocean Optics, which responds to the wavelength range 200-1100 nm with the resolution of 0.035 nm (FWHM) [17]. Unlike other instruments, the spectrometer is interfaced to a computer via a USB 2.0 port and spectra are separately recorded as data files by the software provided with the instrument.

Successive scans in the same area for topographical information of the sample, can show the data repeatability of the system. In Figure 3.7, three consecutive line scans show inclined plane-removed topography on the same region of a grating. These sets of data were acquired without vibration isolation and during the laboratory fabrication with workmen walking in and out regularly. Later on the SNOM system was moved into a new laboratory and placed on an anti-vibration optical table. Therefore, in order to avoid the crashing of the tip into the sample's surface, the scanning was performed at low speed by setting the rest time between each position at 1 second. The speed of scanning is 0.2 $\mu\text{m/s}$. The overall distance of each line scanning in Figure 3.7 is 15 μm with 128 positions of data collecting. As a result, the total time for each scan is 3.36 minutes. The second and third scans were performed after the first scan by 6.05 minutes

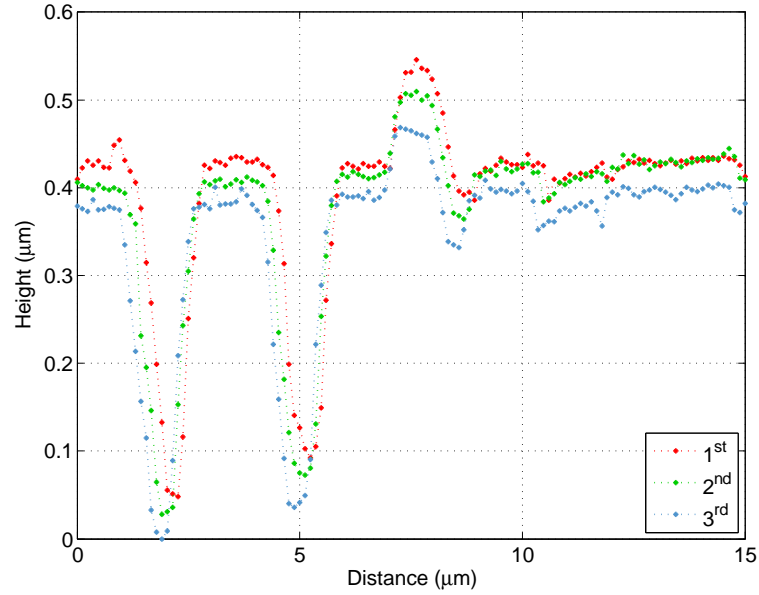


FIGURE 3.7: Three consecutive line scans on the same area of a grating shows the data repeatability of the system. There was no vibration isolation while the data was being acquired and the time for all scans is around 12 minutes. Note that the data has been processed by removing an inclined plane inherent in the data, caused by the sample stage not being entirely flat.

and 17.56 minutes, respectively. As can be observed from Figure 3.7, the height drift from the first scan to the third scan is in the range of around 40 nm. Nevertheless, the observable repeatability of small features on the surface guarantees the reliability of the feedback system.

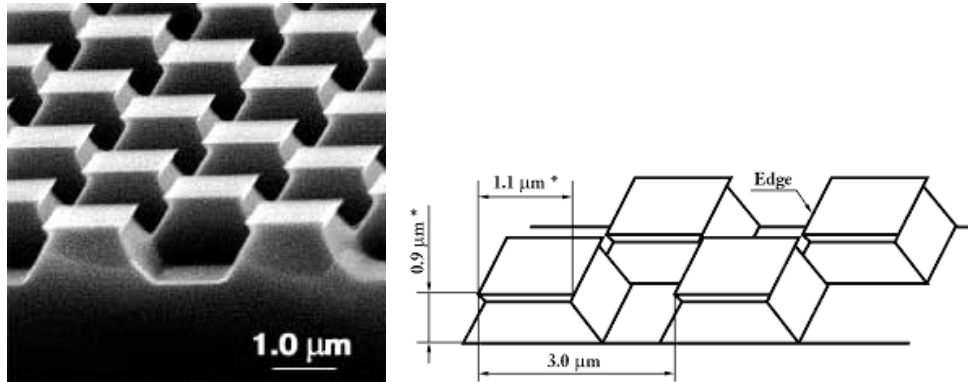


FIGURE 3.8: SEM image and schematic diagram of TGX01 test sample [18]. The grating is the periodic pattern of trapezoidal ridges with the period $3\ \mu\text{m}$. Each ridge is $0.9\ \mu\text{m}$ high and $1.1\ \mu\text{m}$ wide (the top side).

In order to characterise the topographical data acquisition capability of the SNOM system, a calibrated test sample was scanned by an uncoated tapered fibre probe. The test structure is a TGX01 silica grating used for lateral calibration in SPM [18]. The SEM image and dimensions of this grating are shown in Figure 3.8. The grating is a

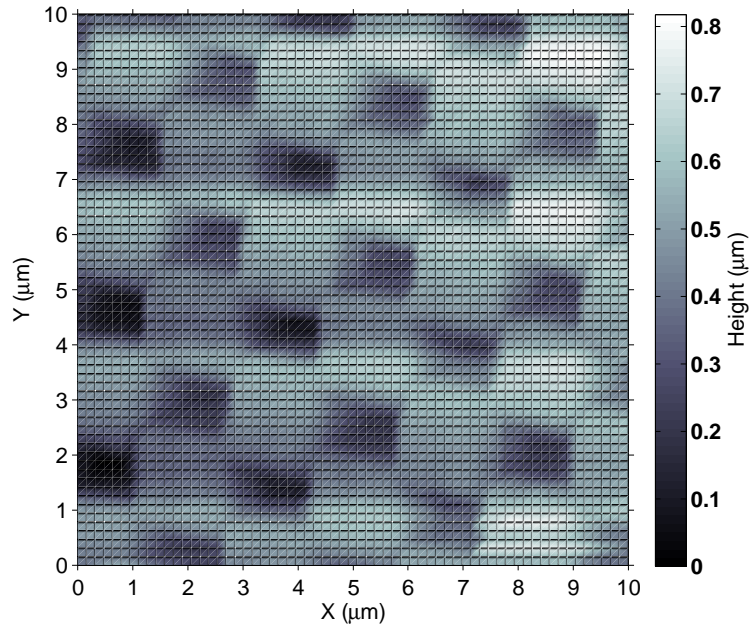


FIGURE 3.9: A false-color topographical image of TGX01. The height data from scanning is plane removed by MATLAB scripts. The pitch of pattern in the image ranges 3.15-3.21 μm whereas the height is around 0.8 μm .

periodic pattern of trapezoidal ridges with a period of 3 μm . Each ridge is 0.9 μm high and 1.1 μm wide (the top side).

The probe height data was collected at different locations on the sample plane by the shear-force feedback system. An oblique plane in the data was removed to increase image contrast and the result is shown in Figure 3.9. The scan direction is along the x -axis, with a scan step of 0.078 μm . The pitch of pattern in the image ranges 3.15-3.21 μm whereas the height is around 0.8 μm . The ridges in the image look oversized. Since the topographical image is the result of a convolution between the tip and surface profiles, the false dimension is probably because the tip has been damaged. Moreover, the scan was performed with the lack of vibration isolation. Disturbances such as air currents from the air conditioner, or people walking by, can dislocate the scanning probe and also cause dimensional discrepancies in the topographical image. The image consequently becomes obscure. Accordingly, the inclusion of vibration isolation will improve spatial accuracy.

3.6 Vibration and thermal effects

The gap distance between the probe and the sample is in the range of nanometres. Hence any disturbance such as mechanical vibration or thermal expansion can induce a great deal of changes in the detected signals. The system should be placed on an anti-vibration table or vibration damping materials such as rubber or foam. Moreover air

currents can also make the probe drift. Therefore the system is covered by a Styrofoam box during scanning.

Material such as brass or aluminium, which is widely used as mounts or bases of instruments, has a thermal expansion too high for the SNOM system. During the placement of the sample, the assistance of a microscope is necessary, but its light can generate heat in the system. This causes the dimensional expansion and later the contraction when the light is turned off. Compared to aluminium with the expansion coefficient $24 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ [19], Invar has an expansion coefficient less than $0.07 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ [20] which can increase the stability of the system more than 20 times. Therefore, the SNOM cantilever has been constructed with Invar.

3.7 Chapter conclusion

The experimental setup for the SNOM system has been described in this chapter. The near-field probe fabrication technique utilised in this research is the heating and pulling technique. The fabrication is performed by a commercial micropipette puller which melts a fibre with a CO₂ laser and then applies a tensile force to break the fibre. As a result, the tapered shape is formed. The probe is utilised as uncoated to avoid introducing some deviation in the spectral and temporal profile of the collected light.

The height control is achieved by the shear-force technique, in which the deviation of the probe oscillation caused by the interaction of shear force is monitored. The probe is attached to a tuning fork which is mechanically driven by a piezoelectric disc with the oscillating signal from a function generator. The dithering of the probe is transformed to an electrical voltage and measured by a lock-in amplifier. A feedback system consisting of a PID unit controls the probe height by monitoring the change in amplitude voltage from the tuning fork since the strength of the shear force depends on the distance between the probe height and the surface of the sample. The feedback system can then provide an appropriate voltage to adjust the height channel of the piezoelectric stage according to the magnitude of the error signal.

The computer-controlled lateral scanning of the probe is performed by providing a step voltage to the stage. At various locations of the sample, topographical and spectral information are collected. Finally vibration and thermal effects were also discussed in the chapter.

3.8 References

- [1] Sutter Instrument Company. Operation manual: P-2000 micropipette puller.
- [2] Stelea, S. Re: P-2000 micropipette puller. Private communication, 25 July 2008.

- [3] C. W. J. Hillman. *Scanning near-field optical microscope characterisation of microstructured optical fibre devices*. PhD thesis, University of Southampton, 2002.
- [4] K. Karrai and R. D. Grober. Piezoelectric tip-sample distance control for near field optical microscopes. *Applied Physics Letters*, 66(14):1842–1844, 1995.
- [5] K. Karrai and R. D. Grober. Piezo-electric tuning fork tip-sample distance control for near field optical microscopes. *Ultramicroscopy*, 61(1-4):197–205, 1995.
- [6] J. U. Schmidt, H. Bergander, and L. M. Eng. Shear force interaction in the viscous damping regime studied at 100 pN force resolution. *Journal of Applied Physics*, 87(6):3108–3112, 2000.
- [7] J. U. Schmidt, H. Bergander, and L. M. Eng. Experimental and theoretical analysis of shear-force interaction in the non-contact regime with 100 pN force resolution. *Applied Surface Science*, 157(4):295–301, 2000.
- [8] Y. Qin and R. Reifenberger. Calibrating a tuning fork for use as a scanning probe microscope force sensor. *Review of Scientific Instruments*, 78(6):63704/1–63704/7, 2007.
- [9] R. C. Dunn. Near-field scanning optical microscopy. *Chemical Reviews*, 99:2891–2927, 1999.
- [10] P. K. Wei and W. S. Fann. The effect of humidity on probe-sample interactions in near-field scanning optical microscopy. *Journal of Applied Physics*, 87(5):2561–2564, 2000.
- [11] J. C. Gates. *Measurement of the local optical phase and amplitude in photonic devices using scanning near-field microscopy*. PhD thesis, University of Southampton, 2003.
- [12] J. D. Mills. *An investigation of phase-mask diffraction patterns and fibre Bragg gratings with scanning near-field optical microscopy*. PhD thesis, University of Southampton, 2001.
- [13] I. I. Smolyaninov, W. A. Atia, S. Pilevar, and C. C. Davis. Experimental study of probe-surface interaction in near-field optical microscopy. *Ultramicroscopy*, 71(1-4):177–182, 1998.
- [14] Stanford Research Systems. Operation manual: Model SR530 lock-in amplifier, 1993.
- [15] Stanford Research Systems. Operation and service manual: Analog PID controller SIM960.
- [16] Melles Griot Ltd. Nanomax-HS 3 axis flexure stage: Instruction handbook for the Melles Griot nanopositioning modular system, 1998.

- [17] Ocean Optics Inc. HR2000+ High-speed fiber optic spectrometer: Installation and operation manual.
- [18] MikroMasch. Gratings for lateral calibration of SPM. Device datasheet retrieved from <http://www.spmtips.com/tgx/>.
- [19] D. R. Lide, editor. *CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data*. CRC Press, Boca Raton, FL, USA, 87th edition, 2006.
- [20] W. F. Schlossera, G. M. Grahama, and P. P. M. Meinckea. The temperature and magnetic field dependence of the forced magnetostriction and thermal expansion of Invar. *Journal of Physics and Chemistry of Solids*, 32(5):927–938, 1971.

Chapter 4

Waveguide characteristics

4.1 Chapter introduction

In this chapter, the physical background relating to the characteristics of the waveguides will be studied. The propagation modes of light guided by total internal reflection will be analysed. The study is focused on only the case of Cartesian symmetry owing to the rectangular shape of the waveguide utilised in this research. It begins with the simple structure of the planar waveguide, with a one-dimensional index profile, in Section 4.2. The propagation constants and modal field distribution will be derived from the Maxwell's equations and the boundary conditions. This one-dimensional mode analysis will be later adapted for the rectangular waveguides in Section 4.3. Such a waveguide consists of a two-dimensional refractive index profile where the problem of solving the propagation modes becomes more complicated, and normally non-analytical solutions are preferred. Among many approaches, the effective index method adopted in this research to determine the propagation modes is probably the simplest technique and is very accurate. The approach is based on the assumption that the spatial dependency of the field in two orthogonal directions is separable. As a result, the problem of the two-dimensional waveguide is converted into the two planar waveguides with their refractive index profiles perpendicular to each other. Then the propagation modes are determined from one of the planar waveguides containing the refractive index profile specified from the propagation modes of the other waveguide. The accuracy can be enhanced further by optimising the equivalent refractive index profile of the effective index method, to nullify the first-order error according to the perturbation theory.

The last section, Section 4.4, will provide the underlying theoretical background of nonlinear pulse propagation in a medium. The evolution of pulses induced by dispersion and nonlinear effects is expressed in the form generally acknowledged as the nonlinear Schrödinger equation. The derivation of the equation starts with the general form of wave equation given by the Maxwell's equations. The induced polarisation, taken as the

response of a nonlinear system with the electric field of the launched light as the input, can be approximated by the Volterra series with the Volterra kernels related to the electric susceptibility of the medium. The derived propagation equation will be further utilised in later chapters for the numerical modeling of the research.

4.2 Planar waveguides

We start by considering the light confinement in a planar waveguide with a one-dimensional distribution of the refractive index. Even though real rectangular waveguides have light confined along two dimensions, which are perpendicular to the propagation direction, their characteristics are based on those of basic one-dimensional waveguides. Moreover, the simple structure of planar waveguides can be utilised as an approximation for waveguides whose length in one dimension is much longer than the other, so that the light confinement in the longer dimension can be neglected.

4.2.1 Field distribution of planar waveguides

The basic structure of planar waveguides is shown in Figure 4.1. The plane of incidence is in the yz plane, whereas the x -axis aligns perpendicular to the plane. The core is $2d$ wide and the light, which is confined within the core area, propagates along the z -direction with the angle of incidence θ . The step-index profile of the waveguide is along the y direction and its dielectric constant profile can be described by a function $\eta^2(y)$ as

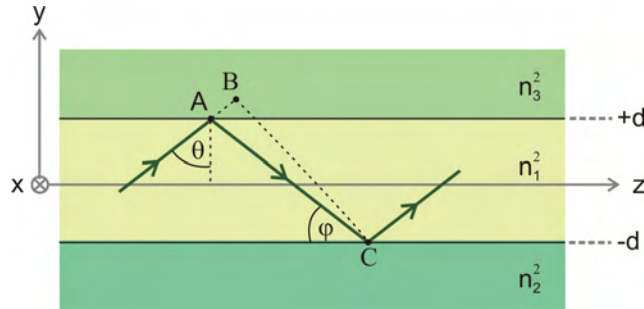


FIGURE 4.1: Total internal reflection in a one-dimensional waveguide

$$\eta^2(y) = \begin{cases} n_1^2, & |y| \leq d \\ n_2^2, & y < -d \\ n_3^2, & y > d \end{cases} \quad (4.1)$$

where $n_1^2 > n_2^2, n_3^2$ and the materials are assumed to be homogeneous. This structure of

the waveguide can represent any arbitrary planar waveguides which are not necessarily symmetric, with the same material for both upper and lower cladding layers. Certainly the analysis of this waveguide can also be applied to the symmetric case by substituting $n_2 = n_3$. For the waveguide utilised in this research, along the height dimension, light is confined by an asymmetric waveguide, whereas by a symmetric waveguide along the width dimension.

Since the dielectric profile varies only along the y -direction, the electric field \mathcal{E} and magnetic field \mathcal{H} for the waveguide in Figure 4.1, are assumed to be independent from the x -direction and can be written as

$$\mathcal{E}(y, z, t) = \mathbf{E}(y) \exp[i(\omega t - \beta z)] \quad (4.2a)$$

$$\mathcal{H}(y, z, t) = \mathbf{H}(y) \exp[i(\omega t - \beta z)] \quad (4.2b)$$

with

$$\mathbf{E}(y) = E_x(y) \hat{\mathbf{x}} + E_y(y) \hat{\mathbf{y}} + E_z(y) \hat{\mathbf{z}} \quad (4.3a)$$

$$\mathbf{H}(y) = H_x(y) \hat{\mathbf{x}} + H_y(y) \hat{\mathbf{y}} + H_z(y) \hat{\mathbf{z}} \quad (4.3b)$$

where $\mathbf{E}(y)$ and $\mathbf{H}(y)$ are the field amplitudes with $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ as the unit vectors, ω is the angular frequency and β is the propagation constant of the waveguide medium which is the z -component of the wave vector $\mathbf{k} = k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}$, i. e.

$$\beta = k_z = k \sin \theta \quad (4.4)$$

There are two types of modes according to the polarisation of the electric and magnetic fields. For the transverse electric (TE) modes, the electric field is perpendicular to the plane of incidence and the magnetic field aligns in the plane. Therefore the nonzero components of the electric and magnetic fields for the TE modes are E_x , H_y and H_z . On the other hand, the transverse magnetic (TM) modes with the magnetic field perpendicular to the plane of incidence have the nonzero components H_x , E_y and E_z .

The fields \mathcal{E} and \mathcal{H} in Equation 4.2 can be determined from the curl Maxwell equations for a charge-free isotropic medium which are

$$\nabla \times \mathcal{E} = -\mu \frac{\partial \mathcal{H}}{\partial t} \quad (4.5a)$$

$$\nabla \times \mathcal{H} = \varepsilon \frac{\partial \mathcal{E}}{\partial t} \quad (4.5b)$$

where ε is the electric permittivity and μ is the magnetic permeability of the waveguide medium. Substituting the fields of Equation 4.2 into Equation 4.5 results in

TE mode:

$$\beta E_x = -\mu\omega H_y \quad (4.6a)$$

$$\frac{\partial E_x}{\partial y} = -i\mu\omega H_z \quad (4.6b)$$

$$-i\beta H_y - \frac{\partial H_z}{\partial y} = i\varepsilon\omega E_x \quad (4.6c)$$

TM mode:

$$\beta H_x = \varepsilon\omega E_y \quad (4.6d)$$

$$\frac{\partial H_x}{\partial y} = i\varepsilon\omega E_z \quad (4.6e)$$

$$i\beta E_y + \frac{\partial E_z}{\partial y} = i\mu\omega H_x \quad (4.6f)$$

By applying the partial derivative $\partial/\partial t$ to Equation 4.6b with the assistance of Equation 4.6a, Equation 4.6c and the relation $\mu\varepsilon = 1/v^2 = \eta^2(y)/c^2$ where v and c is the speed of light in medium and free space respectively, the wave equation of the transverse electric field component E_x for the TE modes is derived.

TE mode:

$$\frac{\partial^2 E_x}{\partial y^2} + (\eta^2(y)k_0^2 - \beta^2) E_x = 0 \quad (4.7a)$$

With a similar procedure, but applied to Equation 4.6e, the wave equation of H_x for the TM modes can be written as

TM mode:

$$\frac{\partial^2 H_x}{\partial y^2} + (\eta^2(y)k_0^2 - \beta^2) H_x = 0 \quad (4.7b)$$

where k_0 is the magnitude of the wave vector in free space. Furthermore, the components of the fields in the plane of incidence, H_y and H_z for TE modes and E_y and E_z for TM modes, can be determined by substituting the solutions of Equation 4.7 into Equation 4.6a, Equation 4.6b, Equation 4.6d, and Equation 4.6e respectively. After

arithmetical rearrangement, H_y , H_z , E_y and E_z can be expressed as

TE mode:

$$H_y = \frac{-\beta}{\mu\omega} E_x \quad (4.8a)$$

$$H_z = \frac{i}{\mu\omega} \frac{\partial E_x}{\partial y} \quad (4.8b)$$

TM mode:

$$E_y = \frac{\beta}{\varepsilon\omega} H_x \quad (4.8c)$$

$$E_z = -\frac{i}{\varepsilon\omega} \frac{\partial H_x}{\partial y} \quad (4.8d)$$

Since the guided field in the waveguide is confined within the core ($|y| < d$) and decays rapidly within the cladding region ($|y| > d$), the solution of Equation 4.7 can be written in the form [1]

TE mode:

$$E_x = \begin{cases} A_1 \cos(k_y y - \vartheta), & |y| < d \\ A_2 e^{\gamma_2(y+d)}, & y < -d \\ A_3 e^{-\gamma_3(y-d)}, & y > d \end{cases} \quad (4.9a)$$

TM mode:

$$H_x = \begin{cases} A_1 \cos(k_y y - \vartheta), & |y| < d \\ A_2 e^{\gamma_2(y+d)}, & y < -d \\ A_3 e^{-\gamma_3(y-d)}, & y > d \end{cases} \quad (4.9b)$$

where $A_{1,2,3}$ are arbitrary constants whose relations can be determined from the boundary conditions, k_y is the propagation constant for the field propagation in the core area, $\gamma_{2,3}$ are the extinction coefficients in the lower and upper cladding layers respectively and ϑ is the arbitrary phase shift.

The fields are sinusoidal inside the core with the feature specified by the propagation constant k_y which is the y -component of the wave vector \mathbf{k}

$$k_y^2 = k^2 \cos^2 \theta = n_1^2 k_0^2 - \beta^2 \quad (4.10)$$

On the other hand, the fields becomes evanescent in the cladding regions and attenuated

exponentially in the lower and upper cladding regions with the extinction coefficients γ_2 and γ_3 respectively. The values of γ_2 and γ_3 can be calculated from

$$\gamma_2^2 = \beta^2 - n_2^2 k_0^2 \quad (4.11)$$

$$\gamma_3^2 = \beta^2 - n_3^2 k_0^2 \quad (4.12)$$

From Equation 4.10 - Equation 4.12, the parameters k_y , γ_2 and γ_3 are related by

$$(k_y d)^2 + (\gamma_2 d)^2 = V_2^2 \quad (4.13a)$$

$$(k_y d)^2 + (\gamma_3 d)^2 = V_3^2 \quad (4.13b)$$

with V_2 and V_3 defined as the normalised frequency corresponding to the quality of light confinement within the core and are also related to the numerical aperture (NA) by

$$V_{2,3} = k_0 d (n_1^2 - n_{2,3}^2)^{1/2} = k_0 d NA \quad (4.14)$$

where

$$NA = (n_1^2 - n_{2,3}^2)^{1/2} \quad (4.15)$$

Since all the mode parameters k_y , γ_2 and γ_3 must be real, their upper limit value is bound by $V_{2,3}/d$. Normally a non-symmetric waveguide has the refractive index of the substrate higher than that of the cover layer, that is, $n_2 > n_3$ and consequently $V_2 < V_3$. This physically means that the critical angle for the lower interface is greater than that of the upper interfaces and therefore light must propagate with the angle of incidence higher than $\theta_{c2} = \sin^{-1}(n_2/n_1)$ for the condition of total internal reflection to be fulfilled at both interfaces. Equivalently k_y is required to be less than V_2/d .

The constants $A_{1,2,3}$ and ϑ in Equation 4.9 are determined by the boundary conditions. At the interface $y = \pm d$, the tangential components to the interfaces, E_x for TE modes and H_x For TM modes, are continuous as long as $A_{1,2,3}$ satisfy

$$A_2 = A_1 \cos(k_y d + \vartheta) \quad (4.16a)$$

$$A_3 = A_1 \cos(k_y d - \vartheta) \quad (4.16b)$$

By substituting Equation 4.16 into Equation 4.9 and taking the constant A_1 as unity, E_x and H_x can be written as

TE mode:

$$E_x = \begin{cases} \cos(k_y y - \vartheta), & |y| < d \\ \cos(k_y d + \vartheta)e^{\gamma_2(y+d)}, & y < -d \\ \cos(k_y d - \vartheta)e^{-\gamma_3(y-d)}, & y > d \end{cases} \quad (4.17a)$$

TM mode:

$$H_x = \begin{cases} \cos(k_y y - \vartheta), & |y| < d \\ \cos(k_y d + \vartheta)e^{\gamma_2(y+d)}, & y < -d \\ \cos(k_y d - \vartheta)e^{-\gamma_3(y-d)}, & y > d \end{cases} \quad (4.17b)$$

On the other hand, the phase constant ϑ can be determined by the condition of continuity for H_z and E_z . By substituting Equation 4.17a and Equation 4.17b into Equation 4.8b and Equation 4.8d respectively, H_z and E_z are continuous at $y = \pm d$ as long as

TE mode:

$$k_y \sin(k_y d + \vartheta) = \gamma_2 \cos(k_y d + \vartheta) \quad (4.18a)$$

$$k_y \sin(k_y d - \vartheta) = \gamma_3 \cos(k_y d - \vartheta) \quad (4.18b)$$

TM mode:

$$n_{21}^2 k_y \sin(k_y d + \vartheta) = \gamma_2 \cos(k_y d + \vartheta) \quad (4.18c)$$

$$n_{31}^2 k_y \sin(k_y d - \vartheta) = \gamma_3 \cos(k_y d - \vartheta) \quad (4.18d)$$

where $n_{21,31}$ are defined as

$$n_{21} = \frac{n_2}{n_1} \quad \text{and} \quad n_{31} = \frac{n_3}{n_1} \quad (4.19)$$

Note that Equation 4.18a and Equation 4.18b are derived on the assumption that the

waveguide's material is nonmagnetic and therefore $\mu_{2,3} \approx \mu_0$ where μ_0 is the permeability of free space. Two sets of coupled equations in Equation 4.18 are solved for ϑ , and finally the phase constant for each mode with mode index n can be determined from [1]

TE mode:

$$2\vartheta = n\pi + \tan^{-1}\left(\frac{\gamma_2}{k_y}\right) - \tan^{-1}\left(\frac{\gamma_3}{k_y}\right) \quad (4.20a)$$

TM mode:

$$2\vartheta = n\pi + \tan^{-1}\left(\frac{\gamma_2}{n_{21}^2 k_y}\right) - \tan^{-1}\left(\frac{\gamma_3}{n_{31}^2 k_y}\right) \quad (4.20b)$$

where n is an integer $n = 0, 1, 2, \dots$. Equation 4.17 accompanied by Equation 4.8 and Equation 4.20, provides the field distribution inside the waveguide which will be later utilised in the calculation for the transverse field profile in Chapter 5.

4.2.2 Propagation modes of planar waveguides

To be guided by the waveguide, the light inside needs to satisfy total internal reflection at the interfaces, as shown in Figure 4.1. The incidence angle θ of the guided modes are required to be more than the critical angle θ_c . There are various modes which satisfy such a condition and propagate with different optical paths. The ray optics of light propagating within the waveguide is adequate enough for solving the propagation constant β of these guided modes. From the ray path in Figure 4.1, after one round trip for the light traveling between the upper and lower cladding layers, the twice reflected wave experiences a phase shift caused by the optical path difference (*OPD*), and the reflection phase shift ϕ_{r2} and ϕ_{r3} at the lower and upper interfaces respectively. For guided modes, the overall phase shift must satisfy

$$k_0(OPD) - \phi_{r2} - \phi_{r3} = 2\pi n \quad (4.21)$$

where n is the mode index number and the optical path difference is determined from

$$\begin{aligned} OPD &= n_1(\overline{AC} - \overline{AB}) \\ &= n_1\left(\frac{2d}{\sin \varphi} - \frac{2d}{\sin \varphi} \cos 2\varphi\right) \\ &= 4n_1 d \sin \varphi \end{aligned} \quad (4.22)$$

The reflection phase shifts ϕ_{r2} and ϕ_{r3} , which depend on the angle of incidence and polarisation, can be determined from the Fresnel equations. The argument of the complex

reflection coefficient provides the phase shift for the reflected wave and can be written as [2]

TE mode:

$$\tan\left(\frac{\phi_{r2}}{2}\right) = \frac{(\sin^2 \theta - \sin^2 \theta_{c2})^{1/2}}{\cos \theta} \quad (4.23a)$$

$$\tan\left(\frac{\phi_{r3}}{2}\right) = \frac{(\sin^2 \theta - \sin^2 \theta_{c3})^{1/2}}{\cos \theta} \quad (4.23b)$$

TM mode:

$$\tan\left(\frac{\phi_{r2}}{2}\right) = \frac{(\sin^2 \theta - \sin^2 \theta_{c2})^{1/2}}{n_{21}^2 \cos \theta} \quad (4.23c)$$

$$\tan\left(\frac{\phi_{r3}}{2}\right) = \frac{(\sin^2 \theta - \sin^2 \theta_{c3})^{1/2}}{n_{31}^2 \cos \theta} \quad (4.23d)$$

where θ_{c2} and θ_{c3} are the critical angles at the bottom and top interfaces of the waveguide. i.e.

$$\sin \theta_{c2} = \frac{n_2}{n_1} \quad \text{and} \quad \sin \theta_{c3} = \frac{n_3}{n_1} \quad (4.24)$$

The phase shift ϕ_{r2} and ϕ_{r3} can be expressed in terms of modal parameters k_y , γ_2 , and γ_3 by applying Equation 4.4, Equation 4.10, Equation 4.11, Equation 4.12 and Equation 4.24 to Equation 4.23. As a result, the simplified forms are derived.

TE mode:

$$\tan\left(\frac{\phi_{r2}}{2}\right) = \frac{\gamma_2}{k_y} \quad (4.25a)$$

$$\tan\left(\frac{\phi_{r3}}{2}\right) = \frac{\gamma_3}{k_y} \quad (4.25b)$$

TM mode:

$$\tan\left(\frac{\phi_{r2}}{2}\right) = \frac{\gamma_2}{n_{21}^2 k_y} \quad (4.25c)$$

$$\tan\left(\frac{\phi_{r3}}{2}\right) = \frac{\gamma_3}{n_{31}^2 k_y} \quad (4.25d)$$

By substituting Equation 4.22 and Equation 4.25 for the *OPD*, and the phase shift ϕ_{r2} and ϕ_{r3} into Equation 4.21, we get

TE mode:

$$2k_y d - \tan^{-1}\left(\frac{\gamma_2}{k_y}\right) - \tan^{-1}\left(\frac{\gamma_3}{k_y}\right) = n\pi \quad (4.26a)$$

TM mode:

$$2k_y d - \tan^{-1} \left(\frac{\gamma_2}{n_{21}^2 k_y} \right) - \tan^{-1} \left(\frac{\gamma_3}{n_{31}^2 k_y} \right) = n\pi \quad (4.26b)$$

Finally, with the assistance of the trigonometry identities,

$$\tan(\theta_1 + \theta_2) = \frac{\tan \theta_1 + \tan \theta_2}{1 - \tan \theta_1 \tan \theta_2} \quad (4.27)$$

and

$$\tan(\theta_1 - n\pi) = \tan \theta_1 \quad (4.28)$$

Equation 4.26 becomes

TE mode:

$$\tan(2k_y d) = \frac{k_y (\gamma_2 + \gamma_3)}{k_y^2 - \gamma_2 \gamma_3} \quad (4.29a)$$

TM mode:

$$\tan(2k_y d) = \frac{k_y ((\gamma_2/n_{21}^2) + (\gamma_3/n_{31}^2))}{k_y^2 - (\gamma_2 \gamma_3 / n_{21}^2 n_{31}^2)} \quad (4.29b)$$

Equation 4.29 is generally referred to as the characteristic equation of the waveguide which specifies the propagation modes with the propagation parameters k_y , γ_2 and γ_3 and consequently β , through the relation Equation 4.10 - Equation 4.11. The equation is transcendental and its roots can be found by graphical means. The left and right terms of Equation 4.29 are plotted against $k_y d$ whose value varies from 0 to $\min[V_2, V_3]/d$. The intercepts between two curves from both sides of the equation specifies the mode propagation constants and the extinction coefficients in the cladding regions. The approach can be viewed in Section 5.3, Chapter 5. The tangential function on the left side of Equation 4.29 is periodic with the period $k_y d = \pi/2$. On the other hand, the term on the right is a decreasing function and its curve intercepts with the tangent curve at every period of the latter. Therefore the number of modes N can be determined from

$$N = \lceil \frac{V_{2,3}}{\pi/2} \rceil = \lceil \frac{4d}{\lambda_0} NA \rceil \quad (4.30)$$

where $\lceil \rceil$ denotes the ceiling function whose value is an integer not less than the argument and λ_0 is the wavelength in free space.

For a symmetrical waveguide in which $n_2 = n_3$, the phase condition in Equation 4.21 becomes

$$k_y d - n \frac{\pi}{2} = \frac{\phi_{r2}}{2} \quad (4.31)$$

By applying the tangent function to both sides of Equation 4.31, the characteristic equation is given by

TE mode:

$$\tan \left(k_y d - n \frac{\pi}{2} \right) = \frac{\gamma^2}{k_y} \quad (4.32a)$$

TM mode:

$$\tan \left(k_y d - n \frac{\pi}{2} \right) = \frac{\gamma^2}{n_{21}^2 k_y} \quad (4.32b)$$

For arbitrary dimensions, the structure of a planar waveguide supports at least one mode. However, in order to remain single mode, Equation 4.30 suggests that the width of the waveguide must be less than $\lambda_0/(2NA)$. According to Miller's rule, high-order susceptibility is proportional to the linear susceptibility and consequently a medium with high optical nonlinearity tends normally to come with high refractive index [3–5]. Therefore, the width dimension for operation in the single-mode regime becomes smaller and probably more tedious to fabricate than a linear waveguide. For example, a guiding medium of highly nonlinear Ta_2O_5 thin film on a silica substrate gives the condition of the waveguide's dimension less than $\sim \lambda_0/3$ which is less than half a micron for the operation wavelength of 800 nm. On the other hand, the upper limit rises up to $\sim \lambda_0/0.9$ for a widely utilised medium in planar waveguide applications such as $\text{SiO}_2/\text{GeO}_2$ on the same kind of substrate [6–8].

4.3 Rectangular waveguides

The scalar wave equations in Equation 4.7 for a planar waveguide is a specific case of the full wave vector equation when there is only one transverse component of the field. For the propagation in a rectangular waveguide with a two-dimensional step index profile and homogeneous medium, the transverse component of the field (\mathbf{E}_t) which contains both the x - and y -components satisfies [9, 10]

$$\nabla_t^2 \mathbf{E}_t + (k_0^2 \eta^2 - \beta^2) \mathbf{E}_t = -\nabla_t \left(\mathbf{E}_t \cdot \frac{\nabla_t \eta^2}{\eta^2} \right) \quad (4.33)$$

where η^2 is the dielectric constant of the waveguide and the transverse del operator ∇_t is defined as

$$\nabla_t = \frac{\partial}{\partial x} \hat{\mathbf{x}} + \frac{\partial}{\partial y} \hat{\mathbf{y}} \quad (4.34)$$

The term $\nabla_t \eta^2$ on the right hand side of Equation 4.33 is the coupling term of various

field components and specifies the polarisation characteristics of the modes. For rectangular shapes of the waveguide, the coupling term can be neglected and the equation can be separated into each component's wave equation [9]

$$\nabla_t^2 E_x + (k_0^2 \eta^2 - \beta^2) E_x = -\frac{\partial}{\partial x} \left(\frac{E_x}{\eta^2} \frac{\partial \eta^2}{\partial x} \right) \quad (4.35a)$$

$$\nabla_t^2 E_y + (k_0^2 \eta^2 - \beta^2) E_y = -\frac{\partial}{\partial y} \left(\frac{E_y}{\eta^2} \frac{\partial \eta^2}{\partial y} \right) \quad (4.35b)$$

with E_x and E_y as the x - and y -components of the field. Two polarisation modes are normally classed according to a much greater strength of electric field component aligning along the x - or y -axis whereas the other orthogonal component, (y or x , respectively), is small and can be ignored. In contrast to the case of a planar waveguide in Section 4.2 in which only the x -axis is recognised as the transverse direction, the notation TE and TM becomes the unclear definition of the polarisation directions for the case of a rectangular waveguide. Therefore in order to avoid such an ambiguity, the notations E_{mn}^x and E_{mn}^y are adopted instead, to represent the propagation modes derived from Equation 4.35a and Equation 4.35b respectively with the superscript stating the direction of the strong electric field and the subscript mn specifying the mode indices for the field distribution along the x - and y -directions respectively.

The problem of determining propagation modes for a two-dimensional waveguide is complicated. The approximated solutions for the propagation modes of the waveguide can be derived by both analytical and numerical methods. The discussion of these methods in detail can be found in Chiang's review [9]. For the rectangular geometry, the effective index method (EIM) [11] is efficient and widely utilised owing to its simplicity and reasonable accuracy, as long as the frequency region in the analysis is not near the cut-off frequency [12].

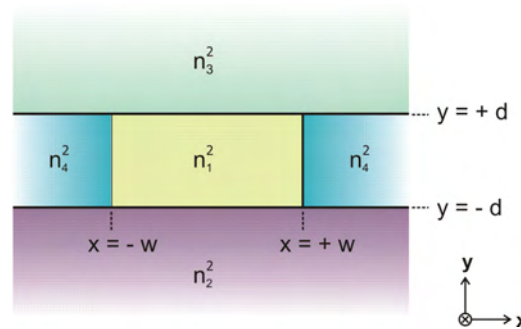


FIGURE 4.2: Basic structure of a rectangular waveguide

The general schematic structure of a rectangular waveguide is displayed in Figure 4.2. The core area is $2d$ high and $2w$ wide with the refractive index n_1 . The length of the waveguide aligns along the z -axis whereas light is confined in two transverse dimensions

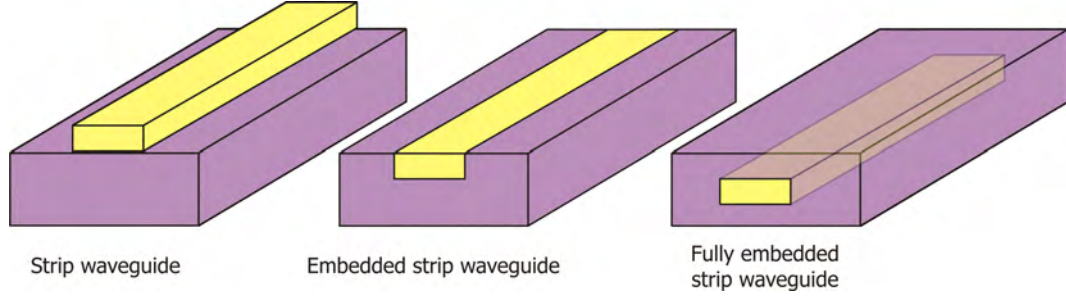


FIGURE 4.3: Various types of rectangular waveguides

x and y by the higher value refractive indices n_2 , n_3 , and n_4 of the cladding area. Such a structure can represent several types of rectangular waveguides such as strip waveguides ($n_3 = n_4 < n_2$), embedded strip waveguides ($n_2 = n_4 > n_3$), or fully embedded strip waveguides ($n_2 = n_3 = n_4$) [2, 12], as portrayed in Figure 4.3.

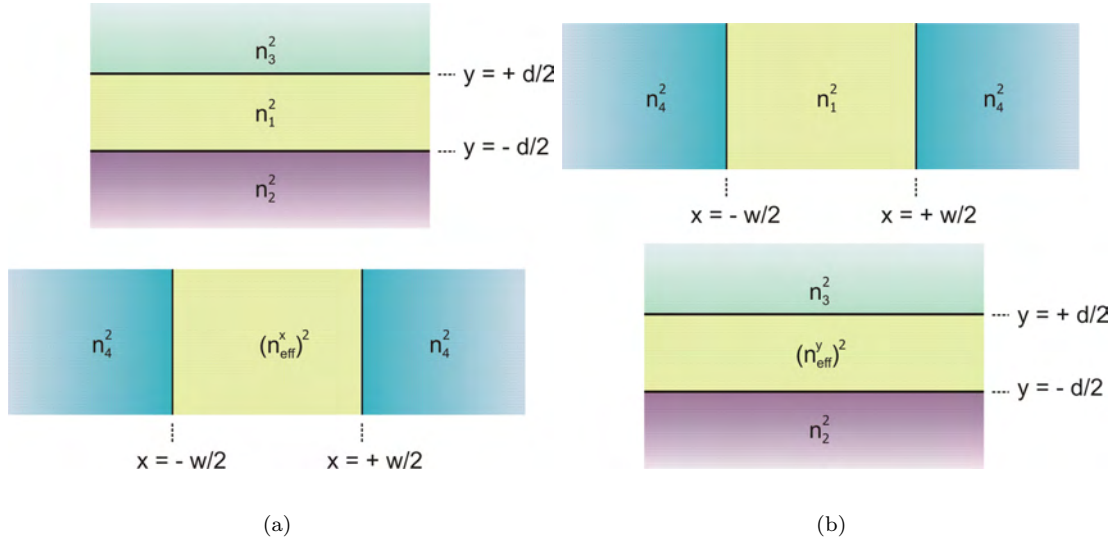


FIGURE 4.4: Two orthogonal planar waveguides for the analysis of a rectangular waveguide by the EIM. There are two choices of utilising the method: (a) determining the propagation modes from the planar waveguide along the x -direction with the effective index of the core given by the propagation constant previously determined from the planar waveguide along the y -direction or (b) vice versa

The EIM is based on the separation of variables which changes a two-dimensional problem of the rectangular waveguide to the problem of two one-dimensional planar waveguides. The analysis of propagation modes in one planar waveguide provides the effective index for the core of the other orthogonal planar waveguide. The procedure of the EIM is portrayed in Figure 4.4 in which there are two choices for applying the method. The propagation modes of the waveguide, are those determined from the planar waveguide with the step index profile along the x -direction, in which the effective index of the core is specified by the propagation constants previously determined from the planar waveguide, aligning parallel to the y -axis as summarised in Figure 4.4(a). On the other hand,

the order of both planar waveguides can be reversed, as in Figure 4.4(b). The effective indices n_{eff}^x or n_{eff}^y of the lower waveguides in Figure 4.4(a) are specified by

$$n_{eff}^x = \frac{\beta^x}{k_0} \quad \text{or} \quad n_{eff}^y = \frac{\beta^y}{k_0} \quad (4.36)$$

where β^x and β^y are the propagation constants determined from the upper waveguides in Figure 4.4(a) and Figure 4.4(b) respectively.

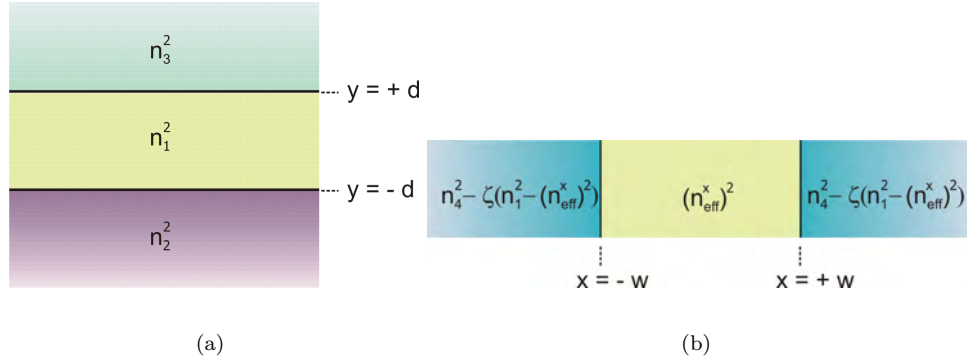


FIGURE 4.5: Schematic diagram of the effective index method with the introduction of the correcting parameter (ζ) in the dielectric constant of the side cladding area. The analysis starts with determining modes in the planar waveguide with the refractive index profile along the y direction in (a), and is then followed by the planar waveguide along the x direction in (b)

However, it has been found that the approach is not commutative. The procedures in Figure 4.4(a) and Figure 4.4(b) do not provide exactly the same solutions and, if the analysis starts with the shorter dimension waveguide, the result is more accurate than the other option [13]. Furthermore, Kumar et al. demonstrated that EIM always gives an overestimated value of the propagation constants. As suggested by Chiang, the accuracy can be enhanced by adding a correcting parameter in the refractive index of cladding area in order to minimise the error between the EIM approximation and the exact value [14], as displayed in Figure 4.5. The dielectric constant of the side cladding regions is different from that of the conventional scheme in Figure 4.4(a) by the amount $\zeta \left(n_1^2 - (n_{eff}^x)^2 \right)$ where ζ is a constant whose value is optimised by a perturbation analysis. In such an approach, the error in the calculation for the propagation constants of the EIM is assumed to have originated from the deviation of the equivalent index profile of the EIM from the waveguide's original index profile.

For a homogeneous medium, $\nabla_t \eta^2 / \eta^2$ becomes zero and the scalar form of Equation 4.35 can be written as

$$\frac{\partial^2 \psi_{mn}(x, y)}{\partial x^2} + \frac{\partial^2 \psi_{mn}(x, y)}{\partial y^2} + (k_0^2 \eta^2(x, y) - \beta_{mn}^2) \psi_{mn}(x, y) = 0 \quad (4.37)$$

where the scalar field $\psi_{mn}(x, y)$, which can represent either E_{mn}^x or E_{mn}^y modes, is a two-dimensional function of x and y and the mode can be identified by two indices m and n corresponding to the x - and y -directions respectively, and $\eta(x, y)$ is the refractive index profile of the waveguide, and β_{mn} is the propagation constant. Since the EIM is based on the assumption that the mode field is a separable function of variables, $\psi_{mn}(x, y)$ can be written as

$$\psi_{mn}(x, y) = \mathcal{X}_m(x) \mathcal{Y}_n(y) \quad (4.38)$$

where $\mathcal{X}_m(x)$ and $\mathcal{Y}_n(y)$ are the mode field distributions for two planar waveguides in Figure 4.5(a) and Figure 4.5(b) respectively. Both field distributions are in the same form as that of Equation 4.17 with a sinusoidal profile of the field inside the core specified by the x and y components of the wave vector (k_x and k_y), and the exponential decay in the lower, upper and side cladding regions is described by the extinction coefficients γ_2 , γ_3 , and γ_4 respectively.

The wave equation in Equation 4.37 can give such solutions if the dielectric constant distribution function $\eta^2(x, y)$ is in the form [13].

$$\eta^2(x, y) = \eta_x^2(x) + \eta_y^2(y) \quad (4.39)$$

By substituting Equation 4.38 and Equation 4.39 into Equation 4.37, the following expression is derived.

$$\frac{1}{\mathcal{X}_m(x)} \frac{d^2 \mathcal{X}_m(x)}{dx^2} + \frac{1}{\mathcal{Y}_n(y)} \frac{d^2 \mathcal{Y}_n(y)}{dy^2} + [k_0^2 (\eta_x^2(x) + \eta_y^2(y)) - \beta_{mn}^2] = 0 \quad (4.40)$$

On the other hand, the modal field $\mathcal{Y}_n(y)$ for the waveguide in Figure 4.5(a) must still satisfy the following one-dimensional wave equation

$$\frac{1}{\mathcal{Y}_n(y)} \frac{d^2 \mathcal{Y}_n(y)}{dy^2} + [k_0^2 \eta_y^2(y) - (\beta_n^x)^2] = 0 \quad (4.41)$$

where the dielectric constant profile along the y -direction $\eta_y^2(y)$ is the same as that in Equation 4.1. By subtracting Equation 4.41 from Equation 4.40, the wave equation for the planar waveguide along the x axis, which provides the final value of the propagation constants β_{mn} of the rectangular waveguide, is derived.

$$\frac{1}{\mathcal{X}_m(x)} \frac{d^2 \mathcal{X}_m(x)}{dx^2} + \left[k_0^2 \left(\eta_x^2(x) + \frac{(\beta_n^x)^2}{k_0^2} \right) - \beta_{mn}^2 \right] = 0 \quad (4.42)$$

Since Equation 4.42 describes the modes of the planar waveguide in Figure 4.5(b), the dielectric constant term in the round brackets must be equal to $\left(n_{eff}^x\right)^2$ in the core region and $n_4^2 - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right)$ in the cladding regions. This condition will be satisfied if the value of $\eta_x^2(x)$ is given by

$$\eta_x^2(x) = \begin{cases} 0, & |x| < w \\ -\left(\left(n_{eff}^x\right)^2 - n_4^2\right) - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right), & |x| > w \end{cases} \quad (4.43)$$

$\begin{aligned} & n_3^2 \\ & - \left(\left(n_{eff}^x\right)^2 - n_4^2\right) \\ & - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$	n_3^2	$\begin{aligned} & n_3^2 \\ & - \left(\left(n_{eff}^x\right)^2 - n_4^2\right) \\ & - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$
$\begin{aligned} & n_4^2 \\ & + (1 - \zeta) \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$	n_1^2	$\begin{aligned} & n_4^2 \\ & + (1 - \zeta) \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$
$\begin{aligned} & n_2^2 \\ & - \left(\left(n_{eff}^x\right)^2 - n_4^2\right) \\ & - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$	n_2^2	$\begin{aligned} & n_2^2 \\ & - \left(\left(n_{eff}^x\right)^2 - n_4^2\right) \\ & - \zeta \left(n_1^2 - \left(n_{eff}^x\right)^2\right) \end{aligned}$

FIGURE 4.6: Equivalent waveguide by the effective index method

The combination of $\eta_x^2(x)$, from Equation 4.43, and $\eta_y^2(y)$, from Equation 4.1, gives the total dielectric constant $\eta^2(x, y)$ for all areas of the waveguide, which can be illustrated by the diagram in Figure 4.6. This is the equivalent structure of the waveguide analysed by the EIM. Compared with Figure 4.2, the dielectric constants in the side cladding regions are increased by $\left(n_1^2 - \left(n_{eff}^x\right)^2\right)$ whereas they are decreased by $\left(\left(n_{eff}^x\right)^2 - n_4^2\right)$ in the corner area in the case of the conventional EIM in which ζ is zero. These alterations in the dielectric constant profile lead to the error in the mode calculation. Introducing the parameter ζ adds extra terms in the dielectric profile which can decrease the original deviation. The deviation of the approximated propagation constant β by the EIM from the exact value $\tilde{\beta}$ can be quantified as the error parameter ε which is defined as

$$\varepsilon = \beta^2 - \tilde{\beta}^2 \quad (4.44)$$

For the first-order approximation by the perturbation theory in which changes in the dielectric constant profiles in Figure 4.2 are taken as a perturbation to the original structure, the error parameter ε can be determined from [10]

$$\varepsilon = k_0^2 \frac{\iint_{-\infty}^{+\infty} (\eta^2 - \tilde{\eta}^2) \psi \tilde{\psi} dx dy}{\iint_{-\infty}^{+\infty} \psi \tilde{\psi} dx dy} \quad (4.45)$$

where η and ψ are the refractive index and mode field of the equivalent waveguide by the EIM and $\tilde{\eta}$ and $\tilde{\psi}$ are those of the original waveguide. For the case of large values of normalised frequencies, most of the energy of the mode field is well confined within the core and therefore $\tilde{\psi}$ can be approximated by ψ . Since $(\eta^2 - \tilde{\eta}^2)$ is not zero only in the side cladding region and corner areas, Equation 4.45 can be simplified as [12]

$$\varepsilon = (1 - \zeta) \left(n_1^2 - (n_{eff}^x)^2 \right) k_0^2 \bar{P}_1 - \left(\zeta \left(n_1^2 - (n_{eff}^x)^2 \right) - \left(n_4^2 - (n_{eff}^x)^2 \right) \right) k_0^2 \bar{P}_2 \quad (4.46)$$

with \bar{P}_1 and \bar{P}_2 as the fractional powers in the side cladding regions and corner areas respectively. The confined field in the core area of the planar waveguide, in Figure 4.5(a), and the decaying field in the cladding area of the lower planar waveguide, in Figure 4.5(b), are contributed to \bar{P}_1 whereas \bar{P}_2 is the result of the fields in the cladding regions of both planar waveguides.

$$\bar{P}_1 = (P_x)_{cladding} (P_y)_{core} \quad (4.47a)$$

$$\bar{P}_2 = (P_x)_{cladding} (P_y)_{cladding} \quad (4.47b)$$

where P_x and P_y are the total fractional powers of the fields of the waveguides in Figure 4.5(a) and Figure 4.5(b) respectively, with the subscripts identifying the portion of the power in the specific areas. These power parameters can be expressed in the terms of the field distribution $\mathcal{X}(x)$ and $\mathcal{Y}(y)$ as

$$(P_x)_{cladding} = \frac{\int_{-\infty}^{-w} \mathcal{X}^2(x) dx + \int_{+w}^{+\infty} \mathcal{X}^2(x) dx}{\int_{-\infty}^{+\infty} \mathcal{X}^2(x) dx} \quad (4.48a)$$

$$(P_y)_{cladding} = \frac{\int_{-\infty}^{-d} \mathcal{Y}^2(y) dy + \int_{+d}^{+\infty} \mathcal{Y}^2(y) dy}{\int_{-\infty}^{+\infty} \mathcal{Y}^2(y) dy} \quad (4.48b)$$

$$(P_y)_{core} = 1 - (P_y)_{cladding} \quad (4.48c)$$

The parameter ζ is optimised when the error parameter ε becomes zero. In such a case, Equation 4.46 results in

$$\zeta_{opt} = 1 - \left(\frac{\bar{P}_2}{\bar{P}_1 + \bar{P}_2} \right) \left(\frac{n_1^2 - n_4^2}{n_1^2 - (n_{eff}^x)^2} \right) \quad (4.49)$$

With Equation 4.47 and Equation 4.48c, the fraction $\bar{P}_2 / (\bar{P}_1 + \bar{P}_2)$ can be simplified as $(P_y)_{cladding}$. By integrating the field $\mathcal{Y}(y)$ according to the expression in Equation 4.48b, the power fraction can be written as [12]

$$\frac{\bar{P}_2}{\bar{P}_1 + \bar{P}_2} = \frac{k_y^2}{\gamma_2 + \gamma_3 + 2d\gamma_2\gamma_3} \left(\frac{\gamma_2}{k_y^2 + \gamma_3^2} + \frac{\gamma_3}{k_y^2 + \gamma_2^2} \right) \quad (4.50)$$

Lastly the expression for ζ_{opt} is obtained by substituting Equation 4.50 into Equation 4.49

$$\zeta_{opt} = 1 - \frac{1}{\gamma_2 + \gamma_3 + 2d\gamma_2\gamma_3} \left[\left(\frac{n_1^2 - n_4^2}{n_1^2 - n_3^2} \right) \gamma_2 + \left(\frac{n_1^2 - n_4^2}{n_1^2 - n_2^2} \right) \gamma_3 \right] \quad (4.51)$$

The characteristic equations in Equation 4.29 are still utilised to determine the propagation constants and the extinction coefficients for the planar waveguides in Figure 4.5. In the case of E_{mn}^x modes, the characteristic equations for TE and TM modes are applied to the planar waveguide in Figure 4.5(a) and Figure 4.5(b) respectively whereas vice versa in the case of E_{mn}^y modes. The sets of the characteristic equation related to both modes can be summarised as the following.

E^x :

$$\tan(2k_y d) = \frac{k_y (\gamma_2 + \gamma_3)}{k_y^2 - \gamma_2\gamma_3} \quad (4.52a)$$

$$\bar{n}^2 k_x \tan(k_x w) = \gamma_4 \quad (4.52b)$$

E^y :

$$\tan(2k_y d) = \frac{k_y \left(\frac{\gamma_2}{n_{21}^2} + \frac{\gamma_3}{n_{31}^2} \right)}{k_y^2 - \frac{\gamma_2\gamma_3}{n_{21}^2 n_{31}^2}} \quad (4.52c)$$

$$k_x \tan(k_x w) = \gamma_4 \quad (4.52d)$$

where \bar{n}^2 is defined as

$$\bar{n}^2 = \frac{n_4^2 - \zeta_{opt} \left(n_1^2 - (n_{eff}^x)^2 \right)}{(n_{eff}^x)^2} \quad (4.53)$$

Note that the relations between $k_{x,y}$ and $\gamma_{2,3,4}$ are still the same as that of the equation Equation 4.14 which can be rewritten as

$$(k_y d)^2 + (\gamma_2 d)^2 = V_2^2 \quad (4.54a)$$

$$(k_y d)^2 + (\gamma_3 d)^2 = V_3^2 \quad (4.54b)$$

$$(k_x d)^2 + (\gamma_4 d)^2 = V_4^2 \quad (4.54c)$$

where

$$V_{2,3,4} = k_0 d \sqrt{n_1^2 - n_{2,3,4}^2} \quad (4.55)$$

The propagation constant β for a rectangular waveguide is related to the x and y components of the wave vector (k_x and k_y) and the extinction coefficients in the cladding area (γ_2 , γ_3 , and γ_4) as

$$\beta^2 = n_1^2 k_0^2 - k_x^2 - k_y^2 \quad (4.56a)$$

$$\gamma_2^2 = \beta^2 + k_x^2 - n_2^2 k_0^2 \quad (4.56b)$$

$$\gamma_3^2 = \beta^2 + k_x^2 - n_3^2 k_0^2 \quad (4.56c)$$

$$\gamma_4^2 = \beta^2 + k_y^2 - n_4^2 k_0^2 \quad (4.56d)$$

4.4 Nonlinear pulse propagation

A pulse propagating in a medium experiences various phenomena such as absorption loss, chromatic dispersion or nonlinearities. The strength of these effects depends on the

responsive characteristics of the medium to the electromagnetic field of the propagating light. The overall response contributes to the induced polarisation density which is related to the electric field. In order to clearly present all the influences on the alteration of the propagating pulse, the equation which relates the input and output of the system must be derived. This can be achieved with the Maxwell's equations. By application of the curl operation to the curl equation in Equation 4.5a, it results in

$$\nabla \times \nabla \times \mathbf{E} = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \frac{\partial^2}{\partial t^2} \nabla \times \mathbf{H} \quad (4.57)$$

The medium is assumed to be non-magnetic where $\mu \approx \mu_0$. From Maxwell's equations, the divergence of the electric field for a medium with no free electric charges is zero whereas the curl of $\nabla \times \mathbf{H}$ is equal to $-\partial \mathbf{D} / \partial t$ where \mathbf{D} is the electric displacement which is related to the electric field \mathbf{E} , the polarisation density \mathbf{P} and the electric permittivity in vacuum ϵ_0 by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (4.58)$$

Therefore by substituting Equation 4.58 in Equation 4.57, it gives

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2} \quad (4.59)$$

The complexity of solving Equation 4.59 for a light pulse can be simplified by the slowly varying envelope approximation (SVEA) in which the change in the field amplitude of the pulse along the distance of propagation is assumed to be very small in comparison with the oscillation of the field with frequency ω_0 . Therefore the scalar form of the field \mathbf{E} can be expressed as

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} [E(\mathbf{r}, t) \exp(-j\omega_0 t) + \text{c.c.}] \quad (4.60)$$

where the rapidly varying phase term $\exp(-j\omega_0 t)$ is separated from the amplitude term $E(\mathbf{r}, t)$ and c.c. denotes the complex conjugate part. The polarisation \mathbf{P} can be assumed to align in the same direction as the electric field and have the similar function form as that of the field. Therefore, \mathbf{P} can be also written as

$$\mathbf{P}(\mathbf{r}, t) = \frac{1}{2} [P(\mathbf{r}, t) \exp(-j\omega_0 t) + \text{c.c.}] \quad (4.61)$$

The electric field $\mathbf{E}(\mathbf{r}, t)$ can be represented as the superposition of monochromatic plane

waves as

$$\mathcal{E}(\mathbf{r}, t) = \int_{-\infty}^{+\infty} \tilde{\mathcal{E}}(\mathbf{r}, \omega) \exp(-j\omega t) \quad (4.62)$$

with the spectral amplitude $\tilde{\mathcal{E}}(\mathbf{r}, \omega)$ given by

$$\begin{aligned} \tilde{\mathcal{E}}(\mathbf{r}, \omega) &= \int_{-\infty}^{+\infty} \mathcal{E}(\mathbf{r}, t) \exp(j\omega t) dt \\ &= \frac{1}{2} \left[\tilde{E}(\mathbf{r}, \omega - \omega_0) + \tilde{E}(\mathbf{r}, \omega + \omega_0) \right] \end{aligned} \quad (4.63)$$

where $\tilde{E}(\mathbf{r}, \omega)$ is the Fourier transform of the field amplitude $E(\mathbf{r}, t)$. Equation 4.63 implies that the spectrum of the field is actually the spectrum of the amplitude function shifted by the amount of the central frequency ω_0 . Similarly the Fourier transform of the polarisation $\mathcal{P}(\mathbf{r}, t)$ can also be written as

$$\tilde{\mathcal{P}}(\mathbf{r}, \omega) = \frac{1}{2} \left[\tilde{P}(\mathbf{r}, \omega - \omega_0) + \tilde{P}(\mathbf{r}, \omega + \omega_0) \right] \quad (4.64)$$

Since the derivation of the pulse propagation equation in the frequency domain is much easier, it is preferable to convert Equation 4.59 to the equation for the Fourier transform of the field. By substituting Equation 4.60 and Equation 4.61 into Equation 4.59 and applying the Fourier transform to both sides of the equation, the equivalent wave equation of the field amplitude $\tilde{E}(\mathbf{r}, \omega - \omega_0)$ in the frequency domain can be expressed as

$$\nabla^2 \tilde{E} + k^2 \tilde{E} = -(\mu_0 \omega^2) \tilde{P} \quad (4.65)$$

where $k = \omega/c$ and the polarisation $\tilde{P}(\mathbf{r}, \omega)$ is linearly and nonlinearly proportional to the field amplitude $\tilde{E}(\mathbf{r}, \omega)$. For the propagation of ultrashort pulse light in a dispersive nonlinear medium whose intensity is high enough to induce significant nonlinearity, the polarisation density $\mathcal{P}(\mathbf{r}, t)$ can be perceived as the output of a nonlinear system responding to the input $\mathcal{E}(\mathbf{r}, t)$. Therefore, $\mathcal{P}(\mathbf{r}, t)$ can be represented by the Volterra series as [2, 15]

$$\begin{aligned} \mathcal{P}(\mathbf{r}, t) &= \varepsilon_0 \sum_{n=1}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \chi^{(n)}(t - t_1, \dots, t - t_n) \\ &\quad \times \mathcal{E}(\mathbf{r}, t_1) \cdots \mathcal{E}(\mathbf{r}, t_n) dt_1 \cdots dt_n \end{aligned} \quad (4.66)$$

where $\chi^{(n)}$ is the n -th order of the Volterra kernel and related to the n -th order electric

susceptibility. Equation 4.66 consists of both linear \mathcal{P}_L and nonlinear \mathcal{P}_{NL} parts which contribute to the total polarisation \mathcal{P} .

$$\mathcal{P} = \mathcal{P}_L + \mathcal{P}_{NL} \quad (4.67)$$

The linear polarisation \mathcal{P}_L is given by the first-order term of the series which can be written as

$$\mathcal{P}_L(\mathbf{r}, t) = \mathcal{P}^{(1)}(\mathbf{r}, t) = \varepsilon_0 \int_{-\infty}^{+\infty} \chi^{(1)}(t - t_1) \mathcal{E}(\mathbf{r}, t_1) dt_1 \quad (4.68)$$

where $\chi^{(1)}(t)$ is the linear susceptibility function of time. By the convolution theorem, the corresponding Fourier transform $\tilde{P}_L(\mathbf{r}, \omega)$ can be determined from

$$\tilde{P}_L(\mathbf{r}, \omega) = \varepsilon_0 \tilde{\chi}^{(1)}(\omega) \tilde{E}(\mathbf{r}, \omega) \quad (4.69)$$

where $\tilde{\chi}^{(1)}(\omega)$ is the Fourier transform of $\chi^{(1)}(t)$ which describes the frequency dependence of the linear susceptibility. If the system consists only of the linear response, the complex $\tilde{\chi}^{(1)}(\omega)$ is related to the dielectric constant $\varepsilon(\omega)$, refractive index $n(\omega)$ and absorption coefficient $\alpha(\omega)$ of the system as [2]

$$\varepsilon(\omega) = \left(n(\omega) + j \frac{\alpha(\omega)}{2k_0} \right)^2 = 1 + \tilde{\chi}^{(1)}(\omega) \quad (4.70)$$

From Equation 4.70, the complex value of $\varepsilon(\omega)$ can be separated to the real (\Re) and imaginary (\Im) parts which can be expressed as

$$\Re[\varepsilon(\omega)] = n^2(\omega) + \frac{\alpha(\omega)^2}{4k_0^2} = 1 + \Re[\tilde{\chi}^{(1)}(\omega)] \quad (4.71a)$$

$$\Im[\varepsilon(\omega)] = \frac{\alpha(\omega)n(\omega)}{k_0} = \Im[\tilde{\chi}^{(1)}(\omega)] \quad (4.71b)$$

For a weakly absorbing medium in which $\alpha(\omega)$ is very small and the $\alpha^2(\omega)$ term can be neglected, $n(\omega)$ and $\alpha(\omega)$ can be determined from

$$n(\omega) \approx \left(1 + \Re[\tilde{\chi}^{(1)}(\omega)] \right)^{1/2} \approx 1 + \frac{1}{2} \Re[\tilde{\chi}^{(1)}(\omega)] \quad (4.72)$$

$$\alpha(\omega) \approx \frac{k_0}{n_0} \Im[\tilde{\chi}^{(1)}(\omega)] \quad (4.73)$$

Note that Equation 4.72 is derived by assuming that $\Re[\tilde{\chi}^{(1)}(\omega)] \ll 1$ and the approximation by the first two terms of the Taylor expansion can be applied, whereas the refractive index at the central frequency ω_0 is utilised instead of $n(\omega)$ on the right side of Equation 4.73. Therefore it can be roughly stated that the linear refractive index and absorption coefficient are given by the real and imaginary parts of the first-order susceptibility.

The nonlinear polarisation begins with the second-order term of Equation 4.66 which is

$$\mathcal{P}^{(2)}(\mathbf{r}, t) = \varepsilon_0 \int \int_{-\infty}^{+\infty} \chi^{(2)}(t - t_1, t - t_2) \mathcal{E}(\mathbf{r}, t_1) \mathcal{E}(\mathbf{r}, t_2) dt_1 dt_2 \quad (4.74)$$

The second-order kernel $\chi^{(2)}(t, t')$ is a function of two variables and related to the second-order susceptibility which becomes zero for a centrosymmetric material. Therefore if the higher-order terms of Equation 4.66 are small enough to be truncated, the nonlinear polarisation \mathcal{P}_{NL} is approximated by the third-order term of Equation 4.66

$$\begin{aligned} \mathcal{P}_{NL}(\mathbf{r}, t) \approx \mathcal{P}^{(3)}(\mathbf{r}, t) = & \varepsilon_0 \int \int \int_{-\infty}^{+\infty} \chi^{(3)}(t - t_1, t - t_2, t - t_3) \\ & \times \mathcal{E}(\mathbf{r}, t_1) \mathcal{E}(\mathbf{r}, t_2) \mathcal{E}(\mathbf{r}, t_3) dt_1 dt_2 dt_3 \end{aligned} \quad (4.75)$$

with the third-order kernel $\chi^{(3)}(t, t', t'')$ related to the third-order susceptibility. In the non-resonant regime, the nonlinear response is proportional to the optical intensity and then Equation 4.75 can be rewritten as [16]

$$\mathcal{P}_{NL}(\mathbf{r}, t) = \varepsilon_0 \chi^{(3)} \mathcal{E}(\mathbf{r}, t) \int_{-\infty}^{+\infty} h(t - t_1) |\mathcal{E}(\mathbf{r}, t_1)|^2 dt_1 \quad (4.76)$$

where $h(t)$ is the impulse response of the system and normalised such that $\int_{-\infty}^{+\infty} h(t) dt = 1$. Several nonlinear phenomena are attributed to the third-order nonlinearity [17]. For the system that lacks phase matching, however, the Kerr and Raman effects are probably the most substantial nonlinear phenomena. The first is the outcome of the instantaneous electronic response of the system whereas the latter is attributed to the time-delayed response owing to phonon relaxation. In such a case, $h(t)$ can be represented by

$$h(t) = (1 - \varrho) \delta(t) + \varrho h_R(t) \quad (4.77)$$

The first term represents the instantaneous phenomenon which is expressed by the delta function whereas the last term configures the Raman response as a function of time $h_R(t)$ and the fractional parameter ϱ specifies the relative strength of the Raman effect

to the total nonlinearities. The model of damped oscillation is exploited to characterise the time-delayed response and therefore $h_R(t)$ can be written as [16]

$$h_R(t) = \frac{\tau_1^2 + \tau_2^2}{\tau_1 \tau_2^2} \exp(-t/\tau_2) \sin(t/\tau_1) \quad (4.78)$$

where τ_1 and τ_2 are parameters characterised according to the Raman gain spectrum whose phonon central frequency and spectral linewidth are given by $1/\tau_1$ and $1/\tau_2$ respectively. However, if the pulse width is much longer than the nonlinear response time, the Kerr effect can be assumed to be much more prevalent and therefore the second term on the right-hand side of Equation 4.77 can be omitted. By substituting $h(t) = \delta(t)$ in Equation 4.76, it can be written as

$$P_{NL}(\mathbf{r}, t) = \varepsilon_0 \varepsilon_{NL} E(\mathbf{r}, t) \quad (4.79)$$

with ε_{NL} denoting the nonlinear permittivity, which is given by

$$\varepsilon_{NL} = \frac{3}{4} \chi^{(3)} |E(\mathbf{r}, t)|^2 \quad (4.80)$$

Note that the factor 3/4 is from the multiplication of $E(\mathbf{r}, t) \exp(-j\omega_0 t)$ and its conjugate terms. Actually \mathcal{P} does not contain only the fundamental frequency ω_0 terms but also the third-harmonic $3\omega_0$ terms, but only the fundamental term with the coefficient 3/4 is retained.

By substituting the Fourier transform of $\mathcal{P}_{NL}(\mathbf{r}, \omega)$ and Equation 4.69 for $\tilde{P}(\mathbf{r}, \omega)$, Equation 4.65 becomes

$$\nabla^2 \tilde{E} + \tilde{\varepsilon}(\omega) k_0^2 \tilde{E} = 0 \quad (4.81)$$

where the complex dielectric constant $\tilde{\varepsilon}(\omega)$ is given by

$$\tilde{\varepsilon}(\omega) = 1 + \tilde{\chi}^{(1)}(\omega) + \varepsilon_{NL} \quad (4.82)$$

In contrast to $\varepsilon(\omega)$ in Equation 4.70, the dielectric constant $\tilde{\varepsilon}(\omega)$ depends on the field intensity and so does the corresponding refractive index $\tilde{n}(\omega)$ and absorption coefficient $\tilde{\alpha}(\omega)$, which are related to $\tilde{\varepsilon}(\omega)$, the same as Equation 4.70. Therefore $\tilde{n}(\omega)$ and $\tilde{\alpha}(\omega)$ can be determined from

$$\tilde{n}^2(\omega) \approx 1 + \Re[\tilde{\chi}^{(1)}(\omega)] + \Re[\varepsilon_{NL}] \quad (4.83)$$

$$\tilde{\alpha}(\omega) \approx \frac{k_0}{n_0} \left(\Im \left[\tilde{\chi}^{(1)}(\omega) \right] + \Im[\varepsilon_{NL}] \right) \quad (4.84)$$

The intensity dependence of both parameters can be emphasized by expressing $\tilde{n}(\omega)$ and $\tilde{\alpha}(\omega)$ in such a form that

$$\tilde{n}(\omega) = n(\omega) + n_2 |E|^2 \quad (4.85)$$

$$\tilde{\alpha}(\omega) = \alpha(\omega) + \alpha_2 |E|^2 \quad (4.86)$$

where $n(\omega)$ and $\alpha(\omega)$ are the linear refractive index and absorption coefficient as already defined in Equation 4.72 and Equation 4.73 respectively, whereas the nonlinear refractive index coefficient, n_2 , and two-photon absorption coefficient, α_2 , can be determined from

$$n_2 = \frac{3}{8n_0} \Re \left[\chi^{(3)} \right] \quad (4.87)$$

$$\alpha_2 = \frac{3}{4n_0} \Im \left[\chi^{(3)} \right] \quad (4.88)$$

Equation 4.81 can be solved by the method of separation of variables in which the z -dependence of $\tilde{E}(\mathbf{r}, \omega)$ is assumed to be separable from the function of x and y . Therefore, $\tilde{E}(\mathbf{r}, \omega)$ can be written as

$$\tilde{E}(\omega - \omega_0) = U(x, y) \tilde{V}(z, \omega - \omega_0) \exp(j\beta_0 z) \quad (4.89)$$

where $U(x, y)$ describes the transverse field profile and $\tilde{V}(z, \omega - \omega_0)$ is a slowly varying amplitude function of z with $V(z, t)$ as the corresponding function in the time domain. By substituting Equation 4.89 into Equation 4.81, two equations of the field are derived.

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \left(\tilde{\varepsilon}(\omega) k_0^2 - \hat{\beta}^2 \right) U = 0 \quad (4.90)$$

$$\frac{\partial \tilde{V}}{\partial z} - j \frac{1}{2\beta_0} \left(\hat{\beta}^2 - \beta_0^2 \right) \tilde{V} = 0 \quad (4.91)$$

Equation 4.91 is based on the assumption that $\partial^2 \tilde{V} / \partial t^2 \simeq 0$ owing to the slowly varying \tilde{V} along the distance of propagation z . Equation 4.90 characterises the transverse field $U(x, y)$ with the eigenvalue $\hat{\beta}$ as the propagation constant of the field. On the other

hand, Equation 4.91 describes the development of a light pulse envelope affected by the dispersion and nonlinearity. Both equations are coupled with the propagation constant $\hat{\beta}$ which can be determined from Equation 4.90, in which $\hat{\beta}$ is the eigenvalue. The perturbation theory can be applied to solve the equation. In this approach, the unperturbed system contains only the linear effect and its eigenvalue can be approximately determined by the effective index method which has already been explained in Section 4.3. Now the nonlinear effect is treated as the perturbation and $\hat{\beta}$ can be approximated by the perturbation theory. In order to separate the linear and nonlinear operators, the dielectric constant $\tilde{\epsilon}(\omega)$ should be expressed as

$$\tilde{\epsilon}(\omega) = \left(\tilde{n}(\omega) + j \frac{\tilde{\alpha}}{2k_0} \right) = (n(\omega) + \Delta n)^2 \approx n^2(\omega) + 2n(\omega)(\Delta n) \quad (4.92)$$

where $n(\omega)$ is the linear refractive index and from Equation 4.85, the perturbing factor Δn in Equation 4.92 is given by

$$\Delta n = n_2 |E|^2 + j \frac{\tilde{\alpha}}{2k_0} \quad (4.93)$$

If $\beta(\omega)$ is the propagation constant, determined from the unperturbed system in which $\tilde{\epsilon}(\omega)$ is replaced by $n^2(\omega)$, by the first-order perturbation theory, the true value $\hat{\beta}(\omega)$ can be determined from

$$\hat{\beta}^2(\omega) = \beta^2(\omega) + \frac{\iint_{-\infty}^{+\infty} (2k_0^2 n \Delta n) |U(x, y)|^2 dx dy}{\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy} \quad (4.94)$$

Note that the denominator $\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy$ is the normalising factor which gives the unity power of transverse field and the total power of the propagating field is given by $|V(z)|^2$. If the perturbation is very small, from Equation 4.94 the eigenvalue $\hat{\beta}$ can be approximated as

$$\hat{\beta}(\omega) \approx \beta(\omega) + \Delta\beta \quad (4.95)$$

where $\Delta\beta$ is defined as

$$\Delta\beta = k_0 \frac{\iint_{-\infty}^{+\infty} \Delta n |U(x, y)|^2 dx dy}{\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy} \quad (4.96)$$

By substituting Equation 4.93 into Equation 4.96 with

$$|E|^2 = \frac{|U(x, y)|^2 |\tilde{V}(z, \omega)|^2}{\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy} \quad (4.97)$$

it can be written as

$$\Delta\beta = \gamma |V(z)|^2 + j \frac{\alpha}{2} \quad (4.98)$$

where γ is the nonlinear parameter which is defined as

$$\gamma = \frac{n_2 k_0}{A_{eff}} \quad (4.99)$$

with A_{eff} as effective mode area which is given by

$$A_{eff} = \frac{\left(\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy \right)^2}{\iint_{-\infty}^{+\infty} |U(x, y)|^4 dx dy} \quad (4.100)$$

The nonperturbed eigenvalue $\beta(\omega)$ can be expanded around the central frequency ω_0 by a Taylor series as

$$\beta(\omega) = \beta_0 + \sum_{m \geq 1} \frac{\beta_m}{m!} \Delta\omega^m \quad (4.101)$$

$$= \beta_0 + \beta_1 \Delta\omega + \frac{1}{2} \beta_2 \Delta\omega^2 + \frac{1}{6} \beta_3 \Delta\omega^3 + \dots \quad (4.102)$$

where

$$\beta_m = \left. \frac{\partial^m \beta}{\partial \omega^m} \right|_{\omega=\omega_0} \quad (4.103)$$

The zeroth-order term of the series in Equation 4.102 is related to the phase velocity of the light whereas the reciprocal of the first derivative β_1 provides the group velocity, v_g , of the pulse envelope.

$$v_g = \frac{\partial \omega}{\partial \beta} = \frac{1}{\beta_1} \quad (4.104)$$

The frequency dependence of the group velocity is described by the group velocity dispersion (GVD) which is related to the second-order term in Equation 4.102. The quantity of GVD is commonly measured by the parameter D which is defined as

$$D = \frac{\partial \beta_1}{\partial \lambda} = \frac{\partial \beta_1}{\partial \omega} \frac{\partial \omega}{\partial \lambda} = -\frac{2\pi c}{\lambda^2} \beta_2 \quad (4.105)$$

The next terms in the series contribute to the higher-order dispersion which can be negligible as long as $\Delta\omega \ll \omega_0$. The approximation becomes inaccurate when the pulse is very short in the order less than picoseconds [18]. However, only the first three terms in Equation 4.102 are utilised for $\beta(\omega)$ in this research. By Equation 4.95, Equation 4.98 and Equation 4.102, the term $\hat{\beta}^2 - \beta_0^2$ can be expressed as

$$\begin{aligned} \hat{\beta}^2 - \beta_0^2 &\approx 2\beta_0 (\beta(\omega) + \Delta\beta - \beta_0) \\ &\approx 2\beta_0 \left(\sum_{m \geq 1} \frac{\beta_m}{m!} (\Delta\omega)^m + \gamma |\tilde{V}(z, \omega)|^2 + j \frac{\alpha}{2} \right) \end{aligned} \quad (4.106)$$

Substituting Equation 4.106 into Equation 4.91 gives

$$\frac{\partial \tilde{V}(z, \omega)}{\partial z} + \frac{\alpha}{2} - j \left(\sum_{m \geq 1} \frac{\beta_m}{m!} (\Delta\omega)^m \right) \tilde{V}(z, \omega) = j\gamma |\tilde{V}(z, \omega)|^2 \tilde{V}(z, \omega) \quad (4.107)$$

The corresponding equation with Equation 4.107 for the pulse envelope in the time domain is obtained by replacing $\Delta\omega$ with $j\partial/\partial t$.

$$\frac{\partial V(z, t)}{\partial z} + \frac{\alpha}{2} + \sum_{m \geq 1} \frac{(j)^{m-1} \beta_m}{m!} \frac{\partial^m V(z, t)}{\partial t^m} = j\gamma |V(z, t)|^2 V(z, t) \quad (4.108)$$

Equation 4.108 is generally referred to as the nonlinear Schrödinger equation which can be either analytically or numerically solved. The process of solving this equation will be described in the following chapters relating to the numerical study of light pulses in nonlinear waveguides.

4.5 Chapter conclusion

The propagation modes and the pulse propagation in a nonlinear waveguide have been explained and the related equations, which will be utilised later in the following chapters, are formulated in this chapter. The modal field and the propagation parameters of the guided modes in a planar waveguide with a one-dimensional step index profile can be described by Maxwell's equations with the boundary conditions and ray optics. The propagation modes, which depend on the polarisation of the light, the refractive index of the waveguide medium and the dimension of the waveguide, are specified by the

characteristic equation which is also the basis for determining the propagation modes in a two-dimensional waveguide. For the case of rectangular waveguides, the effective index method is generally adopted to approximate the propagation constants of the guided modes. The approach is based on the separation of variables by which the two-variable function of the field can be separated into two independent functions of each variable. Therefore the propagation can be solved from two planar waveguides instead. The index profiles of those waveguides are perpendicular to each other where the value of refractive index of one waveguide is specified by the propagation modes of the other. The accuracy can be further improved by the application of perturbation theory. An additional parameter is added to the index profile of the waveguides and optimised to nullify the error approximated by the perturbation theory.

Finally, pulse propagation in a nonlinear medium has been theoretically explained. The response of the medium toward the optical field is in the form of the polarisation density which is linearly and nonlinearly proportional to the field. The linear response, proportional to the first-order susceptibility, corresponds to the dispersion of the light pulses. On the other hand, the higher-order electric susceptibilities contribute to the nonlinear response of which only the instantaneous and intensity-dependent Kerr effect is retained for the numerical study of this research. All these effects upon the electric field are formulated as a nonlinear partial differential equation which will be later numerically solved to describe the pulse propagation in the nonlinear waveguide examined in this research.

4.6 References

- [1] G. P. Agrawal. *Lightwave Technology: Components and Devices*. John Wiley & Sons, Hoboken, NJ, USA, 2004.
- [2] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. John Wiley & Sons Inc., New York.
- [3] R. T. Hammond. Third-order Miller-type rule from coherent states. *Journal of the Optical Society of America B (Optical Physics)*, 7(6):944–945, 1990.
- [4] T. Hashimoto and T. Yoko. Third-order nonlinear optical properties of sol-gel-derived V_2O_5 , Nb_2O_5 , and Ta_2O_5 thin films. *Applied Optics*, 34(16):2941–2948, 1995.
- [5] R. C. Miller. Optical second harmonic generation in piezoelectric crystals. *Applied Physics Letters*, 5(1):1719, 1964.
- [6] D.-G. Chen, B. G. Potter, and J. H. Simmons. GeO_2 - SiO_2 thin films for planar waveguide applications. *Journal of Non-Crystalline Solids*, 178:135–147, 1994.

- [7] G. Lifante. *Integrated Photonics: Fundamentals*. John Wiley & Sons, Chichester, England, 2003.
- [8] A. Martucci, G. Brusatin, M. Guglielmi, C. Strohhofer, J. Fick, S. Pelli, and G. C. Righini. Fabrication and characterization of sol-gel $\text{GeO}_2\text{-SiO}_2$ erbium-doped planar waveguides. *Journal of Sol-Gel Science and Technology*, 13(1-3):535–539, 1998.
- [9] K. S. Chiang. Review of numerical and approximate methods for the modal analysis of general optical dielectric waveguides. *Optical and Quantum Electronics*, 26(3):S113–S134, 1994.
- [10] A. W. Snyder and J. Love. *Optical Waveguide Theory*. Kluwer Academic Publishers, UK, 1983.
- [11] R. M. Knox and P. O. Toullos. Integrated circuits for the millimeter through optical frequency range. In *Symposium on Submillimetre waves*, pages 497–516, New York, NY, USA, 1970. Polytechnic Press, New York, NY, USA.
- [12] K. S. Chiang, K. M. Lo, and K. S. Kwok. Effective-index method with built-in perturbation correction for integrated optical waveguides. *Journal of Lightwave Technology*, 14(2):223–228, 1996.
- [13] A. Kumar, D. F. Clark, and B. Culshaw. Explanation of errors inherent in the effective-index method for analyzing rectangular-core waveguides. *Optics Letters*, 13(12):1129–1131, 1988.
- [14] K. S. Chiang. Analysis of rectangular dielectric waveguides: effective-index method with built-in perturbation correction. *Electronics Letters*, 28(4):388–390, 1992.
- [15] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarigue. An algebraic approach to nonlinear functional expansions. *IEEE Transactions on Circuits and Systems*, 30(8):554–570, 1983.
- [16] K. J. Blow and D. Wood. Theoretical description of transient stimulated Raman scattering in optical fibers. *IEEE Journal of Quantum Electronics*, 25(12):2665–2673, 1989.
- [17] R. W. Boyd. *Nonlinear optics*. Academic Press, San Diego, CA, 2nd edition, 2003.
- [18] G. P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, CA, USA, 3rd edition, 2001.

Chapter 5

Numerical modeling of nonlinear pulse propagation

5.1 Chapter introduction

The spectra from a nonlinear waveguide observed by the SNOM probe in this research contain several unique features which cannot be explained by a simple numerical model of single-mode pulse propagation. The prominent deviation between the measured spectra and the simulated results from the single-mode model, which will be presented later in Chapter 6, is assumed to be owing to the unique multimode property [1, 2]. This motivates the development of a novel model for the nonlinear pulse propagation which can include the multimode effect. In this chapter, the principles of the modeling will be described whereas its results will be presented along with the experimentally measured data in the following chapters.

Firstly, a numerical method, namely the split-step Fourier method, on which the simulation of this research is based will be explained in Section 5.2. The section includes the basic concept of the method which is based on dividing the total span of the waveguide into many tiny segments in which the separation of linear and nonlinear interaction can be assumed with reasonable accuracy. Then the accuracy of the approach and the degree of the approximation error will be discussed.

Section 5.3 explains the determination of the propagation constant and its derivatives whose value will be used for the dispersion effect in the simulation. It starts with the application of the effective index method to determine the propagation modes in the waveguide. To acquire the higher-order dispersion constants, it is necessary to obtain the wavelength dependency of the propagation constant and a differentiation operation is used to determine the derivatives. Also the discrepancies of the approach are discussed in the section. Although the task can be performed by commercial software, a source

code for mode calculation was developed in this research in order to be integrated with the source code of the simulation of the pulse propagation. The accuracy of the mode calculation was testified by comparing with the results generated by the finite difference method (FDM) from a commercial mode-solving software.

In addition, the mode determination provides the mode field distribution, which will be discussed in Section 5.4. The field distribution is elementary for the determination of the intensity-dependent nonlinear processes: self and cross-phase modulations. Moreover, the transverse field is required in the theoretical calculation of the total field perceived by the SNOM probe to simulate the visualisation of the local variation of the nonlinear spectra.

Then Section 5.5 explains the adaptation of the nonlinear Schrödinger equation for pulse propagation in the multimode waveguide. Basically it contains a set of differential equations each of which is specified for the propagation of the particular mode. However these equations are coupled in their nonlinear part with the other contributing modes. The nonlinear coupling is parameterised as the nonlinear parameter whose value depends on the degree of overlapping of the mode field distribution. The envelope field equations are numerically solved individually and then the results are combined with the transverse field to form the SNOM-acquired spectra.

The last section, Section 5.6, will discuss the improvement of the simulation efficiency. The section begins with the investigation of the step-size effect on the numerical error and the optimisation of the size value. Then the scheme adapted from the methods utilised for solving ordinary differential equations will be proposed to enhance the numerical accuracy. Finally there will be another suggestion for improving the calculation efficiency by the adaptive step size algorithm in which instead of being constant the size of the numerical step is adjusted along all the calculation process in accordance with the local error occurring at each step.

5.2 Split-step Fourier Method

5.2.1 Basic algorithm

For the specific case of the nonlinear Schrödinger (NLS) equation which does not contain a perturbation term, the equation becomes integrable and can be analytically solved by applying the inverse scattering transform [3]. The approach is complicated and therefore, in general, the behaviour of a light pulse in a dispersive and nonlinear medium is numerically investigated. One of the most widely utilised numerical algorithms is the split-step Fourier method (SSFM) which was first introduced by Hardin et al. [4] and later applied to an optical system by Fisher et al. [5]. The algorithm divides the waveguide span into many small segments. In each segment the linear and nonlinear

effects are treated separately. The output from the segment is then fed, as the input, into the next segment. This operation is iterated over all the waveguide length.

In each segment, the temporal and spatial evolution of the propagating pulse can be described by the NLS equation which has been already derived in Equation 4.108, Chapter 4, and can be rewritten as

$$\frac{\partial V(z, t)}{\partial z} = \left(\hat{\mathcal{L}} + \hat{\mathcal{N}} \right) V(z, t) \quad (5.1)$$

In Equation 5.1, the waveguide is treated as a system consisting of linear $\hat{\mathcal{L}}$ and nonlinear $\hat{\mathcal{N}}$ operators. The first manipulates the system's dispersion and absorption loss, whereas the effect of nonlinearities are dealt with in the latter. For a basic nonlinear waveguide, which contains only second-order dispersion and self-phase modulation, the operators become

$$\hat{\mathcal{L}} = -\frac{\alpha}{2} - \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} \quad (5.2)$$

$$\hat{\mathcal{N}} = j\gamma |V(z, t)|^2 \quad (5.3)$$

The parameter t is the retarded time with the frame of reference moving with the group velocity v_g and related to the real time, t' , by $t = t' - z/v_g$. The parameter β_2 is the GVD parameter which is the inverse of the group velocity and its basic unit is s^2/m . The nonlinear parameter, γ , the SPM phase change per unit optical power, describes the strength of the nonlinear effect and its unit is W^{-1}/m . The detail of the parameter will be described later in Section 5.5.

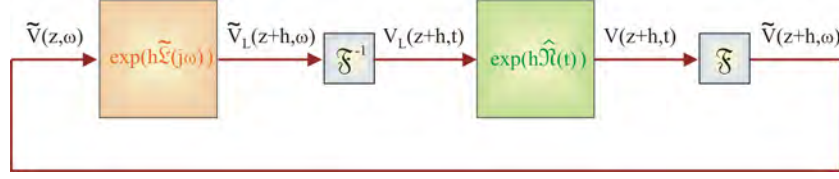
For each segment of length h with the input envelope field $V(z, t)$, the exact solution $V(z + h, t)$ of Equation 5.1 can be written as

$$V(z + h, t) = \exp \left[h \left(\hat{\mathcal{L}} + \hat{\mathcal{N}} \right) \right] V(z, t) \quad (5.4)$$

To simplify, the linear and nonlinear effects are assumed to interact independently on the input pulse at different times and, therefore, $V(z + h, t)$ in Equation 5.4 can be approximated as

$$V(z + h, t) \approx \exp \left(h\hat{\mathcal{L}} \right) \exp \left(h\hat{\mathcal{N}} \right) V(z, t) \quad (5.5)$$

The nonlinear part of the equation can be more easily solved in time whereas analysing the solution of the linear part is achievable in the frequency domain. The linear operation

FIGURE 5.1: 1st-order algorithm for modeling pulse propagation

is mostly related to a differentiation operation which is equivalent to the multiplication operation in frequency. Therefore the Fourier transform is applied to change the field back and forth between two domains. Figure 5.1 shows the scheme of basic SSFM algorithm. Firstly only the linear effect is applied to the input field in frequency for the whole length h of the step, which provides the outcome $\tilde{V}_L(z+h, \omega)$.

$$\tilde{V}_L(z+h, \omega) = \exp\left(h\tilde{\mathcal{L}}(j\omega)\right) \tilde{V}(z, \omega) \quad (5.6)$$

where $\tilde{V}(z, \omega)$ is the Fourier transform of $V(z, t)$ and can be calculated by the relation

$$\tilde{V}(z, \omega) = \int_{-\infty}^{+\infty} V(x, y) \exp[j(\omega - \omega_0)t] dt \quad (5.7)$$

and $\tilde{\mathcal{L}}(j\omega)$ is the frequency response of the linear operator $\hat{\mathcal{L}}(t)$ which is derived from substituting all differentiation terms $\partial^n/\partial t^n$ with $(j(\omega - \omega_0))^n$.

Then $\tilde{V}_L(z+h, \omega)$ is inversely transformed to $V_L(z+h, t)$ in the time domain to be operated by $\exp(h\hat{\mathcal{N}})$. The output of the current step $V(z+h, t)$ will be further forwarded as the input of the next step and the process is repeated.

5.2.2 Numerical errors and 2nd-order approximation

There are two types of error which are related to numerical simulations. The *round-off error* originates from the imperfect representation of real numbers by floating point numbers. For each floating point operation this kind of inaccuracy is bound by $O(\tilde{\epsilon})$ where the big O notation, generally used in mathematics and computer science, describes the growth rate of the error and $\tilde{\epsilon}$ is the smallest value represented by the calculating system. Since all the calculation in this research is executed in MATLAB, which applies the double-precision format by default, the value of $\tilde{\epsilon}$ is 2.22×10^{-16} . This participates in the *local error* which is incurred at every step of the modeling and keeps accumulating with the increasing number of segments during the simulation process. If h is the step size, the overall number of steps is proportional to $1/h$ and therefore the total round-off error of the modeling or the *global error* is $O(\frac{\tilde{\epsilon}}{h})$.

Another type of error is the *truncation error* which is the deviation of the approximation from the true solution owing to omitting some higher-order terms. In contrast to the round-off order, the truncation error varies depending on the algorithm adopted in the simulation. For the SSFM algorithm, utilising the approximation in Equation 5.5, the quantity of the truncation error can be determined by applying the Baker-Campbell-Hausdorff formula. For two non-commutable operators, x and y , it can be written as [6]

$$\exp(x)\exp(y) = \exp\left((x+y) + \frac{1}{2}[x,y] + \frac{1}{12}[x,[x,y]] - \frac{1}{12}[y,[x,y]] - \dots\right) \quad (5.8)$$

where $[x,y] = xy - yx$. If x and y are substituted by $h\hat{\mathfrak{N}}$ and $h\hat{\mathfrak{L}}$ respectively, by comparing with Equation 5.4, the first term on the right side of Equation 5.8 corresponds with the true solution of the NLS equation whereas the term on the left side represents the approximation solution in Equation 5.5. The latter terms on the right side of Equation 5.8 are accordingly: the discrepancy of the approximation starting with the term proportional to h^2 (which is the most significant one) and followed by the terms with higher-order powers of step size. As a result, the local error is bound by $O(h^2)$ and the global error is $O(h)$. Therefore the SSFM in Equation 5.5 is classed as the 1st-order approximation.

To enhance the accuracy, Equation 5.5 can be adjusted into the symmetrised form as [7]

$$V(z+h, t) \approx \exp\left(\frac{h}{2}\hat{\mathfrak{L}}\right) \exp\left(\int_z^{z+h} \hat{\mathfrak{N}}(z') dz'\right) \exp\left(\frac{h}{2}\hat{\mathfrak{L}}\right) V(x, y) \quad (5.9)$$

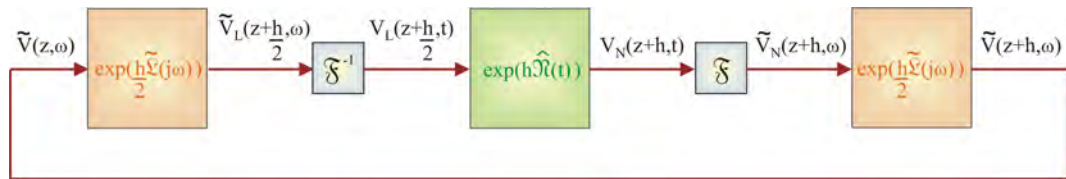


FIGURE 5.2: 2nd-order algorithm for modeling pulse propagation

In Equation 5.9, the nonlinear interaction is sandwiched by the separated linear interaction, applied twice over the half-step distance length, $h/2$, at the first and last halves of the step. The nonlinear operation includes taking the distance-dependency of the nonlinearities into account. However for small step size, $\hat{\mathfrak{N}}$ can be assumed to be constant and the integral term is reduced to $h\hat{\mathfrak{N}}$. The calculation procedure of this symmetrised scheme is displayed in Figure 5.2. Compared with Figure 5.1, the new procedure consists of another linear operation for the second half of the segment which is just the multiplication between the frequency response of the linear operation and the previous output of the nonlinear operation without an additional Fourier transform.

This implies no significant increase in the calculation expense since the main cost of the calculation comes from the performance of the Fourier transform. On the contrary, by applying Equation 5.8 for an extra multiplication between the operators, the discrepancy between the exact and new approximated solution is found to be proportional to h^3 . Therefore the accuracy of the symmetrised scheme is confined to $O(h^3)$ locally and $O(h^2)$ globally. This algorithm is classed as the 2nd-order approximation.

The numerical modeling in this research, whose results will be presented further in later chapters, has adopted the second-order scheme. Although the algorithm of higher-order approximation is also available, the overall efficiency is not improved owing to the increase of calculation cost. The details of the higher-order approximation will be given later in Section 5.6.2.

5.3 Determination of propagation modes and dispersion parameters

The linear operator of the NLS equation consists of the terms relating to the dispersion characteristic of the propagation modes. These effects are parameterised as the propagation constant and its derivatives. The value of these parameters can be derived from the determination of the guided modes of the waveguide. In addition, the wavelength variation of the mode propagation constant is required in order to enable the differentiation for the parameters of high-order dispersion, such as the group velocity dispersion or the third-order dispersion.

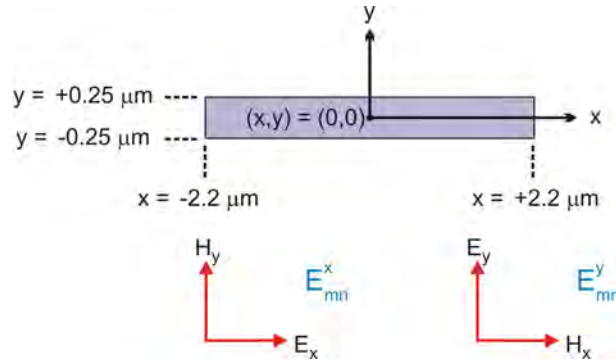


FIGURE 5.3: Waveguide geometry and the detail of polarisation notation

The effective index method (EIM) is an efficient technique to study the mode characteristics of the waveguide used in this research owing to its rectangular geometry. The details of the method, made with enhanced accuracy by eliminating the first-order perturbation error, have been explained in Section 4.3, Chapter 4. The geometry of the waveguide, utilised in the calculation, is shown in Figure 5.3. The x -axis aligns along the width dimension whereas the y -axis is along the height, with the centre of the waveguide as the origin of the coordinate. The location along length of the waveguide is specified by the

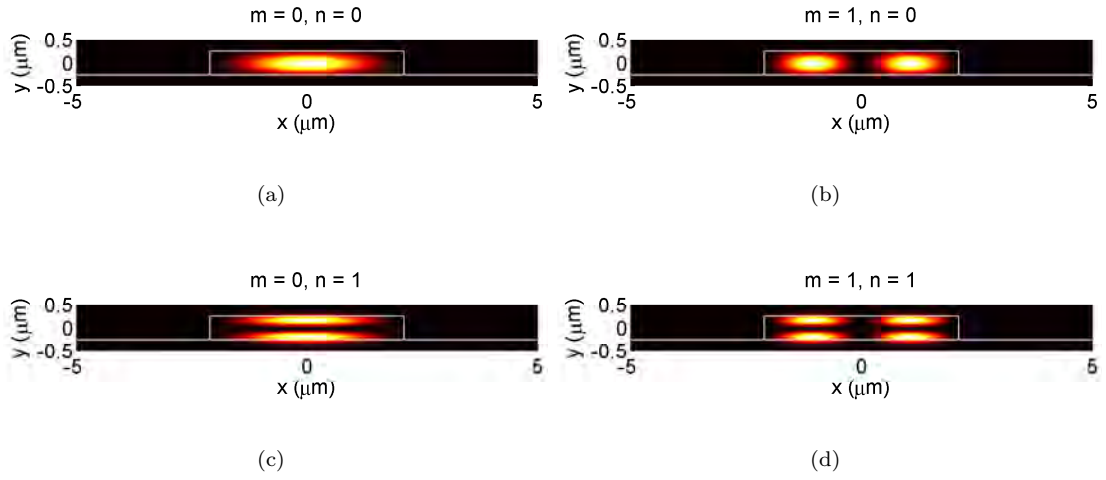


FIGURE 5.4: Calculated intensity distributions for modes with various mode indices m and n . Note the white lines indicating the waveguide profile

z -coordinate. The polarisation notation for a two-dimensional rectangular waveguide is adopted, as already explained in Section 4.3, Chapter 4, as E_{mn}^x and E_{mn}^y (representing the alignment of the electric field along the x - and y -direction respectively) with the two indices m and n related to the number of antinodes in the field distribution along both directions. The examples of the intensity profile of first few modes are shown in Figure 5.4.

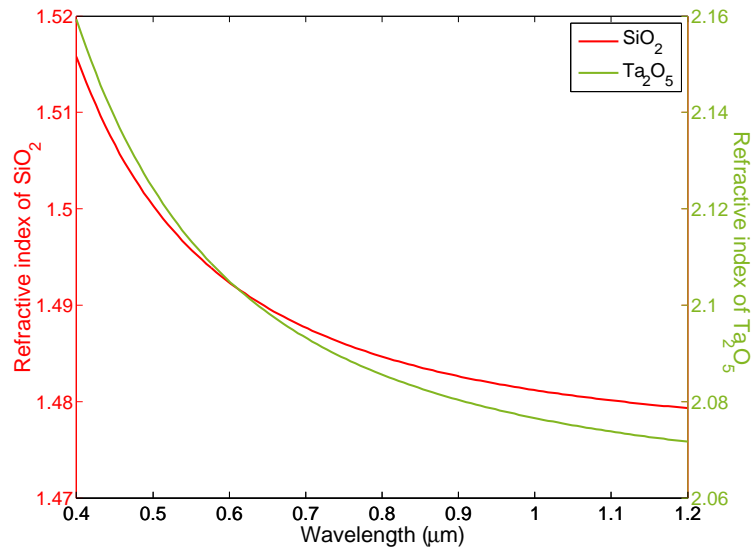


FIGURE 5.5: Refractive index dispersion of Ta_2O_5 and SiO_2 calculated from Equation 5.10 and Equation 5.11 respectively, for the wavelength range 400-1200 nm

The material dispersion of Ta_2O_5 and SiO_2 is also considered in the calculation. The empirical formula for both materials can be found in the literature [8, 9]. The wavelength-dependent refractive index, $n(\lambda)$, of Ta_2O_5 is given by a Cauchy's equation.

$$n_{\text{Ta}_2\text{O}_5}^2(\lambda) = A + \frac{B}{\lambda^2} \quad (5.10)$$

where A and B are 4.2454 and $0.06677 \mu\text{m}^2$ respectively [8]. On the other hand, the refractive index of SiO_2 is calculated from a Sellmeier's equation, which is

$$n_{\text{SiO}_2}^2(\lambda) = 1 + \frac{C\lambda^2}{\lambda^2 - D} \quad (5.11)$$

where C and D are 1.1760 and $0.015 \mu\text{m}^2$ respectively[9]. The plot of both materials' dispersion is shown in Figure 5.5.

The EIM starts its one-dimensional mode calculation firstly on the y -axis. The modal propagation constants are determined from the intercept of the curves plotted from both sides of the transcendental characteristic equations, Equation 4.52a and Equation 4.52c, Chapter 4, against $k_y d$ where k_y is the propagation constant component along the y -axis and d is half the length of the guide's height.

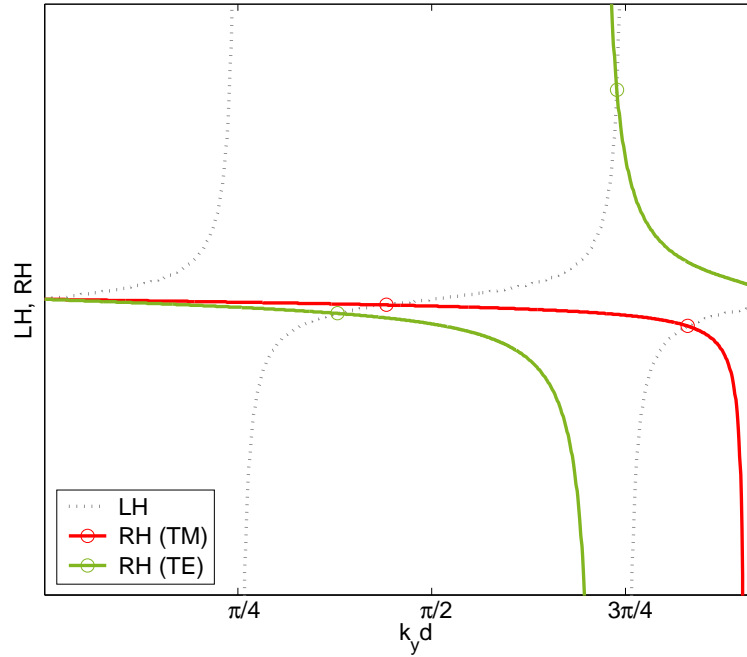


FIGURE 5.6: Mode determination along the y -direction. The left-hand (LH) and right-hand (RH) sides of the characteristic equations in Equation 4.52a and Equation 4.52c.

Figure 5.6 presents the graphical method applied to determine the modes. The notations TE and TM, for \mathbf{E} along the x - and y -directions respectively, are utilised to distinguish the polarisation modes of the planar waveguide from those of the rectangular waveguide. The tangent function of the left-hand side of the characteristic equations in Equation 4.52a and Equation 4.52c is displayed by the grey dotted curve whereas the right-hand side is shown in the red curve for TM polarisation (Equation 4.52c) and the green curves for TE polarisation (Equation 4.52a). Since the guide's height is in

Refractive index	E^y		E^x	
	$n = 0$	$n = 1$	$n = 0$	$n = 1$
Effective index	1.9624	1.6078	1.9957	1.7183
Side cladding index	0.6586	0.4759	0.6671	0.5653

TABLE 5.1: Effective index and side cladding index of the waveguide with the index profile along the x direction. The values are determined from the one-dimensional mode calculation along the y -axis.

the wavelength scale, the structure cannot support many modes but only two modes for each polarisation as identified by the intercept points in the graphs. Those are denoted as symmetric ($n = 0$) and antisymmetric modes ($n = 1$) according to their field distribution along the y axis. The first is the even function without any nodes in the field distribution whereas the latter is the odd function with one node at the centre. The effective index corresponding to these two modes is further used for the next mode calculation along the x -axis as the index of the core. On the other hand, the index of the side cladding layers is adjusted from $n_4 = 1$ to $\sqrt{1 - \zeta_{opt} \left(n_1^2 - (n_{eff}^x)^2 \right)}$, by which means the accuracy of the EIM is enhanced according to the first-order perturbation theory as already described in Section 4.3, Chapter 4. The values of all these refractive indices are displayed in Table 5.1.

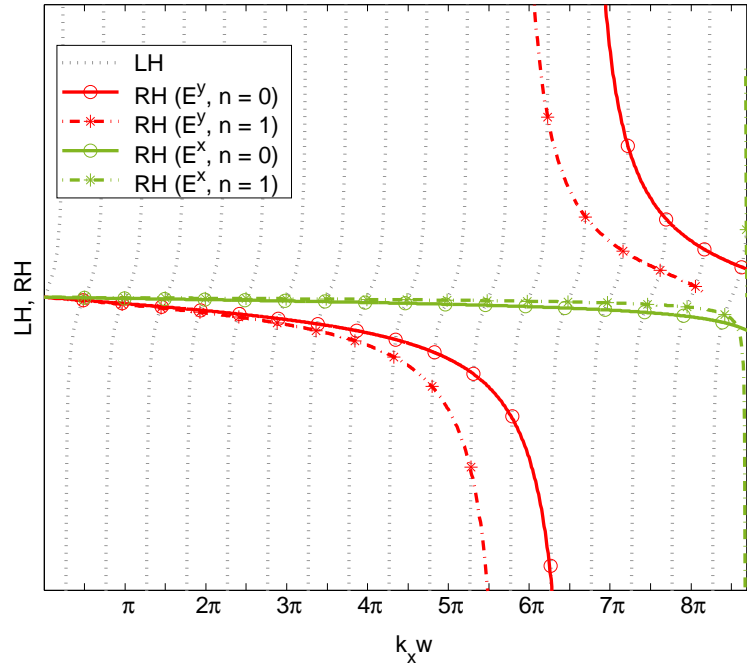


FIGURE 5.7: Mode determination along the x -direction

Since a calculation for modes in a planar waveguide deals with a set of refractive indices for core and cladding, there are two different waveguides and characteristic equations for each kind of polarisation to be calculated for the modes along the x -axis. Similarly the

Mode	$\beta \text{ (m}^{-1}\text{)}$	Mode	$\beta \text{ (m}^{-1}\text{)}$
E_{00}^y	1.5395×10^7	E_{01}^y	1.2607×10^7
E_{10}^y	1.5344×10^7	E_{11}^y	1.2546×10^7
E_{20}^y	1.5259×10^7	E_{21}^y	1.2442×10^7
E_{30}^y	1.5138×10^7	E_{31}^y	1.2295×10^7
E_{40}^y	1.4982×10^7	E_{41}^y	1.2105×10^7
E_{50}^y	1.4789×10^7	E_{51}^y	1.1867×10^7
E_{60}^y	1.4558×10^7	E_{61}^y	1.1581×10^7
E_{70}^y	1.4287×10^7	E_{71}^y	1.1242×10^7
E_{80}^y	1.3974×10^7	E_{81}^y	1.0847×10^7
E_{90}^y	1.3615×10^7	E_{91}^y	1.0387×10^7
$E_{10,0}^y$	1.3209×10^7	$E_{10,1}^y$	0.9856×10^7
$E_{11,0}^y$	1.2750×10^7	$E_{11,1}^y$	0.9226×10^7
$E_{12,0}^y$	1.2232×10^7	$E_{12,1}^y$	0.8530×10^7
$E_{13,0}^y$	1.1652×10^7	$E_{13,1}^y$	0.7692×10^7
$E_{14,0}^y$	1.0995×10^7	$E_{14,1}^y$	0.6688×10^7
$E_{15,0}^y$	1.0252×10^7	$E_{15,1}^y$	0.5439×10^7
$E_{16,0}^y$	0.9403×10^7	$E_{16,1}^y$	0.3777×10^7
$E_{17,0}^y$	0.8425×10^7		
$E_{18,0}^y$	0.7286×10^7		

(a)

TABLE 5.2: Propagation constant β for modes determined from Figure 5.7 at the wavelength 800 nm (a) E^y -polarisation modes (b) E^x -polarisation modes

modes are determined from the intercepts of curves from two sides of the characteristic equations but this time the curves are plotted against $k_x w$ where w is the half length of the waveguide's width and the intercepts come from two pairs of curves relating to different symmetry of modes along the y -axis, as shown in Figure 5.7. Even though there are four tangent function curves, they are not much different and superimpose on one another. Therefore only one set is displayed in the figure. The right side of the equation for E^y (red curves) and E^x (green curves) polarisations are separated into those of symmetric (solid curves) and asymmetric (dash-dot curves) modes with the circles and asterisks marking those modes respectively. Owing to the fact that the waveguide width is around 8 times longer than that the height, a much greater number of modes is present for the planar waveguide along the x -axis. The propagation constants β of

Mode	β (m ⁻¹)	Mode	β (m ⁻¹)
E_{00}^x	1.5657×10^7	E_{01}^x	1.3475×10^7
E_{10}^x	1.5604×10^7	E_{11}^x	1.3413×10^7
E_{20}^x	1.5515×10^7	E_{21}^x	1.3308×10^7
E_{30}^x	1.5390×10^7	E_{31}^x	1.3161×10^7
E_{40}^x	1.5228×10^7	E_{41}^x	1.2970×10^7
E_{50}^x	1.5027×10^7	E_{51}^x	1.2732×10^7
E_{60}^x	1.4787×10^7	E_{61}^x	1.2444×10^7
E_{70}^x	1.4505×10^7	E_{71}^x	1.2105×10^7
E_{80}^x	1.4178×10^7	E_{81}^x	1.1708×10^7
E_{90}^x	1.3805×10^7	E_{91}^x	1.1248×10^7
$E_{10,0}^x$	1.3380×10^7	$E_{10,1}^x$	1.0718×10^7
$E_{11,0}^x$	1.2901×10^7	$E_{11,1}^x$	1.0106×10^7
$E_{12,0}^x$	1.2359×10^7	$E_{12,1}^x$	0.9397×10^7
$E_{13,0}^x$	1.1749×10^7	$E_{13,1}^x$	0.8568×10^7
$E_{14,0}^x$	1.1060×10^7	$E_{14,1}^x$	0.7583×10^7
$E_{15,0}^x$	1.0280×10^7	$E_{15,1}^x$	0.6380×10^7
$E_{16,0}^x$	0.9394×10^7	$E_{16,1}^x$	0.4863×10^7
$E_{17,0}^x$	0.8393×10^7	$E_{17,1}^x$	0.3648×10^7
$E_{18,0}^x$	0.7340×10^7		

(b)

TABLE 5.2: Propagation constant β for modes determined from Figure 5.7 at the wavelength 800 nm (a) E^y -polarisation modes (b) E^x -polarisation modes (cont)

those markers are summarised in Table 5.2. As can be seen from the figure, for each kind of polarisation, the geometry of the waveguide supports around 30 modes at the pump wavelength of 800 nm. Since the wavelength is close to the cut-off wavelength of the antisymmetric modes, which is around 1 μ m, for the index profile along the y direction, the field has greater penetration into the upper and lower cladding layers.

To determine the dispersion characteristic of the modes, the propagation constant must be calculated at various wavelength ranges covering the range of spectra used in study. Figure 5.8 shows the calculation of β over 400-950 nm for some of lower-order modes. As expected from the inversely proportional relation between β and λ , the value declines while the wavelength is increasing. The trends of the curves are parallel with the de-

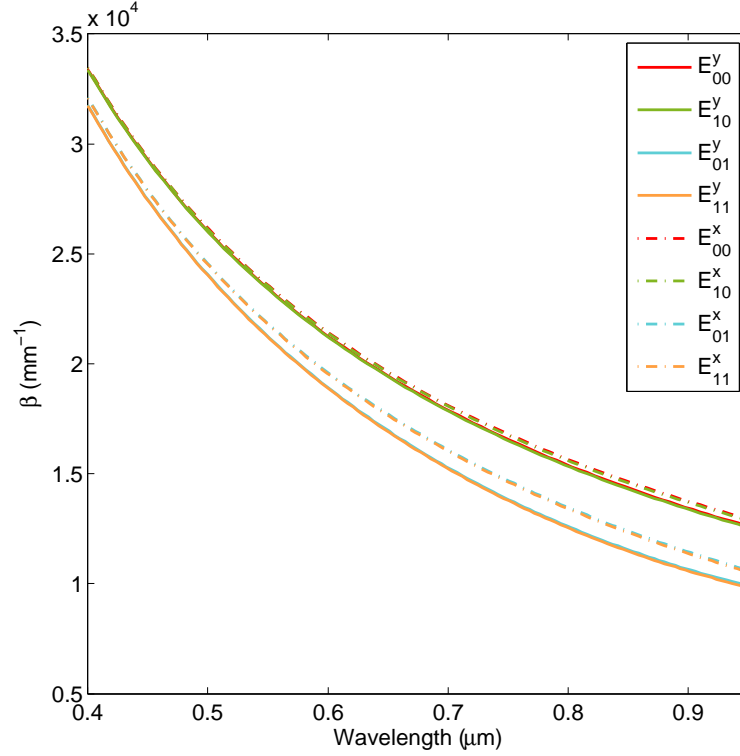


FIGURE 5.8: The variation of propagation constant β for some of the low-order modes calculated by the EIM over the wavelength 400 nm - 950 nm

viation originated by the difference in polarisation and modal symmetry. However, the first accounts for less significant divergence that the latter does.

To confirm the accuracy of the mode calculation programme written in this research, the finite difference method (FDM) from an optical mode solving software, namely CV2's OympIOs, is also utilised to calculate the propagation constant for some modes and the comparison of the values from two techniques is quantified as the percentage difference, shown in Figure 5.9. The curves of the difference between the EIM and FDM, displayed in Figure 5.9(a), do not grow beyond 0.13% for symmetric modes E_{00}^x , E_{10}^x , E_{00}^y and E_{10}^y at the wavelength region of 400 - 950 nm. On the contrary, a much higher value of percentage difference can be viewed in Figure 5.9(b) for asymmetric modes E_{01}^x , E_{11}^x , E_{01}^y and E_{11}^y . This is owing to the fact that the inaccuracy of the EIM increases with the optical power in the cladding area [10, 11] into which the modal field of the asymmetrical modes can penetrate deeper.

Directly differentiating methods may be not suitable to be utilised to determine the higher-order derivatives of β because of the discontinuous nature of the data. Otherwise an uncertainty could arise in the attained values. The process of the calculation begins with determining the value of the propagation constant β at various wavelengths from 400 nm to 950 nm. The data is then interpolated by a polynomial function of the angular frequency (ω) with the degree up to 20. Therefore the propagation constant as a function of frequency, $\beta(\omega)$, can be written as

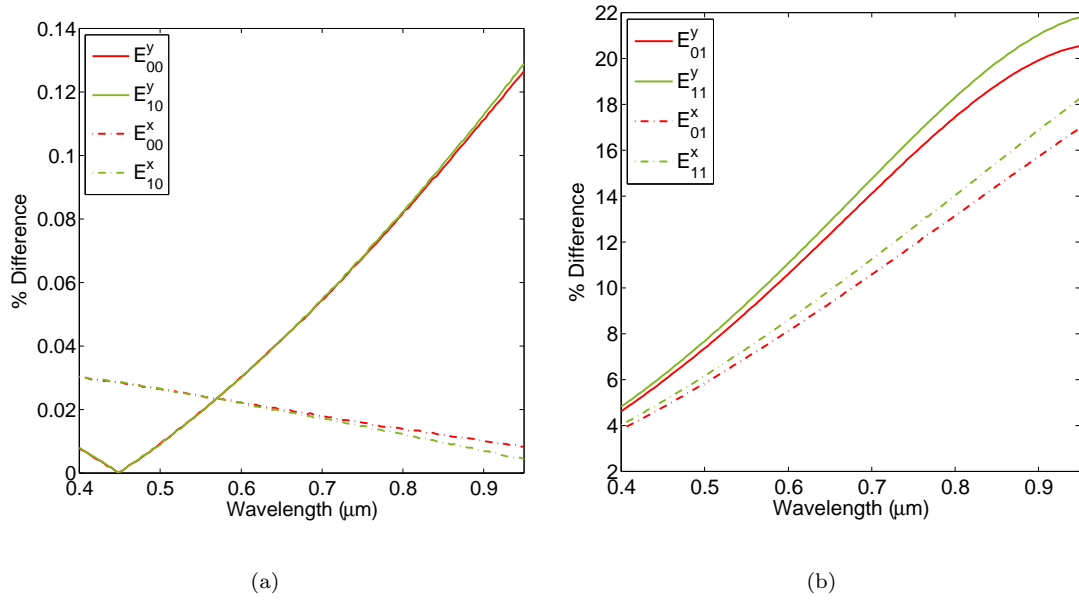


FIGURE 5.9: Percentage difference between the propagation constant, β , calculated by EIM of the mode calculation programme written in this research and FDM of the commercial software OympIOS for (a) symmetric modes and (b) antisymmetric modes

$$\beta(\omega) = \sum_m a_m \omega^m \quad (5.12)$$

where a_m is the function coefficients. From the interpolating function, the n^{th} derivative of $\beta(\omega)$ can be calculated from

$$\beta_n = \frac{d^n \beta(\omega)}{d\omega^n} = \sum_m \frac{m!}{(m-n)!} a_m \omega^{m-n} \quad (5.13)$$

For the sake of clarity, the approach of interpolating the data before determining its derivatives will be referred to later in the text as the *non-direct* method, in contrast to the *direct* method which simply directly differentiates the data of the propagation constant for β_n . The results are displayed in Figure 5.10 showing the first-order, second-order and third-order terms of the Taylor expansion. From the data in the figures, the value of those dispersion parameters at any particular wavelength λ_c is determined from the linear relationship of two adjacent points which have λ_c in-between, as shown in Figure 5.11. Therefore the corresponding value of the dispersion parameters at wavelength λ_c can be calculated from

$$\beta_c = \beta_a + \left(\frac{\beta_b - \beta_a}{\lambda_b - \lambda_a} \right) (\lambda_c - \lambda_a) \quad (5.14)$$

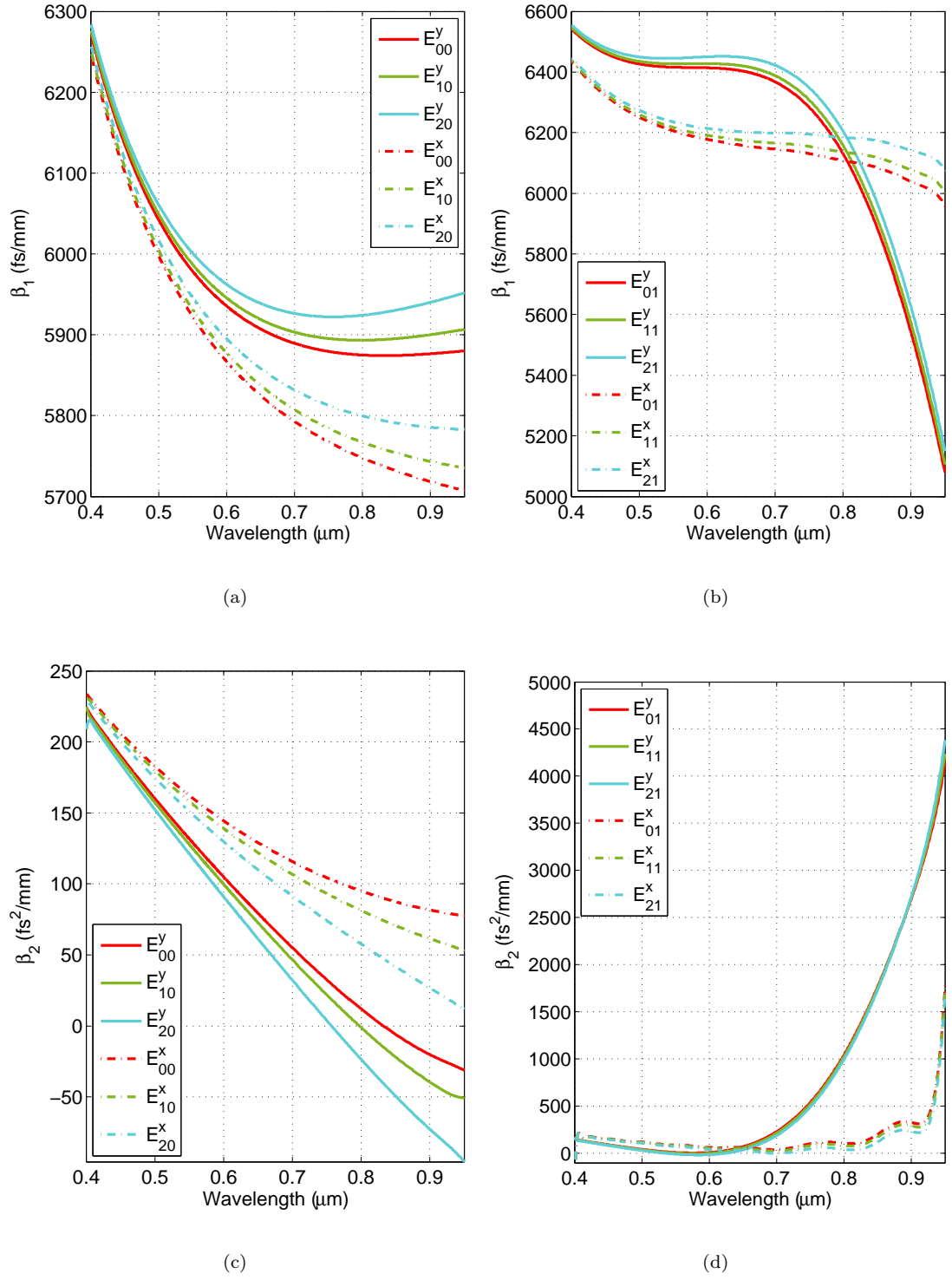


FIGURE 5.10: The dispersion of the propagation constant over the wavelength range 400 - 950 nm. On the left column ((a), (c) and (e)) are for the symmetric modes whereas the asymmetric modes are shown on the right ((b), (d) and (f)). The dispersion of β_1 , β_2 , β_3 are shown in the first row ((a),(b)), second row ((c),(d)) and last row ((e),(f)) respectively.

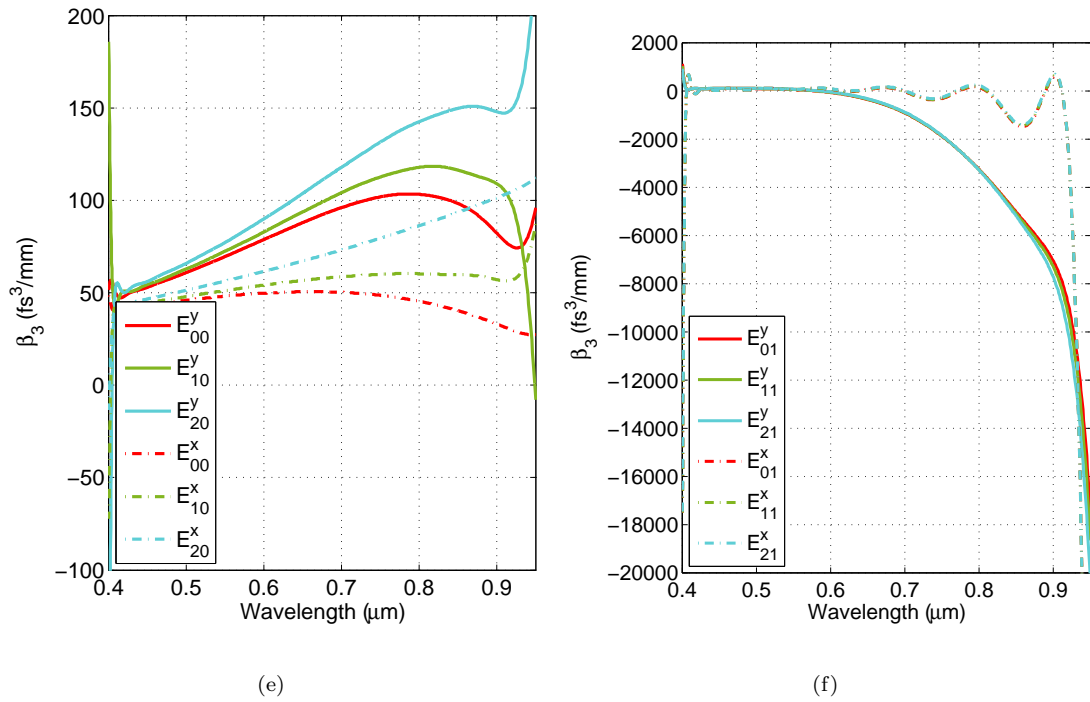


FIGURE 5.10: The dispersion of the propagation constant over the wavelength range 400 - 950 nm. On the left column ((a), (c) and (e)) are for the symmetric modes whereas the asymmetric modes are shown on the right ((b), (d) and (f)). The dispersion of β_1 , β_2 β_3 are shown in the first row ((a),(b)), second row ((c),(d)) and last row ((e),(f)) respectively. (cont)

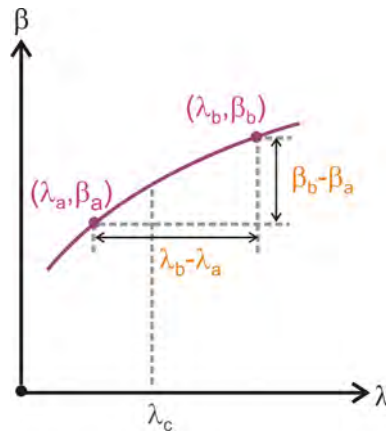


FIGURE 5.11: Evaluation of the values of the dispersion parameters at a particular wavelength

The calculated values of the dispersion constants, at the wavelength 800 nm, are shown in Table 5.3.

Mode	β_1 (fs/mm)	β_2 (fs ² /mm)	β_3 (fs ³ /mm)
E_{00}^y	5.8750×10^3	11.9404	103.2344
E_{10}^y	5.8933×10^3	-1.1554	117.8598
E_{20}^y	5.9242×10^3	-23.5758	142.5681
E_{30}^y	5.9682×10^3	-56.1934	179.7312
E_{40}^y	6.0264×10^3	-100.3087	231.5426
E_{50}^y	6.1000×10^3	-157.8550	301.7020
E_{01}^y	6.1303×10^3	1.0457×10^3	-3.2722×10^3

(a)

Mode	β_1 (fs/mm)	β_2 (fs ² /mm)	β_3 (fs ³ /mm)
E_{00}^x	5.7477×10^3	95.0119	45.3786
E_{10}^x	5.7670×10^3	81.2189	60.2500
E_{20}^x	5.7996×10^3	57.6281	86.2623
E_{30}^x	5.8462×10^3	23.2465	125.0608
E_{40}^x	5.9077×10^3	-23.3822	179.7024
E_{50}^x	5.9855×10^3	-84.4071	253.8276
E_{01}^x	6.1080×10^3	108.3749	134.7366

(b)

TABLE 5.3: The propagation constant and its derivatives at the wavelength 800 nm for (a) E^y -polarisation modes (b) E^x -polarisation modes

The left side of Figure 5.10 is for symmetric modes and the right side for asymmetric modes. The first-order term β_1 , in Figure 5.10(a)-Figure 5.10(d), is inversely proportional to the group velocity of the pulses and therefore provides the information of time delay. The variation of the values among different modes results in the modal separation in time. Figure 5.10(c)-Figure 5.10(d) presents the group velocity dispersion (GVD) parameter, or β_2 , which contributes to the broadening process of the pulse. At the wavelength 800 nm, the waveguide is in the normal dispersion regime for most of the lower-order modes and the dispersion of the symmetric E^y modes becomes anomalous when the mode index increases. The frequency dependency of GVD is connected with

the third-order dispersion (TOD), or β_3 , which is shown in Figure 5.10(e)- Figure 5.10(f).

For a small region of frequency, this effect can be negligible and excluded in the linear operator of the simulation. However, the contribution of TOD to the modeling will be discussed in Section 7.3, Chapter 7. The results in Figure 5.10(e)-Figure 5.10(f) show a kind of calculation instability especially at the border region of the wavelength range. The stability of this calculation approach is shown to be decreasing with the increasing order of derivatives. The accuracy can be enhanced with the increased number of discrete data points in the calculation.

The effect of data size on determining the value of the propagation constant and its derivatives is also studied. Firstly, the value of the propagation constant, β , of E_{00}^y at various wavelengths from 400 nm to 950 nm is determined from the mode calculation algorithm, which was described previously, while varying the parameter of the wavelength data size from 2^5 to 2^{12} . The non-direct method interpolates the data as Equation 5.12 and the higher-order dispersion constants are determined from Equation 5.13. Then the value of the dispersion parameters at the wavelength of 800 nm is determined by Equation 5.14. The effect of the wavelength data size is quantified by calculating the percentage difference between the value determined from utilising different wavelength array data sizes and that derived from the maximum data size of 2^{12} . This percentage difference is plotted against the wavelength data size, as shown in Figure 5.12. The effect of the wavelength data size on the direct method is also included. Direct differentiation is performed by the numerical gradient function of MATLAB. The n^{th} -order derivative comes from applying the gradient function to the $(n-1)^{\text{th}}$ -order one. For the 0^{th} -order derivative, or β_0 , the direct method determines the value directly from the original data of the propagation constant from the mode calculation programme, whereas the non-direct method acquires the value indirectly from the interpolating function in Equation 5.12.

In Figure 5.12(a), the effect of data size is not evident and the deviation is less than 10^{-4} . Since there is still not any differentiating process, the error is directly related to the method of mode determination. There are slight discrepancies between the two curves and these are much reduced when the data size increases. For the deviation of first derivatives, in Figure 5.12(b), the non-direct method gives a value of less than 5% whereas it is in the range of 10^{-3} for the direct method. However, in the case of higher-order derivatives, as can be observed in Figure 5.12(c) and Figure 5.12(d), the deviation of the value for the direct differentiating approach increases enormously compared with that calculated from the non-direct method. As can be seen in all the figures, the deviation of the non-direct method stabilises when the data size is more than 2^{10} . The value should be taken as the lower limit for the data size in the calculation of the dispersion of the propagation constants to guarantee the reliability of the calculated values.

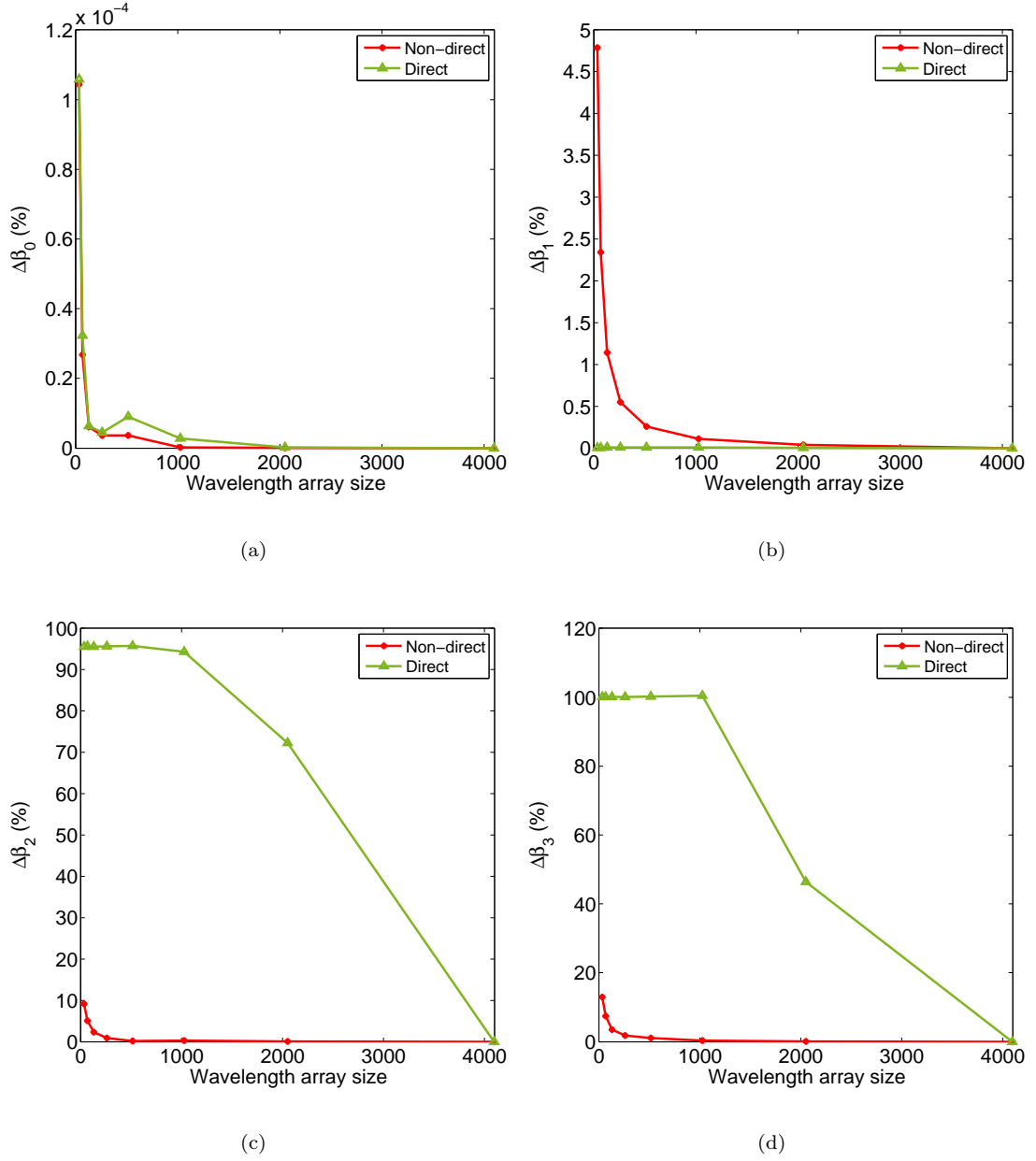


FIGURE 5.12: The effect of wavelength data size on determining (a) the dispersion constant and (b)-(d) its first three orders of derivatives at the wavelength 800 nm for the E_{00}^y mode. The effect is quantified as the percentage difference between the value determined from utilising different wavelength data sizes ranging from 2^5 to 2^{12} and that derived from the maximum data size of 2^{12} . Two methods of determining the derivatives of the dispersion constants can also be compared in the plots: red curves are for the non-direct method whereas green curves are for the direct method.

5.4 Mode transverse field distribution

Each mode calculated in the last section has its own corresponding modal field distribution whose feature is specified by the mode's propagation constants. The transverse field distribution does not play any important role in the single-mode simulation which assumes that the transverse field is constant over the distance of propagation. Therefore the dispersion effect and nonlinear phase modulation is applied to the envelope field only. To calculate the approximated total field experienced by the SNOM probe, the value of the transverse field at each location along the diameter of the probe is multiplied with the envelope field and the outcome is integrated over the diameter. Although the probe can sense the intensity variation across the waveguide, the dominant feature perceived by the probe is from the envelope field which notably develops along the length of the waveguide. On the contrary, when multimode propagation is simulated, the transverse fields of contributing modes strongly participate in defining the spectral feature which this time varies across the width of the waveguide. The details of this will be discussed later in the next chapter. Additionally the field distribution participates in building the intermodal nonlinearity. Therefore, for the simple model of nonlinear multimode propagation, not only self-phase modulation (SPM) but also cross-phase modulation (XPM) are the main nonlinear processes, as will be seen in Section 5.5.

The transverse field can be calculated by solving the Helmholtz equation with the boundary condition requiring the continuity of the field at the top and bottom faces of the waveguide. The derivation of the modal field has been previously discussed in Chapter 4. The transverse fields for some of the first few modes are shown in Figure 5.13. The transverse field is normalised such that it provides a total intensity of 1, $\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy = 1$. This will allow the envelope field to hold the total energy of the pulse and represent itself as the input pulse for the waveguide system in the simulation. The number of nodes observed in each figure corresponds with the indices of the mode label. There are discontinuities at the boundaries along the y -axis for the E^y modes, as can be observed in Figure 5.13(a)-Figure 5.13(c), whereas such discontinuities are not present in the case of E^x in Figure 5.13(d). This is owing to the fact that, for E^y polarisation, the boundary condition requires only the continuity of magnetic field at the top and bottom surfaces of the waveguide but not for the electric field.

The integration of the mode field intensity is related to the nonlinear parameter specifying the degree of nonlinearity, which will be discussed in the next section. Since, in numerical calculations, the indefinite integration of continuous data is replaced by the definite integration of discrete data, the appropriate large number of data points and wide range of x - and y -values are required in order to obtain the accurate approximation of the integration. However, in the case of the guided field, confining itself within the waveguide, the range can be much reduced without inducing much reduction in the accuracy. On the contrary, the widening of the range for the fixed number of data points

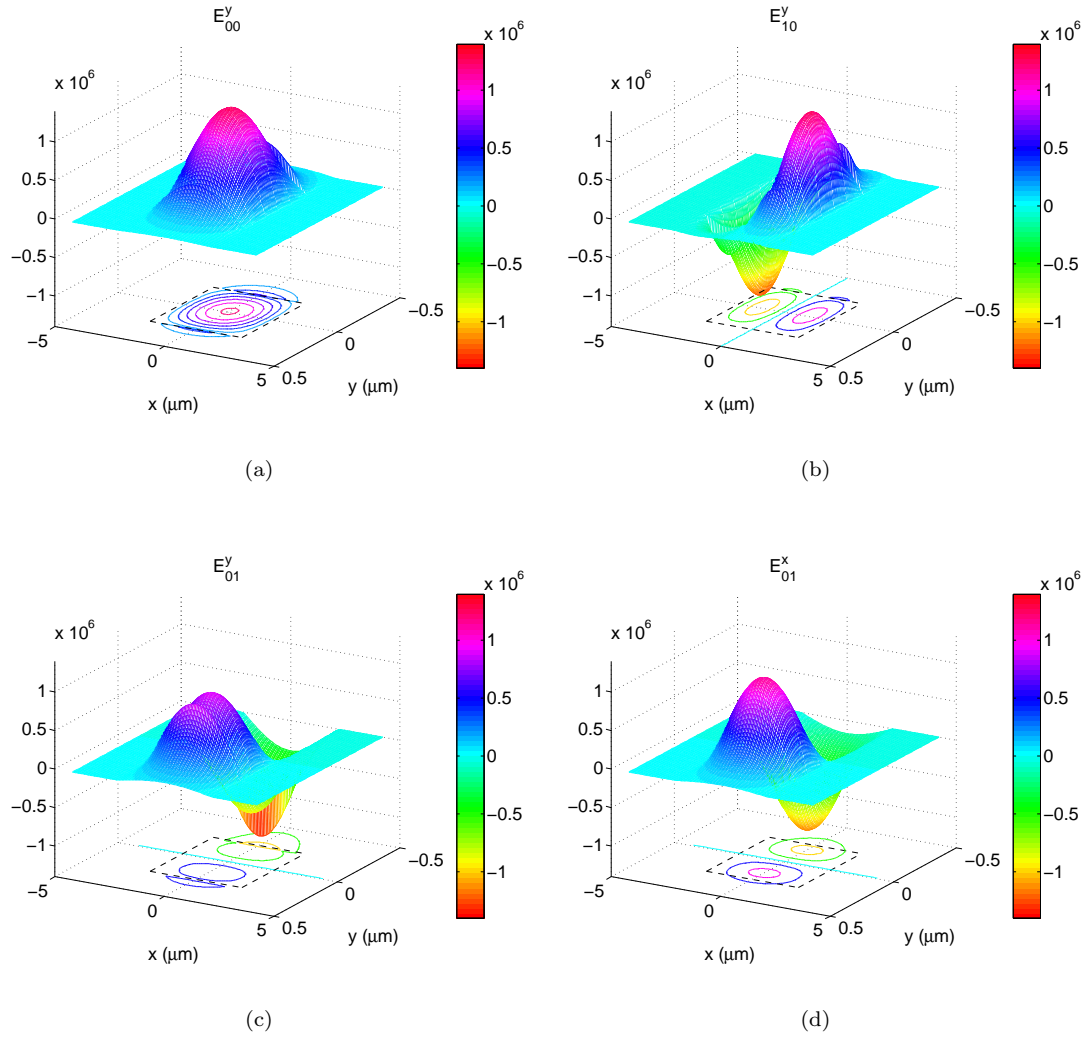


FIGURE 5.13: Normalised field distribution of various propagation modes: (a) E_{00}^y (b) E_{10}^y (c) E_{01}^y and (d) E_{01}^x . The waveguide area is also displayed as a dashed rectangle.

raises the values of Δx and Δy which are the discrete approximation of dx and dy . Consequently the convergence of the discrete integration approaching to the real continuous integration is reduced. The effect of this can be viewed in Figure 5.14, which displays the values of the intensity integration of various modes against the parameter ρ which is the ratio between the relative range of x (or y) and the number of data points.

$$\rho = \frac{L_{x,y}/\ell_{x,y}}{N} \quad (5.15)$$

where $L_{x,y}$ is the overall range of x (or y) value, $\ell_{x,y}$ is the dimensional length of the waveguide along the x (or y) direction, and N is the number of data points. The parameter ρ is directly proportional to Δx and Δy which are $L_{x,y}/N$. Therefore the larger the value the lower is the accuracy of the intensity integration. The significant

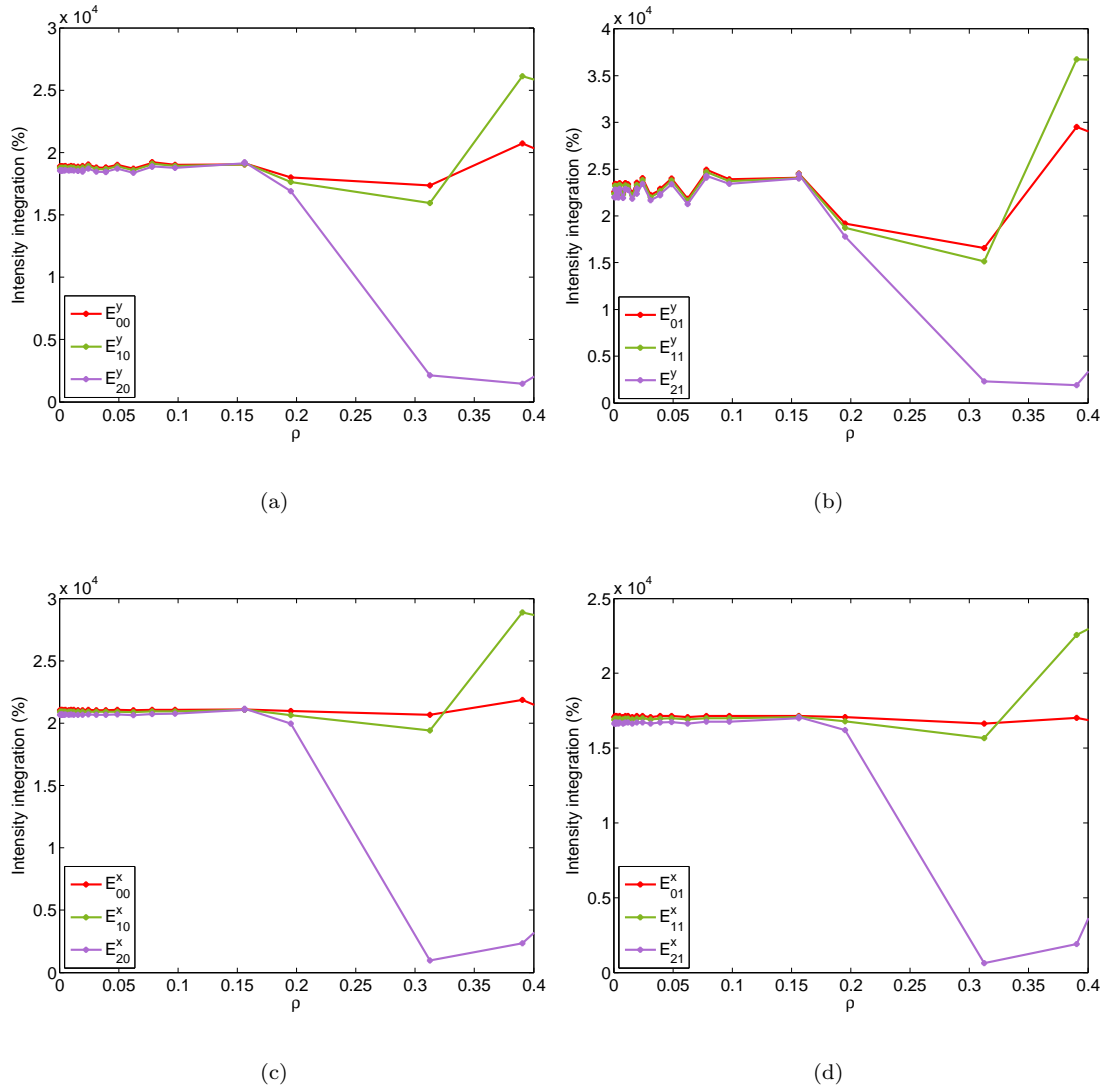


FIGURE 5.14: Intensity integration $\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy$ against the parameter ρ which is the ratio between the relative range of x (or y) and the number of data points, as defined in Equation 5.15: (a) symmetric E^y -polarisation modes (b) asymmetric E^y -polarisation modes (c) symmetric E^x -polarisation modes (d) asymmetric E^x -polarisation modes

deviation of the integration values occurs when the ratio is, approximately, more than 0.15 and therefore this is taken as the limit for selecting the value range and the number of data points.

The effect of the x and y data size on the field intensity integration is shown in Figure 5.15. The range of x and y is chosen to be 5 times wider than the width and height dimensions of the waveguide respectively. The displayed value is the percentage deviation of the intensity integration $\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy$ calculated from other data sizes compared with that generated by the maximum data size of $2^{12} \times 2^{12}$. As can be seen for both polarisation and mode symmetry, the highly abrupt changes of the value at

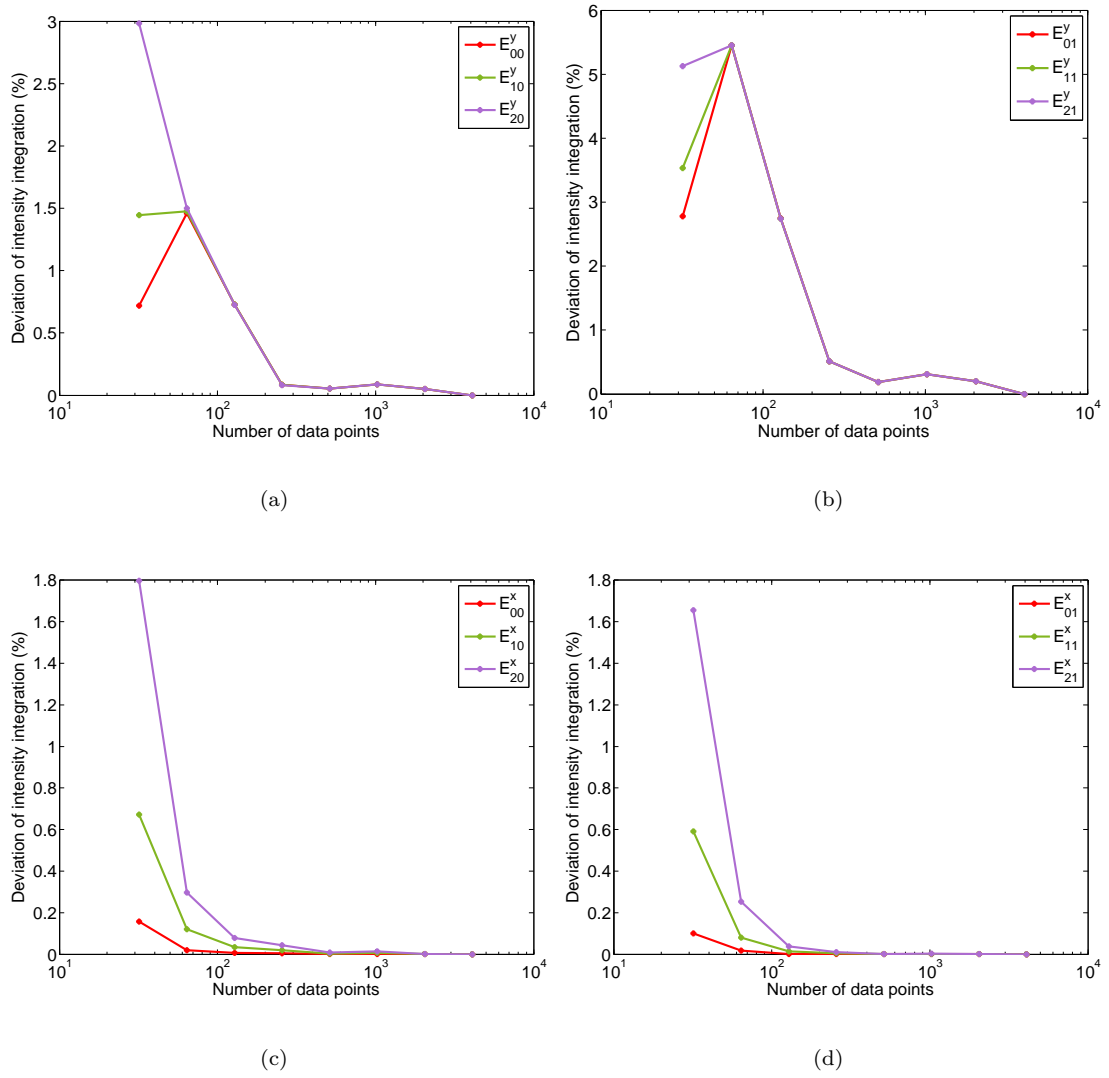


FIGURE 5.15: Deviation of $\iint_{-\infty}^{+\infty} |U(x, y)|^2 dx dy$, calculated by utilising various numbers of data points, from that generated by the maximum data size of $2^{12} \times 2^{12}$ (a) symmetric E^y -polarisation modes (b) asymmetric E^y -polarisation modes (c) symmetric E^x -polarisation modes (d) asymmetric E^x -polarisation modes

the small number of data points starts to stabilise when the field is calculated from the matrix size chosen to be more than $2^8 \times 2^8$.

5.5 NLS equation for multimode pulses

In order to better describe the characteristics of nonlinear pulses propagating in a nonlinear multimode waveguide in this research, the NLS equation whose theoretical derivation has been already explained in Section 4.4, Chapter 4, is required to be adapted for multimode pulses. In contrast to the single mode case, the nonlinear phase modulation

occurring in each mode is not only caused by the intensity distribution of the mode itself but also that of all contributing modes. The additional cross-phase modulation term can be added to the nonlinear propagation equation as [12]

$$\begin{aligned} \frac{\partial V^{(p)}(z, t)}{\partial z} = & \left(\beta_1^{(1)} - \beta_1^{(p)} \right) \frac{\partial V^{(p)}(z, t)}{\partial t} - \frac{j\beta_2^{(p)}}{2} \frac{\partial^2 V^{(p)}(z, t)}{\partial t^2} \\ & + j \left[\sum_{q=1}^N \gamma^{(p)(q)} \left| V^{(q)}(z, t) \right|^2 \right] V^{(p)}(z, t) \end{aligned} \quad (5.16)$$

where $V^{(p)}(z, t)$, $\beta_1^{(p)}$ and $\beta_2^{(p)}$ are the slowly varying field envelope, group velocity parameter, and GVD parameter for mode p , respectively. The frame of reference is chosen such that it moves with the lowest-order mode. Therefore the first term at the right side of Equation 5.16 is related to the relative phase change of each mode to the fundamental mode E_{00}^x (or E_{00}^y), caused by their different group velocities, and the term disappears when the mode is the fundamental mode or $p = 1$. The second term results in the group velocity dispersion of each mode in the pulse. Any inaccuracy introduced by the truncation of high-order dispersion terms will be further discussed in Chapter 7.

The nonlinear term in Equation 5.16 retains only self- and cross-phase modulations, which are shown to be the most important nonlinear effects in this research system. Other nonlinearities such as the Raman effect and four-wave mixing are neglected since the input power in the study is far below the Raman threshold, and the phase-matching condition is not satisfied in the waveguide. The quantity of both nonlinear effects depends on the nonlinear refractive index n_2 of the guide material and the overlapping integrals of the transverse mode intensities, which is parameterised as the nonlinear parameter $\gamma^{(p)(q)}$, which can be formulated by [12].

$$\gamma^{(p)(q)} = n_2 k_0 h^{(p)(q)} \frac{\int \int_{-\infty}^{+\infty} |U^{(p)}(x, y)|^2 |U^{(q)}(x, y)|^2 dx dy}{\left(\int \int_{-\infty}^{+\infty} |U^{(p)}(x, y)|^2 dx dy \right) \left(\int \int_{-\infty}^{+\infty} |U^{(q)}(x, y)|^2 dx dy \right)} \quad (5.17)$$

The value of n_2 for Ta₂O₅, which is experimentally determined in the literature [13] and based on the similar waveguide's structure, is $7.23 \times 10^{-19} \text{ m}^2/\text{W}$. The constant $h^{(p)(q)}$ in Equation 5.17 is 1 for $p = q$ and 2 for $p \neq q$. This gives the integral term the value of each mode's effective area and the intermodal effective area respectively. The field is calculated over the area $x = -10.50 \text{ } \mu\text{m}$ to $x = +10.50 \text{ } \mu\text{m}$ and $y = -1.25 \text{ } \mu\text{m}$ to $y = +1.25 \text{ } \mu\text{m}$. The discrete integration is done over the number of data points $2^{10} \times 2^{10}$. The calculated values of $\gamma^{(p)(q)}$ for E^x and E^y polarisations are shown in Table 5.4, in which each value in a cell is for the contributing modes in the corresponding row and

column. The values are in the range $3\text{--}5 \text{ W}^{-1}/\text{m}$ for SPM and in the range $2\text{--}6 \text{ W}^{-1}/\text{m}$ for XPM. The parameters in are for SPM which decrease with the increase of the mode index. Those in are the XPM parameter for modes having the same symmetry along the y -axis and is for the modes containing the same mode index along the x -axis, whereas is for the modes lacking in similarity in the mode field in both directions. Obviously the XPM parameters depend on the degree of overlapping of contributing mode fields.

Mode	E_{00}^x	E_{10}^x	E_{20}^x	E_{30}^x	E_{01}^x	E_{11}^x	E_{21}^x	E_{31}^x
E_{00}^x	4.4614	5.9470	5.9444	5.9407	5.3864	3.6020	3.5989	3.5976
E_{10}^x		4.4593	5.9430	5.9393	3.5787	5.3848	3.6071	3.6004
E_{20}^x			4.4557	5.9368	3.5800	3.5705	5.3820	3.6117
E_{30}^x				4.4505	3.5786	3.5747	3.5610	5.3778
E_{01}^x					3.8610	5.1477	5.1472	5.1465
E_{11}^x						3.8606	5.1469	5.1462
E_{21}^x							3.860	5.146
E_{31}^x								3.85907

(a)

Mode	E_{00}^y	E_{10}^y	E_{20}^y	E_{30}^y	E_{01}^y	E_{11}^y	E_{21}^y	E_{31}^y
E_{00}^y	4.6011	6.1339	6.1325	6.1278	3.7559	2.4934	2.4951	2.4946
E_{10}^y		4.6001	6.1317	6.1297	2.5135	3.7548	2.4865	2.4910
E_{20}^y			4.5984	6.1283	2.5105	2.5180	3.7528	2.4785
E_{30}^y				4.5959	2.5090	2.5117	2.5222	3.7499
E_{01}^y					3.0272	4.0354	4.0340	4.0321
E_{11}^y						3.0262	4.0332	4.0313
E_{21}^y							3.0245	4.0299
E_{31}^y								3.0221

(b)

TABLE 5.4: Nonlinear parameter $\gamma^{(p)(q)}$ for (a) E^x -polarisation modes, and (b) E^y -polarisation modes. Each value in a cell is determined from Equation 5.17 by applying the transverse field of the modes in the corresponding row and column for $U^{(p)}(x, y)$ and $U^{(q)}(x, y)$

The numerical simulation of pulse propagation for each mode, based on Equation 5.16, is performed individually by SSFM. The outcome is each mode's envelope field $V^{(p)}(z, \omega)$.

By assuming that there are insignificant changes in the mode transverse field along the propagation distance, the total field perceived by the SNOM probe can be determined from

$$E(x, y, z, \omega) = \sum_p \tilde{V}^{(p)}(z, \omega) U^{(p)}(x, y) \exp\left(j \left[\beta_0^{(p)} + \beta_1^{(1)} \omega\right] z\right) \quad (5.18)$$

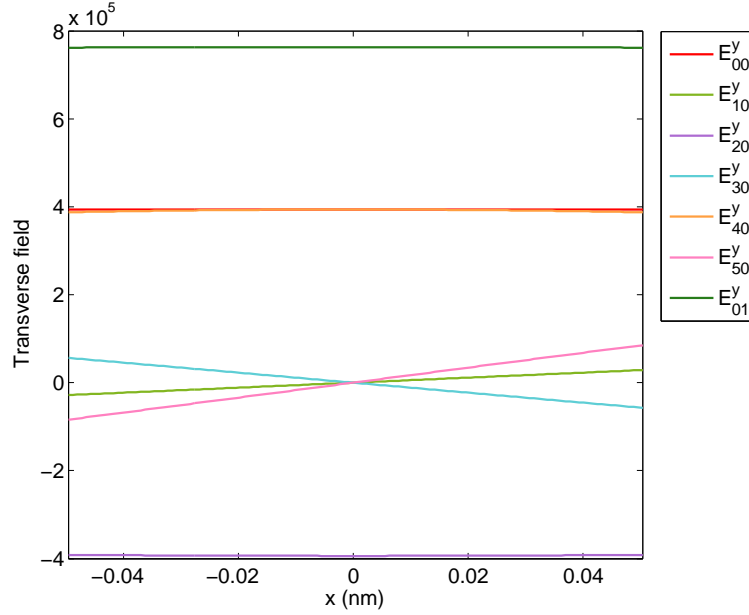


FIGURE 5.16: Transverse field distribution of various modes along the diameter of the SNOM probe, which is placed at the central position of the waveguide and 20 nm above its top surface

where $U^{(p)}(x, y)$ is the mode transverse field, which is normalised in such a way that the total input power is distributed among the mode envelope fields $V^{(p)}$ with different contributing ratios. The result of Equation 5.18 is not the total output from the waveguide but the spectra collected by the SNOM probe and therefore the transverse field $U^{(p)}(x, y)$ is a part of the total transverse field intercepted by the probe area. Figure 5.16 displays the transverse field distribution of various modes along the 100 nm diameter of the SNOM probe placed at the central position of the waveguide with the gap distance 20 nm from the waveguide's top surface. The value of the field is normalised for the unity transverse field integration. The modes E_{10}^y , E_{30}^y , and E_{50}^y have an asymmetric field distribution around the central axis of the waveguide $x = 0$. Therefore these modes would disappear in the time profile observed by the probe. On the other hand, the field distribution of the modes E_{00}^y , E_{20}^y and E_{40}^y is symmetric around $x = 0$ and the sign of the value alternates back and forth between positive for every mode index $m = 4(p - 1)$ and negative for every mode index $m = 4(p - \frac{1}{2})$ where $p = 1, 2, 3$, and so on. The higher value of the transverse field of the mode E_{01}^y is due to the deep penetration of the field since the modes with the mode index $n = 1$ are nearly the cut-off mode for the operation wavelength of 800 nm.

The exponential phase term in Equation 5.18 consists of the general propagation phase shift $\exp(j\beta_0^{(p)}z)$ and the compensating phase shift $\exp(j\beta_1^{(1)}\omega z)$ for the relative transformation with the fundamental mode. In contrast to the transverse field, the envelope field is complex and can be written as

$$\tilde{V}^{(p)}(z, \omega) = \left| \tilde{V}^{(p)}(z, \omega) \right| \exp(j\varphi^{(p)}(z, \omega)) \quad (5.19)$$

where $(\varphi^{(p)}(z, \omega))$ is the spectral phase of the envelope field which is formed by the interplay of the dispersion and nonlinearity in the waveguide.

The intensity observed by the probe can be determined from the multiplication of the field in Equation 5.18 with its conjugate and the result can be written as

$$\begin{aligned} |E(x, y, z, \omega)|^2 &= \sum_p \left| \tilde{V}^{(p)}(z, \omega) \right|^2 \left(U^{(p)}(x, y) \right)^2 \\ &\quad + 2 \sum_p \sum_{\substack{q \\ q \neq p}} \left| \tilde{V}^{(p)}(z, \omega) \right| \left| \tilde{V}^{(q)}(z, \omega) \right| U^{(p)}(x, y) U^{(q)}(x, y) \\ &\quad \times \cos \left(\Delta\phi^{(p)(q)}(z, \omega) \right) \end{aligned} \quad (5.20)$$

with

$$\Delta\phi^{(p)(q)}(z, \omega) = \left(\beta_0^{(p)} - \beta_0^{(q)} \right) z + \left(\varphi^{(p)}(z, \omega) - \varphi^{(q)}(z, \omega) \right) \quad (5.21)$$

The first term on the right side of Equation 5.20 is the intensity of each mode whereas the second term originates from the mode interference. The variation of the spectral intensity is modulated by the factor $\Delta\phi^{(p)(q)}(z, \omega)$ which consists of the propagation phase difference and the envelope field's phase difference. As can be directly deduced from Equation 5.20, the greater the number of modes the higher the complexity of the spectral modulation.

However, what the SNOM system really detects is the intensity of the guided field from the fibre probe. Hence, the calculation for the total intensity above is a simple approximation which neglects the coupling of the near-field into the SNOM probe. With some approximations, the coupling between the evanescent field and the fibre mode field can be simply represented by the scalar amplitude coupling function [14]. The coupling function becomes constant in the case of detecting the field from waveguides because the spatial spectra of the field are narrow. Therefore the SNOM signal quite faithfully represents the intensity distribution of the guided field, as evident in previous experiments [15–17]. In Balistreri et al.'s work [18, 19], the SNOM technique was found to introduce

the polarisation conversion to the detected field and consequently the interference pattern, which is caused by two orthogonal waveguide modes, can be observed. However, the measured beat length agrees well with the calculated value of the propagation modes inside the waveguide.

The results of the multimode simulation will be presented and discussed together with the experimental results in the following chapters.

5.6 Enhancement of numerical simulation

5.6.1 The effect of step size

The accuracy of the SSF scheme depends on the step size. By intuition, the smaller the step the better is the accuracy of the calculation expected. However, it comes with the expense of the computational time, since the number of steps grows with the reduction of the step size. Therefore the trade-off can be optimised by analysing the error trend in response to different values of step size.

The total error of the simulation is the combination of the round-off and the truncation errors, whose definition is already given in Section 5.2.2. For the second-order approximation, the total error ϵ over all the calculation span can be written as

$$\epsilon \sim \frac{\tilde{\epsilon}}{h} + h^2 \quad (5.22)$$

From Equation 5.22, the minimum error ϵ_{min} occurs at the optimum value h_{opt} of the step size, which is approximately equal to

$$h_{opt} \sim \left(\frac{\tilde{\epsilon}}{2} \right)^{\frac{1}{3}} \quad (5.23)$$

$$\sim 10^{-6} \quad (5.24)$$

This provides the error $\epsilon_{min} \sim 3 \left(\frac{\tilde{\epsilon}}{2} \right)^{\frac{2}{3}}$ or $\epsilon_{min} \sim 10^{-12}$. Reducing the step size to less than h_{opt} does not give more accuracy, because the simulation is subject to the round-off error, but increases the calculation time instead. On the contrary, the truncation error becomes prominent when the step size is increased to more than h_{opt} . This behaviour is confirmed by Figure 5.17 which displays the effect of step size on the simulation error of the single-mode (E_{00}^y) propagation in the waveguide over the distance of 6 mm. The relative error, shown in the graph, is calculated from

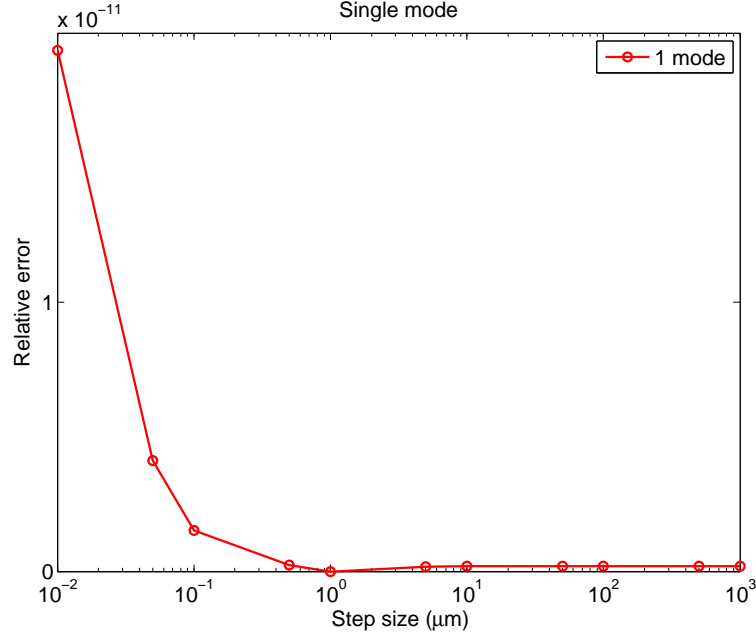


FIGURE 5.17: Relative error from single-mode (E_{00}^y) simulations calculated by using different values of step size

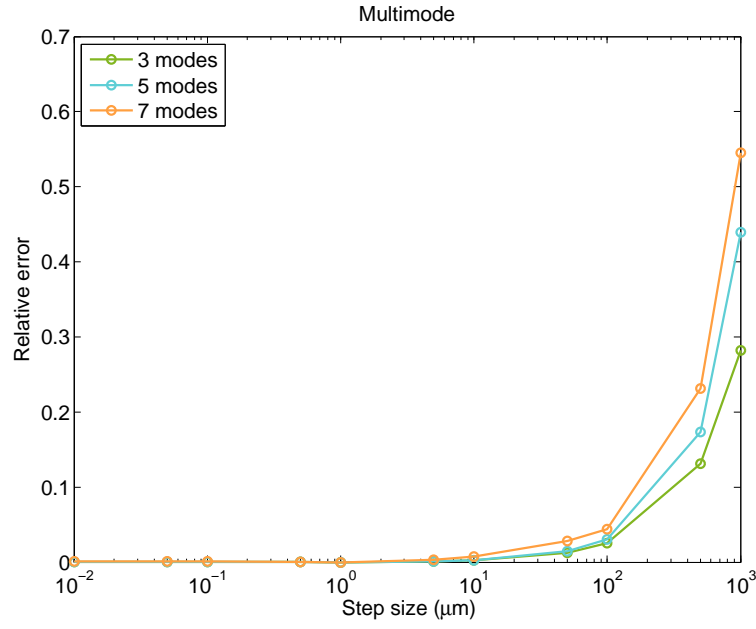
$$\epsilon = \frac{\|V_a - V_r\|}{\|V_r\|} \quad (5.25)$$

where the operator $\|V\|$ is defined as

$$\|V\| = \left(\int_{-\infty}^{+\infty} |V|^2 d\omega \right)^{1/2} \quad (5.26)$$

Since there is no available exact solution, the spectrum V_r , simulated by using the step size $1 \mu\text{m}$, is adopted as the reference spectrum to be compared with the spectrum V_a calculated by using other values of step size. From the figure, the higher degree of deviation of the error is present before the $1 \mu\text{m}$ point. The inverse proportion to the step size suggests the predominance of the round-off error over the truncation error. However when the step size is increased to more than $1 \mu\text{m}$ the degree of truncation error becomes able to compensate the round-off error and results in the plateau after this point. The inconspicuous growth of the truncation error with the step size indicates that the linear and nonlinear effects are still not strong enough over the length range displayed in Figure 5.17 to introduce significant changes in the pulse shape.

Similarly the relative error from various multimode simulations, which is calculated by the same criteria as previously applied to the single-mode case, is plotted against the step size and displayed in Figure 5.18. The simulations of 3-mode, 5-mode and 7-mode pulse propagation are represented by the green, blue and orange curves respectively. The details of contributing modes and their relative intensity ratio, displayed in the table,



(a)

No of Modes	Relative intensity ratio						
	E_{00}^y	E_{10}^y	E_{20}^y	E_{30}^y	E_{40}^y	E_{50}^y	E_{01}^y
3 modes	0.7	0.15	0.15	—	—	—	—
5 modes	0.7	0.1	0.1	0.05	0.05	—	—
7 modes	0.7	0.05	0.05	0.05	0.05	0.05	0.05

(b)

FIGURE 5.18: (a) Relative error from multimode simulations calculated by using different values of step size (b) Details of contributing modes and the relative intensity ratio

are the same as that of the simulation result which will be discussed in Section 6.3, Chapter 6. In contrast to the case of single mode, the truncation error in a multimode simulation grows much more rapidly with the step size and completely suppresses the effect of round-off error. The rise of the error when the step size is reduced, as occurring in Figure 5.17, is not noticeable anymore. As predictable, the error also increases when there are more modes contributing.

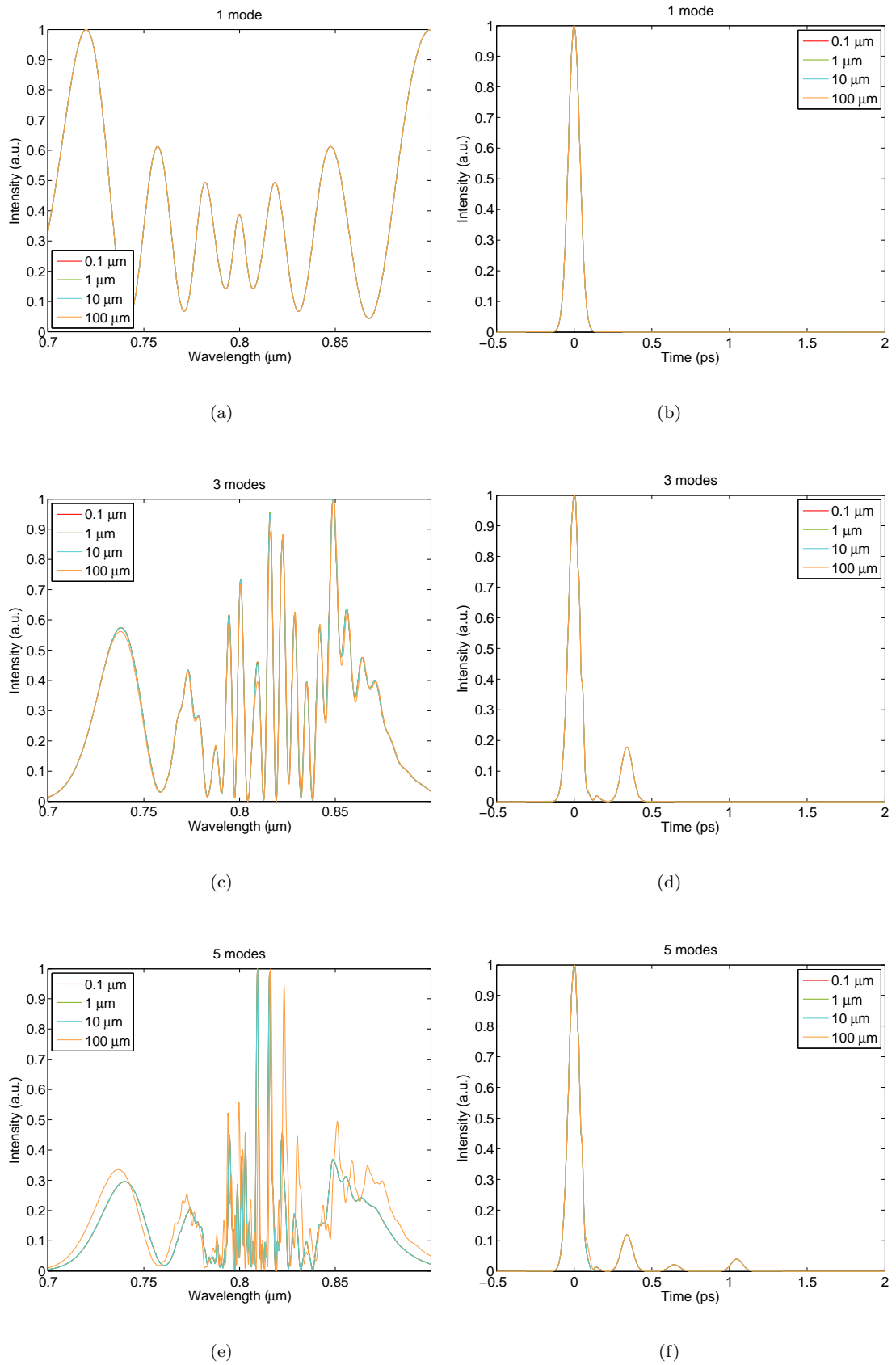


FIGURE 5.19: Spectra and time profiles through the probe simulated from different step sizes, shown in the inset box. (a),(b) Single-mode simulation (c),(d) 3-mode simulation (e),(f) 5-mode simulation. Note that the deviation of the spectra in (a) and time profiles in (b), (d) and (f) for various step sizes is hardly noticeable and the curves are quite superimposed

The variation of simulation results caused by different step sizes is shown in Figure 5.19. The simulation is performed for single-mode, 3-mode and 5-mode propagation over the distance of 6 mm with the same details of the contributing modes and the relative intensity ratio as those in Figure 5.18. The step size extends over four orders of magnitude from $10^{-1} \mu\text{m}$ to $10^2 \mu\text{m}$. Spectra through the probe (Figure 5.19(a), Figure 5.19(c), and Figure 5.19(e)) are shown on the left whereas their corresponding time profiles (Figure 5.19(b), Figure 5.19(d), and Figure 5.19(f)) are on the right. In the case of single-mode simulation, there is no observable deviation in the simulation results generated from the step size less than $10^2 \mu\text{m}$. However, the deviation clearly appears in the multimode simulation and its degree is sensitive to the number of modes, as can be observed in Figure 5.19(c) and Figure 5.19(e). In contrast to the spectra, the time profile is much less sensitive to the step size over the range investigated.

The total error of the simulation can be further reduced by improving the approximation of the NLS equation to a higher-order approximation which decreases the truncation error. This issue will be discussed later in the next subsection.

5.6.2 Higher-order approximation

The scheme of higher-order approximation for the NLS equation is based on the numerical techniques such as the Runge-Kutta method for solving ordinary differential equations (ODE). Higher-order approximation is carried out by combining the results calculated from lower-order approximations with full and half step sizes. If the combination is done with the proper coefficients, the local error can be canceled out and the next higher-order error becomes the most significant error instead [20]. For example, the third-order approximation $V^{(3)}(z+h, t)$ for the envelope field can be found by calculating two 2nd-order solutions $V_h^{(2)}(z+h, t)$ and $V_{h/2}^{(2)}(z+h, t)$. The first is calculated from Equation 5.9 by using the step size h , whereas the step size $h/2$ is used for the latter with twice the calculations in order to fulfil the same amount of propagation distance h . If $V(z+h, t)$ is the exact solution, those approximations can be written as

$$V_h^{(2)}(z+h, t) = V(z+h, t) + \rho h^3 \quad (5.27)$$

$$V_{h/2}^{(2)}(z+h, t) = V(z+h, t) + 2\rho \left(\frac{h}{2}\right)^3 \quad (5.28)$$

If Equation 5.27 and Equation 5.28 are combined such that

$$V^{(3)}(z+h, t) = a_1 V_h^{(2)}(z+h, t) + a_2 V_{h/2}^{(2)}(z+h, t) = V(z+h, t) + O(h^4) \quad (5.29)$$

the coefficients a_1 and a_2 must satisfy

$$a_1 + a_2 = 1 \quad (5.30a)$$

$$a_1 + \frac{a_2}{4} = 0 \quad (5.30b)$$

The coupled equations Equation 5.30 can be arithmetically solved and the values of the coefficients are $a_1 = -\frac{1}{3}$ and $a_2 = \frac{4}{3}$. Therefore the third-order approximation can be rewritten as [20]

$$V^{(3)}(z+h, t) = \frac{4}{3}V_{h/2}^{(2)}(z+h, t) - \frac{1}{3}V_h^{(2)}(z+h, t) \quad (5.31)$$

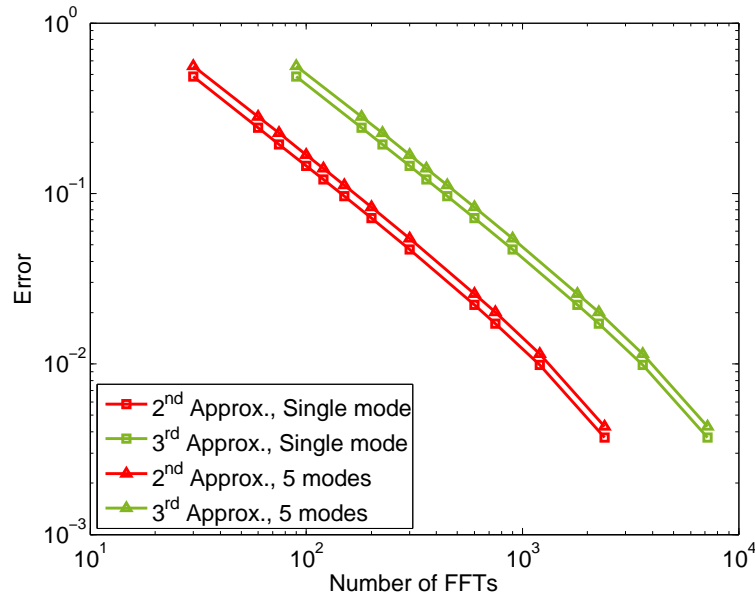


FIGURE 5.20: Comparison of the efficiency of 2nd-order and 3rd-order algorithms

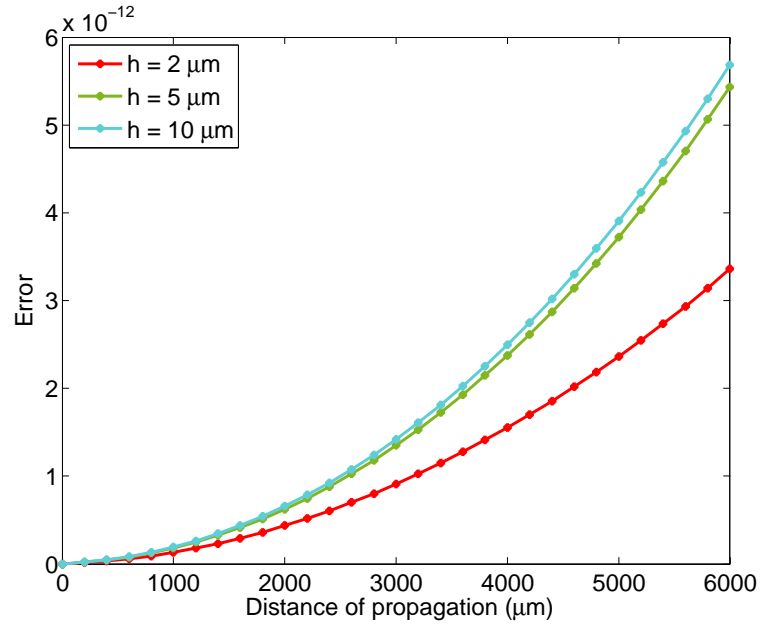
Although Equation 5.31 provides a higher accuracy than Equation 5.9 by one order, the calculation cost is tripled owing to additional calculation for $V_{h/2}^{(2)}(z+h, t)$. The justification of numerical efficiency for both algorithms should be based on the comparison of the accuracy against the cost of the calculation, as pictured in Figure 5.20. The graph expresses the calculation cost in terms of the number of fast Fourier transforms (FFTs) required in each simulation whereas the relative error is calculated by the same error formula, Equation 5.25, previously used. The reference field is from the 2nd-order approximation in Equation 5.28 with the step size $1 \mu\text{m}$, which has been proved to be the optimum value providing the accuracy at the level of machine precision. The simulations were carried out on the single-mode (square) and 5-mode (triangle) pulse propagation

over a distance of 6 mm in the waveguide with the red line from the 2nd-order approximation in Equation 5.28 and the green line from the 3rd-order approximation in Equation 5.31. The single-mode pulse is the E_{00}^y mode whereas the 5-mode pulse is the product of the mixing of E_{00}^y , E_{10}^y , E_{20}^y , E_{30}^y , and E_{40}^y by the relative intensity ratio 0.7:0.1:0.1:0.05:0.05. The plot clearly shows that the degree of accuracy from the 3rd-order approximations for both the single-mode and 5-mode cases is not high enough to compensate for the increased expense. Therefore the 2nd-order algorithm is more efficient and this was used by all the simulations in this research.

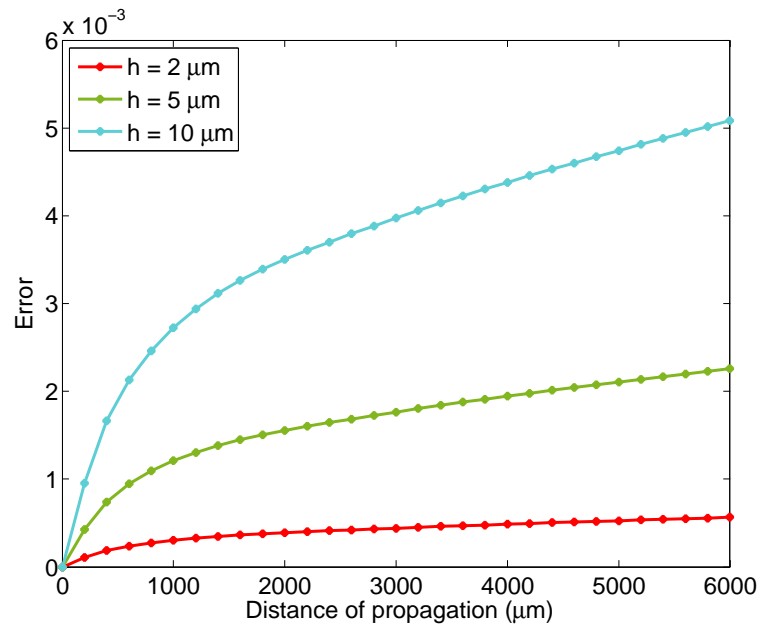
5.6.3 Adaptive step size

The adaptive step size is a numerical technique that can enhance the computation efficiency by investigating the local error occurring at each step of the simulation and adjusting the step size for the next step. The simulation algorithm which has been discussed so far uses a fixed step size in the simulation. Such a scheme accumulates the error generated in each step and therefore the total error increases with the number of the simulation steps. The accumulation of the error can be observed in Figure 5.21 which displays the relative error of the simulations utilising a step size 2 μm (red), 5 μm (green), and 10 μm (blue) plotted against the propagation distance. Like the previous error value, the error in Figure 5.21 is also calculated by Equation 5.25 with the spectral results simulated by utilising a step size 1 μm as the reference.

In Figure 5.21(a), which is from the single-mode propagation, the error curves are in a small scale of 10^{-12} , with those from the longer step size being on the higher level. The relationship of the step size and the error is not linear as can be witnessed by the proximity of the 5 μm and 10 μm curves compared with the separation from the 2 μm curve. The range of error value rises to the scale of 10^{-3} for the case of 5-mode propagation, as displayed in Figure 5.21(b). The details of the contributing modes and their relative intensity ratio are the same as those of Figure 5.20. The rapid increase of the value observable at the beginning of the propagation is owing to the nonlinear phase modulation of the multimode case, that is intensified by XPM. The strong phase change at each step in the simulation amplifies the influence of the step size. However, the phenomenon is reduced at longer propagation distances. The slope of the curves decreases after the propagation distance of $\sim 1 \mu\text{m}$ where the mode separation takes place and therefore XPM is slackened. The details of the mode separation will be discussed later on in Chapter 7.



(a)



(b)

FIGURE 5.21: The spectral error accumulated along the propagation distance, calculated from the simulation results utilising the step size $2 \mu\text{m}$ (red), $5 \mu\text{m}$ (green) and $10 \mu\text{m}$ (blue) with (a) single-mode propagation and (b) 5-mode propagation

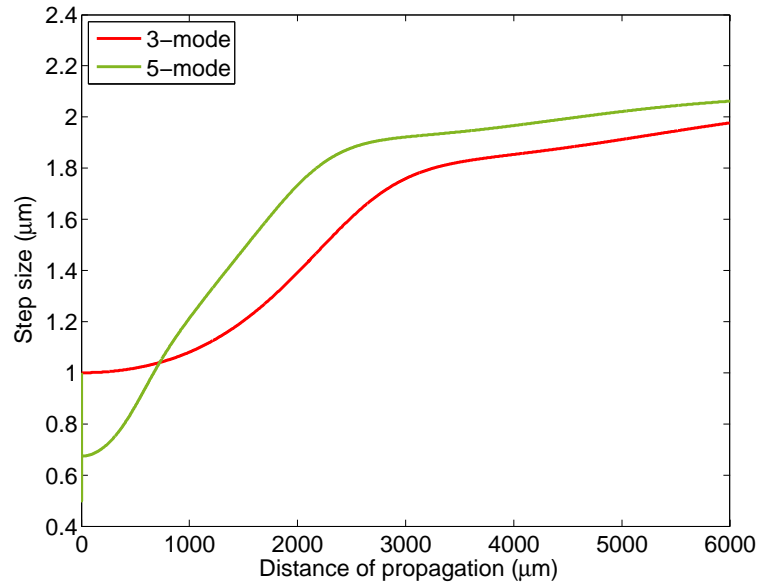
Obviously the direct method to reduce the error is by decreasing the step size in the numerical simulation. Although simple, such an approach is inefficient because in some cases the small step size is required only for certain parts of the simulation. For example, the characteristic of the accumulated error of the multimode simulation (previously mentioned) suggests the shortening of the step size over the beginning of the propagation, whereas the action is redundant for later distances. A more efficient alternative is attained by adopting the adaptive step size algorithm which has been widely used for numerical solutions of many problems [21, 22]. The algorithm allows the adjustment of the step size in order to control the local error presenting at each step in the simulation. For the numerical analysis, by utilising the $(n + 1)^{st}$ -order approximation bearing the n^{th} -order local truncation error, the local error ϵ will be bound by the error tolerance τ if the step size h is adjusted to the new value \tilde{h}

$$\tilde{h} \leq h \left(\frac{\tau}{\epsilon} \right)^{1/n} \quad (5.32)$$

In the case that the local truncation error is unknown, owing to the lack of the exact solution, the error control can focus on the relative deviation between the result simulated from the full and half step size [22]. The error control scheme is applied here in the multimode simulation, for 3-mode and 5-mode propagation, and Figure 5.22(a) displays the value of step size adjusted along the simulated propagation of the pulse in order to retain the local error of each step according to the error tolerance. The details of contributing modes and their intensity ratio are shown in Figure 5.22(b). In order to make the data comparable, both curves use the same value of tolerance, 1.2728×10^{-7} , which is the error value of the first step in the simulation of 3-mode propagation with the initial step size of $1 \mu\text{m}$.

In the case of 5-mode simulation, the initial step size is shortened down to maintain the local error at the same level as that of the 3-mode simulation. As can be observed in both curves, the step size is increased by the error control along the propagation distance due to the abating of nonlinear phase modulation caused by the mode separation. Note that the walk-off of the higher-order modes for the 5-mode case leaves less energy in the main pulse than that of the 3-mode case. For example, at the distance where E_{40}^x and E_{50}^x separate themselves from the main pulse, the residual energy, contributing in nonlinear interaction of the 5-mode pulse is decreased by 10% below that of the 3-mode pulse. Therefore the lower strength of the energy-dependent nonlinear phase modulation enables the step size of the 5-mode simulation to be expanded. This explains why the green curve of the 5-mode simulation is above the red curve of the 3-mode simulation in Figure 5.22.

Additionally, Figure 5.23 illustrates the relative error of the spectral results calculated from the error-controlled simulation. Like those in Figure 5.21, the error values are compared with the results from the simulation utilising a constant step size of $1 \mu\text{m}$,



(a)

Curve	Relative intensity ratio				
	E_{00}^y	E_{10}^y	E_{20}^y	E_{30}^y	E_{40}^y
—	0.7	0.15	0.15	—	—
—	0.7	0.1	0.1	0.05	0.05

(b)

FIGURE 5.22: (a) The step size in the multimode simulation, adjusted along the distance of propagation in order to maintain the tolerance error of 1.2728×10^{-7} . The red curve is for 3-mode simulation whereas the green curve is for 5-mode simulation. (b) Details of contributing modes and their intensity ratio

which is the same reference for the error as in Figure 5.21. The aggregation of the error with the increasing number of simulation steps disappears and instead it stays constant for the whole simulation.

Nevertheless, for such a small system as the waveguide in this research, the efficiency of the adaptive step size may not be significant in comparison with the extra computational cost for calculating the local error at each step of the simulation. Therefore, the constant step size of $1 \mu\text{m}$ is still utilised in the simulation in this research. On the other hand, for a system with a long distance of propagation or strong nonlinear and dispersion effects, the constant step size may result in a large amount of computation time and the adaptive step size can be considered as an alternative algorithm to maintain the accuracy and alleviate the computational time burden.

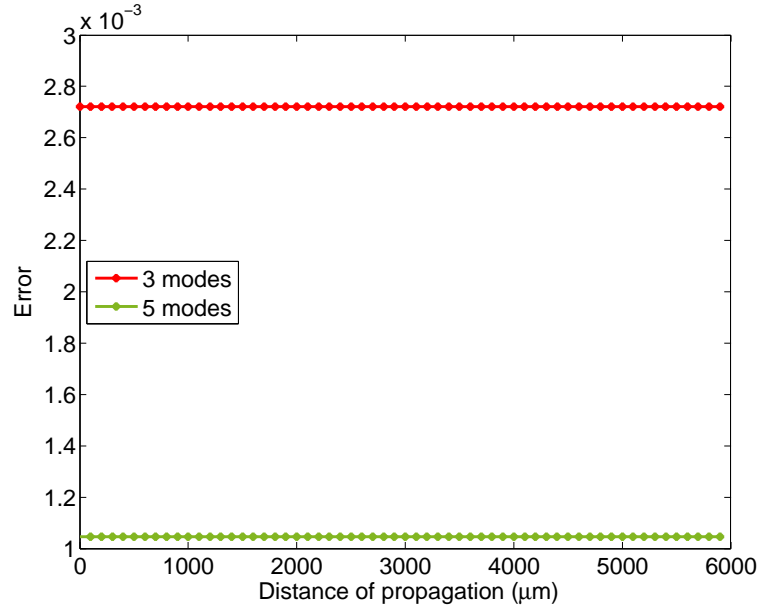


FIGURE 5.23: The spectral relative error of the error-controlled simulation compared with the results from the constant step size of $1 \mu\text{m}$ for 3-mode (red) and 5-mode (green) pulse propagations. In contrast to Figure 5.21, the error stays constant for all over the distance of propagation.

5.7 Chapter conclusion

This chapter concentrates on the general knowledge of the numerical simulation employed in this research. It starts with the explanation of the split-step Fourier method, a numerical algorithm for solving the nonlinear propagation equation. The background of the method relies on the assumption that the linear and nonlinear effects interact individually on the light pulse over the small section of the waveguide. The outcome from one section is delivered into the next section and the process is repeated across all the waveguide length. The error of the approximation is investigated and the algorithm can be adjusted to achieve the second-order accuracy.

Then the determination of the propagation constant, participating in the linear effect of the simulation, is described. The effective index method is applied to determine the propagation modes at various wavelengths and is followed by the application of interpolation and differentiation to evaluate the higher-order dispersion constant. Corresponding with the mode propagation constant, the transverse field is solved from the Helmholtz equation. The mode fields' distribution and their degree of overlapping mainly defines the intensity-dependent nonlinear effect which is quantified as the nonlinear parameter in the NLS equation.

From the basic form of single mode, the NLS equation is adapted for multimode pulses by constructing a set of differential equations. The equations represent the propagation of each mode in the waveguide and are coupled by the intermodal nonlinear effect.

Numerical solving is performed separately and the results are later combined with the transverse field to define the observed spectra measured by the SNOM probe.

Several aspects of the numerical approximation scheme are also studied. The effect of the step size on the error of the simulation is measured in order to estimate the optimised value of the step size. With the application of a numerical method, used in solving ODE, the third-order approximation for the solution of the NLS equation can be achieved but with a triple calculation cost. Finally, the adaptive step size technique, in which the step size is altered corresponding with the local error control, can improve the efficiency of the numerical simulation. The technique provides the future prospects for the application of the numerical models in a system with a long length of pulse propagation or with greater strength of linear and nonlinear phase modulations.

5.8 References

- [1] T. Chaipiboonwong, J. D. Mills, M. D. B. Charlton, M. E. Zoorob, M. C. Netti, J. J. Baumberg, and W. S. Brocklesby. Visualization of optical continuum evolution along a nonlinear waveguide. In *Conference on Lasers and Electro-Optics/Quantum Electronics and Lasers Science Conference (CLEO/QELS 2006)*, Long Beach, CA, USA, 2006.
- [2] J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, M. E. Zoorob, C. Netti, and J. J. Baumberg. Observation of the developing optical continuum along a nonlinear waveguide. *Optics Letters*, 31(16):2459–2461, 2006.
- [3] V. E. Zakharov and A. B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Soviet Physics-JETP*, 34:62–69, 1972.
- [4] R. H. Hardin and F. D. Tappert. Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations. *SIAM Review*, Chronicle 15:423, 1973.
- [5] R. A. Fisher and W. K. Bischel. Numerical studies of the interplay between self-phase modulation and dispersion for intense plane-wave laser pulses. *Journal of Applied Physics*, 46(11):4921–4934, 1975.
- [6] J. G. F. Belinfante and B. Kolman. *A Survey of Lie Groups and Lie Algebra with Applications and Computational Methods*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 1989.
- [7] J. A. Fleck, Jr., J. R. Morris, and M. D. Feit. Time-dependent propagation of high energy laser beams through the atmosphere. *Applied Physics*, 10(2):129–160, 1976.

- [8] D. Smith and P. Baumeister. Refractive index of some oxide and fluoride coating materials. *Applied Optics*, 18(1):111–115, 1979.
- [9] M. Jerman, Z. Qiao, and D. Mergel. Refractive index of thin films of SiO_2 , ZrO_2 , and HfO_2 as a function of the films' mass density. *Applied Optics*, 44(15):3006–3012, 2005.
- [10] A. Kumar, D. F. Clark, and B. Culshaw. Explanation of errors inherent in the effective-index method for analyzing rectangular-core waveguides. *Optics Letters*, 13(12):1129–1131, 1988.
- [11] K. S. Chiang. Performance of the effective-index method for the analysis of dielectric waveguides. *Optics Letters*, 16(10):714–716, 1991.
- [12] G. P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, CA, USA, 3rd edition, 2001.
- [13] C.-Y. Tai, J. S. Wilkinson, N. M. B. Perney, M. C. Netti, F. Cattaneo, C. E. Finlayson, and J. J. Baumberg. Determination of nonlinear refractive index in a Ta_2O_5 rib waveguide using self-phase modulation. *Optics Express*, 12(21):5110–5116, 2004.
- [14] S. I. Bozhevolnyi, B. Vohnsen, and E. A. Bozhevolnaya. Transfer functions in collection scanning near-field optical microscopy. *Optics Communications*, 172(1-6):171–179, 1999.
- [15] A. G. Choo, H. E. Jackson, U. Thiel, G. N. De Brabander, and J. T. Boyd. Near field measurements of optical channel waveguides and directional couplers. *Applied Physics Letters*, 65(8):947–949, 1994.
- [16] Y. Toda and M. Ohtsu. High spatial resolution diagnostics of optical waveguides using a photon-scanning tunneling microscopes. *IEEE Photonics Technology Letters*, 7(1):84–86, 1995.
- [17] S. Bourzeix, J. M. Moison, A. F. Mignard, F. Barthe, A. C. Boccara, C. Licoppe, B. Mersali, M. Allovon, and A. Bruno. Near-field optical imaging of light propagation in semiconductor waveguide structures. *Applied Physics Letters*, 73(8):1035–1037, 1998.
- [18] M. L. M. Balistreri, J. P. Korterik, L. Knipers, and N. F. van Hulst. Local observations of phase singularities in optical fields in waveguide structures. *Physical Review Letters*, 85(2):294–297, 2000.
- [19] M. L. M. Balistreri, A. Driessen, J. P. Korterik, L. Kuipers, and N. F. Van Hulst. Quasi interference of perpendicularly polarized guided modes observed with a photon scanning tunneling microscope. *Optics Letters*, 25(9):637–639, 2000.

- [20] O. V. Sinkin, R. Holzlohner, J. Zweck, and C. R. Menyuk. Optimization of the split-step Fourier method in modeling optical-fiber communications systems. *Journal of Lightwave Technology*, 21(1):61–68, 2003.
- [21] R. L. Burden and J. D. Faires. *Numerical Analysis*. Brooks Cole, Pacific Grove, CA, USA, 7th edition, 2000.
- [22] P. Linz and R. Wang. *Exploring Numerical Methods: An Introduction to Scientific Computing Using MATLAB*. Jones and Bartlett Publishers, Inc., Sudbury, MA, USA, 2003.

Chapter 6

Visualisation of nonlinear pulse propagation by SNOM

6.1 Chapter introduction

In this chapter, a SNOM-based technique will be described as a visualising tool for studying the spectral evolution along a nonlinear multimode waveguide. The measured data demonstrates several unique features which are only observable in a multimode waveguide. The multimode simulation results, performed by the modeling technique already described in the previous chapter, will be presented and compared with the measured data from the experiment. This will provide the better understanding of the underlying phenomenon of spectral evolution in the waveguide.

The chapter begins with Section 6.2 which provides the physical description of the continuum light source which was investigated in this research. In addition, the experimental setup utilised to acquire the spectral information from the evanescent field of the waveguide will be described in detail. The effect on the spectra caused by the wavelength dependence of the evanescent intensity experienced by the SNOM was also taken into consideration and will be discussed in the section.

In Section 6.3, the experimental data and the multimode simulation of the spectral evolution along the waveguide will be presented. The measured spectra display complicated features which suggests the complex interplay of various propagating modes. The assumption is qualitatively confirmed by the simulation results, which utilise similar parameters to those of the experiment although several significant experimental parameters are missing such as the power loss of light coupled into the waveguide, the absorption loss within the waveguide, and the intensity ratio of contributing modes. Similarly, simulated spectra possess complicated fine detail which are caused by the modal interference and this obviously grows with the number of modes.

The growth of the spectra in the waveguide is also studied in Section 6.4. The growth rate of the spectra from the experiment is rapid at the beginning of the propagation and then slows down at the later stage. The behaviour can be explained with the simulation modeling. The results show that in the multimode case the spectral broadening is enhanced by the additional cross phase modulation whereas only self-phase modulation participates in the broadening in the single-mode propagation. However, the nonlinear phase modulation is suppressed by the separation of modes caused by the difference in the group velocities of various modes and therefore results in the depression of the spectral growth at later distances.

The SNOM detection system makes the subwavelength observation of the spectral evolution possible and the visualisation of high resolution spectral variation along the waveguide from the experiment will be showed in Section 6.5. With the study of the simulations from various categories of mode mixing, such subwavelength variation is proven to exist when modes with different symmetry along the y -axis are mixed. The evidence, therefore, suggests that the origin of the variation over such a small scale length is the phase velocity difference of the contributing modes.

In the final part of this chapter, Section 6.6, another uncommon spectral feature observed from multimode nonlinear waveguides, namely its variation across the waveguide, will be discussed. The simulations will show that such a feature is not present in single-mode spectra. Instead, it is caused by the superposition of propagation modes with different phase and transverse field distributions.

6.2 SNOM setup for detecting continuum spectra

The conventional study of continuum spectra has been based on investigating the spectra at the end face of the waveguide by the far-field detector. Even though the acquired data can provide the information about contributing dispersion and nonlinear processes in constructing the spectrum, the nature of the data is based on the accumulated outcome from all distances of propagation. However with the application of a near-field scanning optical microscopy system, which possesses the ability of subwavelength detection, the new prospect of localised continuum generation has been revealed. The technique can provide the visualisation of developing phenomena at different points inside the waveguide. This additional knowledge can be useful for the characterisation of continuum generation sources.

In this section, the experimental setup will be described.

6.2.1 Continuum light sources

Continuum generation is a process of spectral broadening in which several nonlinear effects contribute. The phenomenon has been observed in many forms of material such as crystals and glasses [1], liquids [2], gases [3], dielectrics and semiconductors [4]. Recently, continuum generation in photonic crystal fibers and tapered fiber has been demonstrated at much lower threshold input powers than that of a bulk medium, owing to high nonlinearity caused by the effective confinement of light in small areas [5, 6].

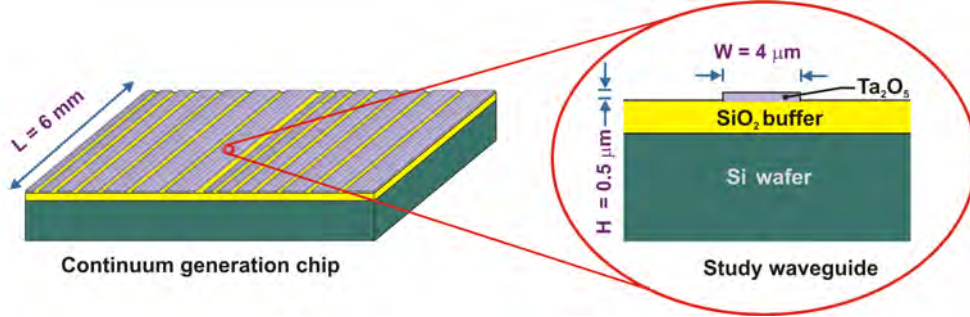


FIGURE 6.1: Continuum generation chip (CGC)

However, the continuum light source utilised in this research is another novel design based on silicon chip technology which provides future prospects for the application to all-optical integrated circuits. The continuum generation chip (CGC), fabricated by Mesophotonics Ltd, is 1 cm long 0.6 cm wide and contains sets of rib waveguides with various widths in the range of $1\ \mu\text{m}$ - $5\ \mu\text{m}$, as displayed in Figure 6.1. The structure of each waveguide guide consists of a ridge of tantalum pentaoxide (Ta_2O_5) film on a silicon substrate, which has a buffer layer of silicon dioxide (SiO_2) in between, and the rib structure is built by photolithography and argon-ion milling. Owing to the high nonlinear refractive index (n_2) of Ta_2O_5 , which is one order higher than silica glass [7], the input pulse energy can be of the order of tens of nJ. Moreover the CGC has more compactness and robustness to fracture damage, which can be caused by short wavelength energy in the broadband spectrum, as is observed in fibres.

6.2.2 Collecting continuum spectra

The experimental setup is displayed in Figure 6.2. The Gaussian light pulse is generated from a COHERENT Ti:Sapphire laser with repetition rate 76 MHz and wavelength 800 nm. The pulse width, determined from the Fourier transform of the intensity spectra, is around 86 fs, whereas the autocorrelation measurement gives around 100 fs. The spectral FWHM of the input pulse is around 11 nm and the pulse energy varies from 0.8 nJ to 2.1 nJ. The pulse was focused and launched into the waveguide with the assistance of an aspherical lens with 0.68 NA and 3 mm focal length. Then the light propagating in the waveguide is confined within the higher refractive index layer of Ta_2O_5 that is

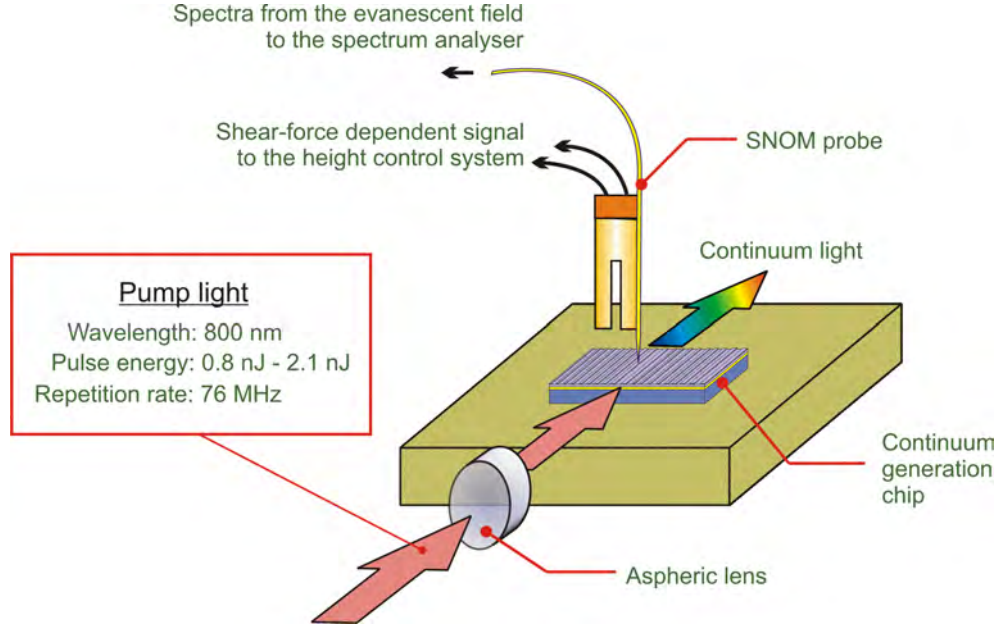


FIGURE 6.2: Experimental setup for visualising continuum spectra developing inside the waveguide

surrounded by the lower-refractive index of air on the top and SiO_2 on the bottom. Theoretically, full confinement of light is achieved for the light ray incident with the angle more than the critical angle of the Ta_2O_5 - SiO_2 interface, which is around 45° at the wavelength of 800 nm with $n_{\text{SiO}_2} = 1.48$ and $n_{\text{Ta}_2\text{O}_5} = 2.08$.

The SNOM probe was placed at around 10 nm - 20 nm above the surface of the waveguide and controlled by the shear force feedback system, as explained in Section 3.4, Chapter 3. The probe detects the evanescent field induced by the total internal reflection at the air- Ta_2O_5 interface of the waveguide. The intensity decays exponentially along the gap distance, in a direction away from the surface of the guide, by the relation

$$I_n(y) \propto \exp(-2\gamma_n y) \quad (6.1)$$

where γ_n is the extinction coefficient for the mode field with the mode index n . The extinction coefficient, which is the transverse component of the wavevector, is also determined in the process of the calculation of the propagation modes, which has already been described in Section 5.3, Chapter 5. The value of the coefficient can be more physically viewed as the penetration depth (δ), which is the depth in the cladding area at which the intensity is reduced down to $1/e \sim 36.79\%$ of the maximum intensity at the interface. Therefore

$$\delta_n = \frac{1}{2\gamma_n} \quad (6.2)$$

Polarisation	Penetration depth (nm)	
	$n = 0$	$n = 1$
E^y	37.70	50.56
E^x	36.85	45.56

TABLE 6.1: Penetration depth for the evanescent field at the interface of air and Ta_2O_5 of the waveguide

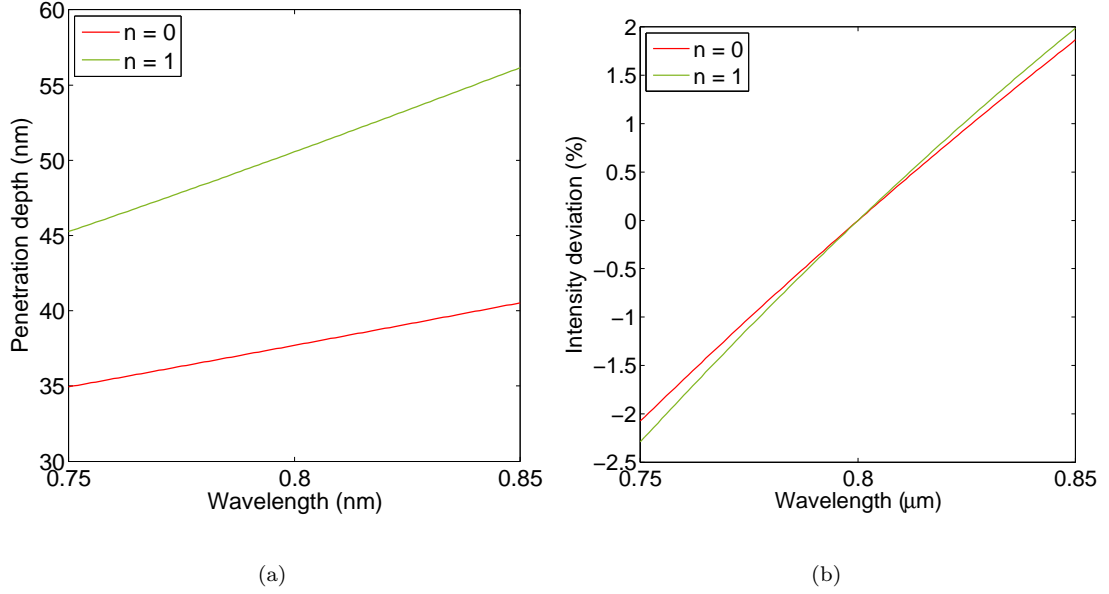


FIGURE 6.3: (a) Variation of penetration depth over the study wavelength range (b) Intensity variation caused by the wavelength dependency of the penetration depth when the probe height is 20 nm above the surface

As previously mentioned in Section 5.3, Chapter 5, for the light confinement along the height dimension of the waveguide, only modes with the mode index number $n = 0$ and $n = 1$ can exist. The penetration depths, at the input wavelength 800 nm, for these mode fields at the upper surface of the guide are given in Table 6.1. The higher-ordered mode and E^y polarisation can penetrate deeper and therefore a greater amount of the field will be collected by the SNOM probe. Moreover the shape of the spectra can also be affected by the wavelength dependency of the penetration depth. Figure 6.3(a) shows the variation of the penetration depth as a function of wavelength. Over the wavelength range covering the studied spectra, the penetration depth varies around 15% - 20%. However, the effect causes much less variation of intensity experienced by the probe as shown in Figure 6.3(b). The intensity collected by the probe is plotted in terms of the percentage deviation from the value at the wavelength of 800 nm. The deviation is only 5% over the wavelength 750 nm - 850 nm.

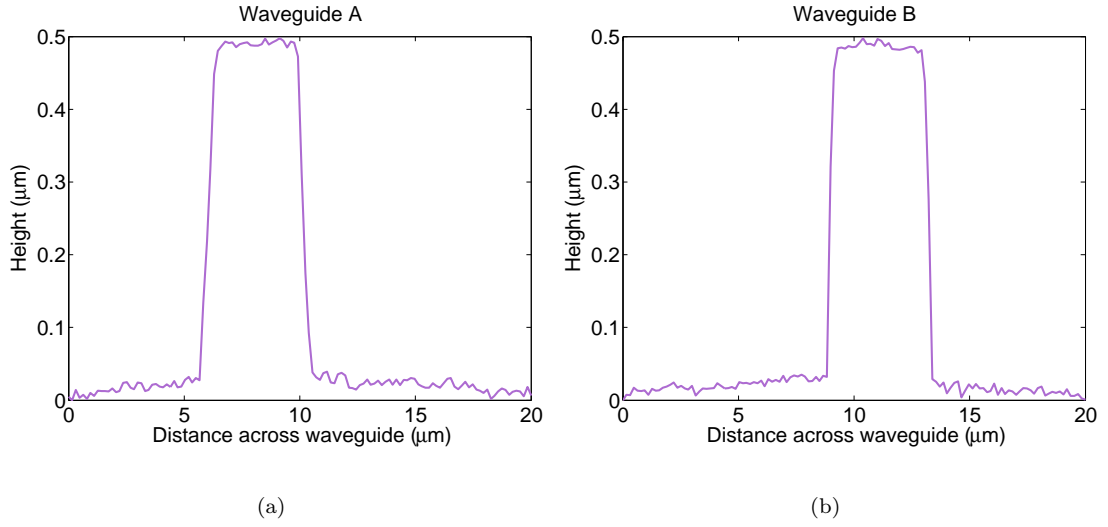


FIGURE 6.4: Topography of (a) Waveguide A (b) Waveguide B

Waveguide label	Average width from topographical data(μm)
A	$4.12 \pm 4.35\%$
B	$4.29 \pm 1.86\%$

TABLE 6.2: Label and average width of waveguides from topographical data. The SNOM probe scanned across the waveguide at various distances from the front face of the waveguide and the precision in the table is calculated from the standard deviation of the measured data.

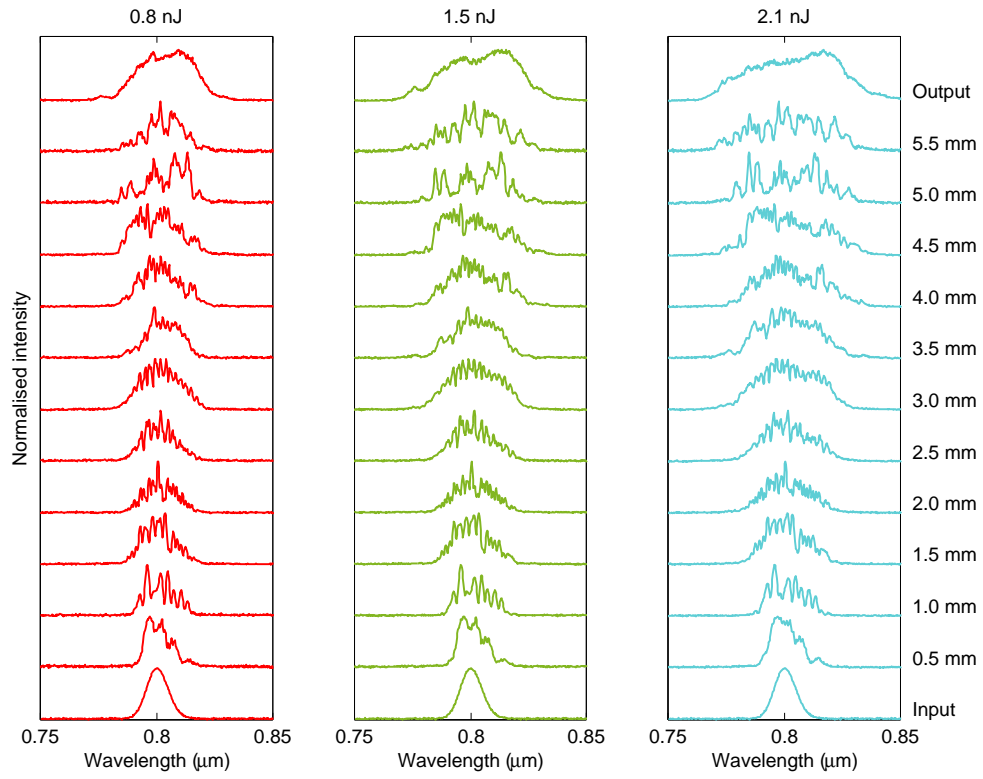
The light collected from the probe was passed through a computer-controlled high resolution spectrometer with the details of instrument specification already given in Section 3.5, Chapter 3. Two adjacent waveguides on the chip were chosen to use in the experiment. Both are $0.5 \mu\text{m}$ high and 6 mm long but with different widths. The label for each waveguide and its width are shown in Table 6.2. The width is the average value measured from the topographical data of across-waveguide scanning at different locations along the length of the guide. Some of the topographical data are shown in Figure 6.4. Certainly, the width values acquired by the SNOM technique are larger than the real physical width of the waveguide since the topographical data is the outer contour of the waveguide which includes the gap distance between the probe and the waveguide surface. Additionally, the accuracy also depends on the cleanliness of the surface and the sharpness of the probe.

Note that the setting-up of the instruments and the collection of the experimental results of this research were assisted by Dr. John Mills [8–10] whereas the simulation work is based on discussions with Dr. Peter Horak.

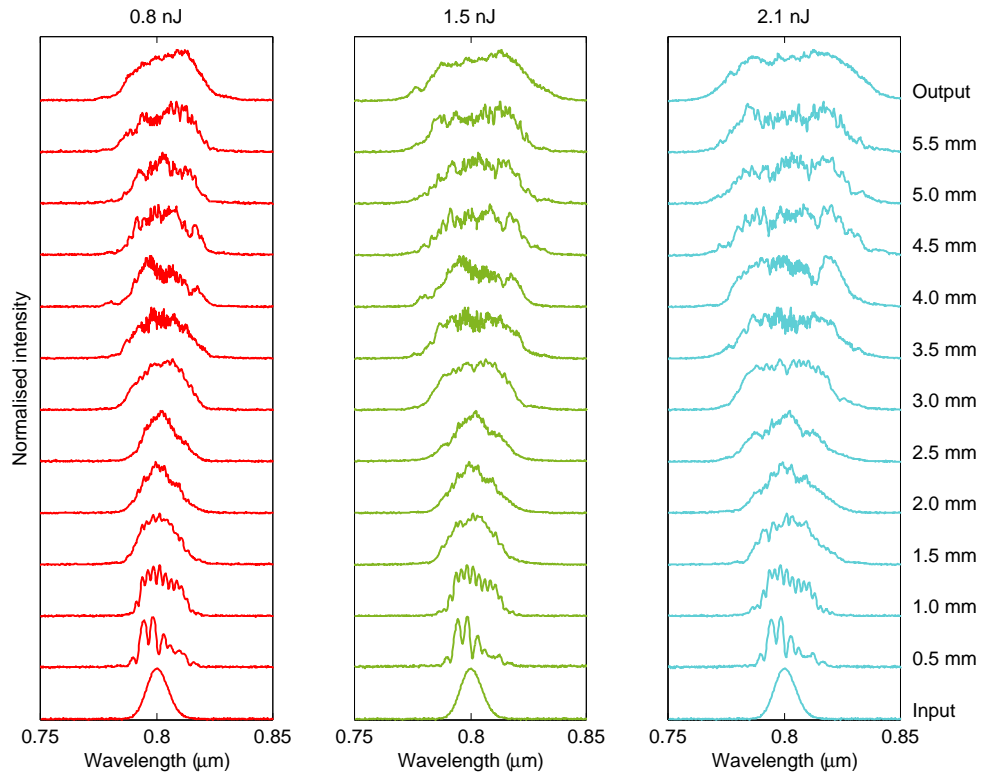
6.3 Spectral evolution along the waveguide

6.3.1 Experimental data

Figure 6.5 displays the spectra collected by the system. The spectra are in the wavelength scale and the intensity of the spectra shown is normalised to unity at the maximum intensity. Unlike the spectra at other locations, the input and output spectra from the waveguide are measured by a far-field detector. Both are added at the ends in the figures, whereas the SNOM-acquired spectra collected by the probe placed at the position of $z = 0.5$ mm to $z = 5.5$ mm are shown in-between. As can be clearly observed, the spectra are broadened along the length of the waveguide, owing to the increase of the dispersion and nonlinearity effects with the distance of propagation. By comparing the spectra from different input pulse energies, the increase of the broadening phenomenon with the input pulse energy is certainly apparent, as can be expected owing to the dependency of nonlinear processes on the intensity. The locality of the information is confirmed by the reappearance of several distinctive peaks in the spectra from the same propagation distance at different input energies. The complex appearance of the spectra from the probe may illustrate the localised and high spatial resolution nature of the SNOM data. On the contrary, such fine details do not appear in the output spectrum which is theoretically the integrated spectral data from all points on the exit face of the waveguide and, as a result, contains only spatially averaged information. Note that there is a noticeable reduction in the oscillatory behaviour of the spectra collected at the distance 2 and 3 mm for Waveguide B. Although the cause of this appearance cannot be specifically explained, changes in the oscillatory pattern of the spectra at various propagation distances may be related to the interference of propagation modes, as observed in Balistreri et al.'s research [11]. In that study, the near-field probe could visualise the spectral interference of two waveguide modes and the variation of the period length in the interference pattern along the length of the waveguide.



(a)



(b)

FIGURE 6.5: Continuum evolution along the waveguide at different input pulse energies: 0.8 nJ (—), 1.5 nJ (—) and 2.1 nJ (—) for (a) Waveguide A and (b) Waveguide B respectively [9]

6.3.2 Simulation of multimode pulse propagation along the waveguide length

In order to explain the unique features observed in the experiment, the NLS equation corresponding to the nonlinear system of the experiment was numerically solved. However, since the system is not a single mode system but a multimode one, the NLS equation needs to be adapted, as was previously discussed in Section 5.5, Chapter 5. Therefore Equation 5.16, Chapter 5 was simulated by using the parameters similar to those of the experiment. However, there are some other unknown parameters such as the number of guided modes induced by the pump light and the division of the input energy among the modes which depends on the correlation between the field distribution of the pump light and that of the guided modes [12]. Therefore the simulation results, which will be shown in this chapter and the following one, are not a theoretical replica of the experimental results. Instead they provide the understanding of the underlying physical background in a complicated spectral evolution in multimode nonlinear waveguides.

The simulation starts with creating the input envelope field fed into the modeling. The input pulse is constructed from a Gaussian function which is in the form as

$$V(t) = V_0 \left(\frac{2a}{\pi} \right)^{\frac{1}{4}} \exp(-at^2) \quad (6.3)$$

The Equation 6.5 is normalised such that

$$\int_{-\infty}^{+\infty} |V(t)|^2 dt = V_0^2 \quad (6.4)$$

The parameter V_0 in Equation 6.5 is the pulse amplitude whose squared value is the peak energy of the pulse. Therefore V_0 can be written in terms of the pump power P_i and the repetition rate Γ as

$$V_0 = \left(\frac{P_i}{\Gamma} \right)^{\frac{1}{2}} \quad (6.5)$$

The parameter a in the exponential term of Equation 6.5 specifies the width of the Gaussian function and relates to the FWHM width of the pulse intensity profile Δt_{FWHM} via the relation

$$a = \frac{2 \ln 2}{\Delta t_{\text{FWHM}}^2} \quad (6.6)$$

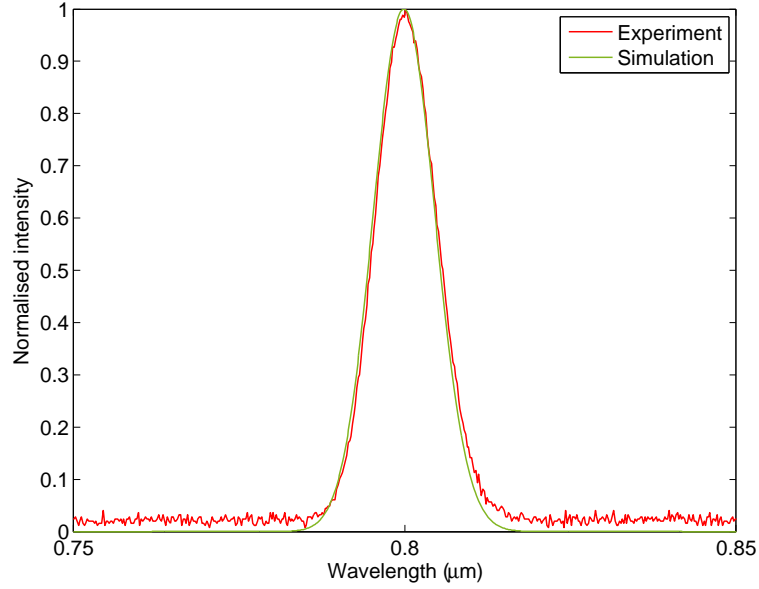


FIGURE 6.6: The spectrum of the input pulse, where the red curve is from the measured spectra and the green curve is from the simulation

The simulated input pulse is created by using 2^{12} data points. This covers the wavelength range of 0.6 - 1.3 nm. The time window size is 10 ps, which is around 150 times the pulse width. The value of the parameter Δt_{FWHM} is 85.707 fs. The Fourier transform of the time profile provides the spectra with the FWHM width $\Delta \lambda_{\text{FWHM}} = 11$ nm. The value is equal to that of the measured input spectrum in the experiment. Figure 6.6 displays the Fourier transform of the simulated input pulse (green curve) in comparison with the measured input spectrum from the experiment (red curve).

The input pulse energy of 2.1 nJ is chosen for the simulation. However, in practice, not all the amount of energy was coupled into the waveguide. The coupling loss is attributed to the power spill over the top, bottom and side faces of the waveguide and the reflection loss at the front face. The exact figure of the loss is unknown. The coupling efficiency of 0.03 is arbitrarily chosen for the simulation. This parameter is defined as the ratio between the coupled power through the waveguide to the incident power. The total energy is divided among the contributing modes, as specified by the parameter of relative intensity ratio. The step size in the simulation is $1 \mu\text{m}$ which is justified to provide numerical stability, as already discussed in Section 5.6.1, Chapter 5 for the effect of step-size on the numerical results. The dispersion and nonlinear parameters corresponding to the modes are available from the calculation criteria, which were previously explained in Section 5.3 and Section 5.5, Chapter 5. The resultant spectrum is calculated from the integration of the evanescent field at each location over the diameter of a 100nm SNOM tip, at the height of 20 nm above the surface of the guide. The area of the tip is parallel to the xz -plane of the waveguide geometry but the integration is performed one-dimensionally. Only the field in the range of the probe diameter, 100 nm, along the x -axis is gathered whereas the variation along the z -axis is negligible because

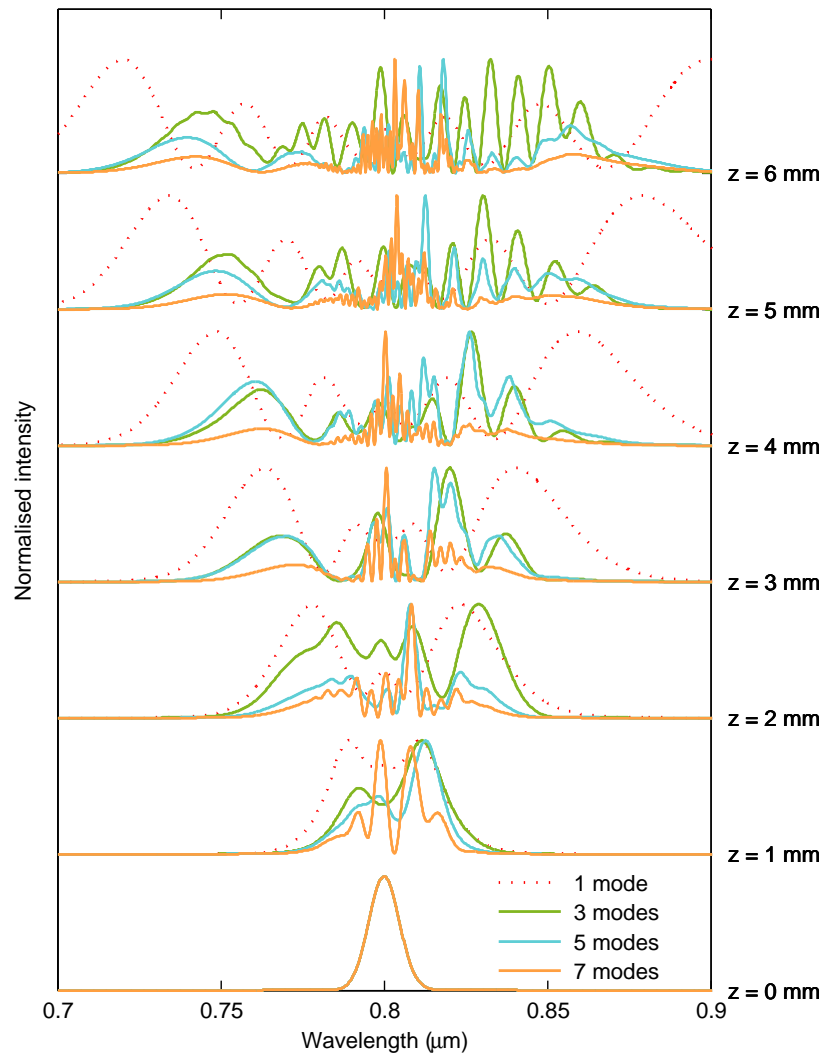
the transverse field is assumed to be z -independent and the envelope field changes insignificantly over such a distance, which is 100 times less than the simulation step size. The details of field integration was described in Section 5.5, Chapter 5.

The transverse field $U(x, y)$ is approximated to be constant over the wavelength range of the spectra. Its value at the central wavelength 800 nm is utilised in the calculation. The effect of the wavelength dependency of $U(x, y)$ will be later discussed in Section 7.4, Chapter 7.

The simulated spectral evolution along the length of the waveguide, of various multimode pulses, is shown in Figure 6.7(a) with details of the corresponding modal intensity ratio given in Figure 6.7(b). The total power, which is equal for all simulations, is distributed among various modes in such a way that the main part of the input energy, 70% in this case, is concentrated in the fundamental mode and the rest is divided quite evenly among the higher-order modes. Since there is no measurement of power distribution among the waveguide modes from the experiment, the intensity ratio of the simulation is just chosen arbitrarily for demonstrating the effect of contributing modes.

The modeling of the pulse propagation is done over the 6 mm length of the waveguide with the simulated position along the x -axis of the SNOM measuring probe set at $x = 200$ nm, in order to avoid a zero contribution from the field of the odd modes which possess an antisymmetric field distribution around the central axis of the guide ($x = 0$). The intensity is normalised to unity at the maximum intensity in order to compare spectral shapes.

Unlike the SPM spectral profile of the single-mode (dotted red curve), the multimode spectra in Figure 6.7(a) lack symmetry, exhibiting a more complex structure which is similar to results gained from the experiment. Indeed the fine details of interference become more complicated with increasing modal contribution, and distance traveled. This is due to the interference of individual modes with different linear and nonlinear phase shifts. However, the simulated spectra seem to possess more features than the measured one. This lower refinement is possibly caused by fluctuations of experimental conditions such as thermal changes in the refractive index of waveguide material (the temperature coefficient of the refractive index of Ta_2O_5 varies in the range of $3.64 \times 10^{-6} \text{ K}^{-1}$ - $7.89 \times 10^{-6} \text{ K}^{-1}$ at 632.8 nm for the temperature range of 298 K - 328 K [14]), pulse-to-pulse energy variation or spatial vibration of the waveguide and the probe.



(a)

Curve	Relative intensity ratio						
	E_{00}^y	E_{10}^y	E_{20}^y	E_{30}^y	E_{40}^y	E_{50}^y	E_{01}^y
.....	1	—	—	—	—	—	—
—	0.7	0.15	0.15	—	—	—	—
—	0.7	0.1	0.1	0.05	0.05	—	—
—	0.7	0.05	0.05	0.05	0.05	0.05	0.05

(b)

FIGURE 6.7: (a) Simulation of the spectral evolution of multimode pulse along the length of the waveguide. The simulated spectra are observed by the probe which is placed above the waveguide by 20 nm and shift from the central axis of the waveguide by 200 nm (b) Details of contributing modes and their relative intensity ratio [13]

6.4 Spectral growth along the distance of propagation

6.4.1 Experimental data

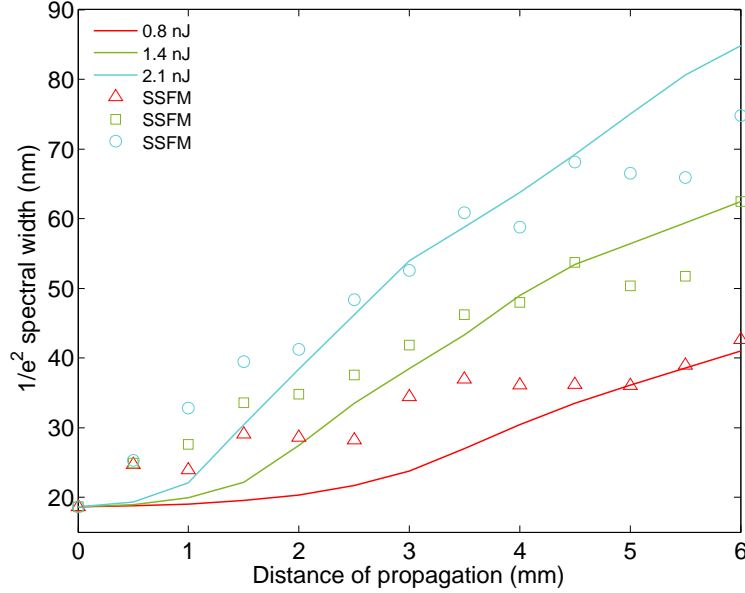


FIGURE 6.8: Comparison of the width of the measured spectra along the length of the waveguide, which is illustrated in Figure 6.5(b), with that from the single mode simulation [9]

The spectral growth is studied by measuring the spectral width of the evolving spectra, displayed in Figure 6.5(b), as a function of propagation distance and the result can be viewed in Figure 6.8. The width is measured at the point where the intensity is $1/e^2 \sim 13.53\%$ of the peak. The data is compared with the result from the single-mode simulation, which is performed by approximating the value of launch efficiency to fit with the output from the intermediate input energy 1.4 nJ. The same value of the efficiency is used to simulate the results for other input energies, 0.8 nJ and 2.1 nJ.

The significant deviation between the measured value and the single-mode simulation is noticeable, especially at the beginning of the propagation. The continuous growth of the simulated spectra cannot fit with the spectral growth of the experimental data, which has retarded spectral growth rate at various points along the length of the waveguide. Since the experiment is under the low-energy regime, in which the nonlinear process is least complicated, the deviation suggests that the multimode nature of the nonlinear propagation yields some specific characteristics.

6.4.2 Spectral growth of the multimode simulation data

The measurement of width is also applied to the simulated spectra, shown in Figure 6.7(a). However, for such complicated features of the simulated spectra, the root-

mean-square (RMS) width is more preferable for describing the spectral broadening. The RMS width Δt_{RMS} of the pulse specified by the field $E(x, y, z, t)$ is defined as [15]

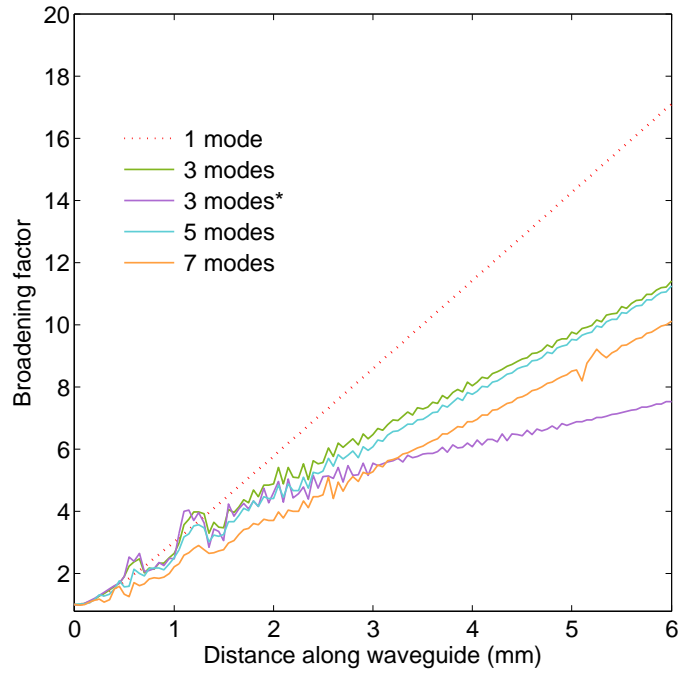
$$\Delta t_{\text{RMS}} = (\langle t^2 \rangle - \langle t \rangle^2)^{\frac{1}{2}} \quad (6.7)$$

where

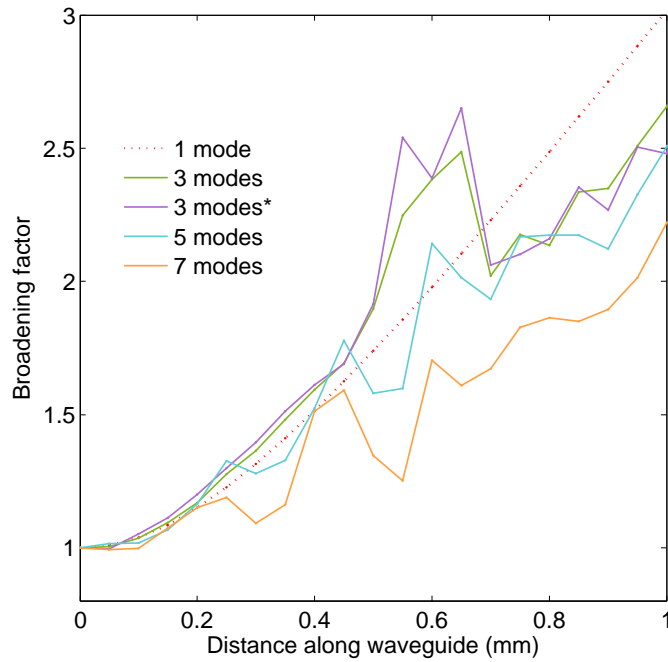
$$\langle t^n \rangle = \frac{\int_{-\infty}^{+\infty} t^n |E(x, y, z, t)|^2 dt}{\int_{-\infty}^{+\infty} |E(x, y, z, t)|^2 dt} \quad (6.8)$$

Figure 6.9(a) displays the spectral width of the simulated spectral data as it develops along the waveguide, over the propagation distance of 6 mm, whereas Figure 6.9(b) is the closeup of Figure 6.9(a) at the first one mm from the beginning of the propagation distance. The data is from various mode mixing, which has already been displayed in Figure 6.7, with an additional 3-mode curve (violet) which is the mixing of E_{00}^y , E_{10}^y and E_{20}^y with intensity ratios 0.4:0.3:0.3. The width shown in the figure is normalised in terms of the pulse width broadening factor, which is defined as $\Delta t_{\text{RMS}}/\Delta t_{\text{RMS},0}$ where $\Delta t_{\text{RMS},0}$ is the pulse width at the beginning of the propagation.

The results agree with the experiment, in which the growth of the spectrum is slower than expected, and cannot be fitted by a simple single-mode, SPM calculation. At the beginning of the propagation shown in Figure 6.9(b), when mode separation is insignificant, the strength of XPM between the modes can cause multimode broadening that is stronger than the single-mode case, as can be observed in the 3-mode and 5-mode curves. In contrast to the single-mode curve, there is fluctuation of pulse width along the propagation distance and the degree of fluctuation rises with the number of mixing modes. Such a feature is caused by the difference of phase velocity of modes, the term $(\beta_0^{(p)} - \beta_0^{(q)})z$ in Equation 5.20, Chapter 5. In order to observe the interplay of the nonlinear phase modulation and the mode separation more clearly, the phase velocity difference is removed from the data in Figure 6.9(a) and the outcome is shown in Figure 6.10. The curves in the latter figure display the overall trend of the spectral growth, which is not obscured by rapid changes caused by the phase velocity difference. The deviation between both figures is higher at the beginning of the propagation where more rapid fluctuations in pulse width are noticeable, as shown in Figure 6.10. This is due to the strong coupling of modes which intensifies the phase velocity difference before the mode separation at a later distance.



(a)



(b)

FIGURE 6.9: (a) RMS pulse width broadening factor along the length of the waveguide of the simulated multimode spectra shown in Figure 6.7, except for an additional 3-mode curve (violet) which is the mixing of E_{00}^y , E_{10}^y and E_{20}^y with intensity ratio 0.4:0.3:0.3
 (b) Details of (a) at the first 0.5 mm distance of propagation

Figure 6.10 clearly demonstrates that, in order to strengthen the effect, a higher amount of energy must be concentrated in higher-order modes. The evidence can be viewed by comparing both 3-mode curves in which the violet curve has twice the energy concentrated within the higher-order modes. The spectral growth of the violet 3-mode curve, which initially has the greatest rate of growth, diminishes due to the separation of E_{20}^y at around the distance of 1.5 mm. After this point, the main contribution of spectral broadening is due to the interaction of E_{00}^y and E_{10}^y . Consequently, the green curve, which has higher energy within these modes, continues broadening at a rate such that it overtakes the violet curve. The 5-mode (blue) and 7-mode (orange) curves also show the retardation of nonlinear phase modulation caused by mode separation. In the 5-mode case the rate is reduced to less than that of the single-mode case at a shorter distance, due to the early separation of E_{30}^y and E_{40}^y , whereas broadening in the 7-mode system cannot be higher than in the single-mode case owing to the temporal walk-off of higher-order modes at the very beginning of the propagation.

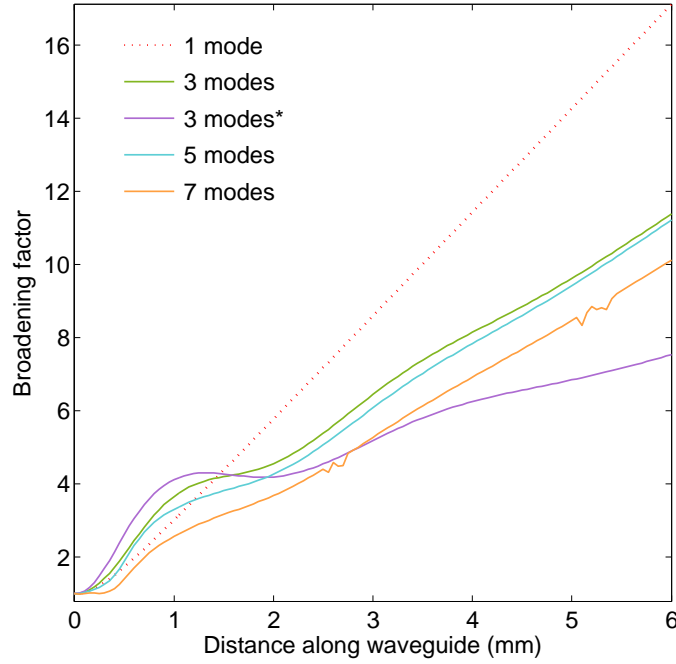
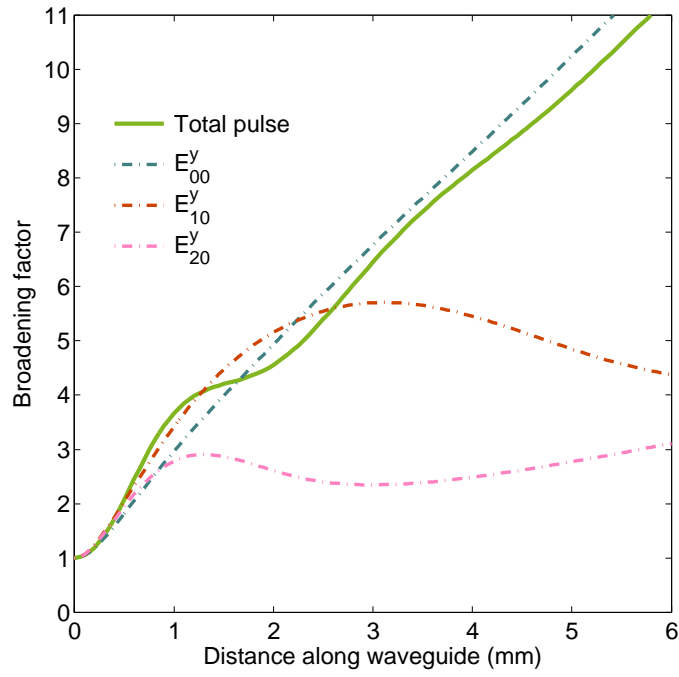


FIGURE 6.10: Same as Figure 6.9(a) but the phase velocity difference of modes is neglected [13]

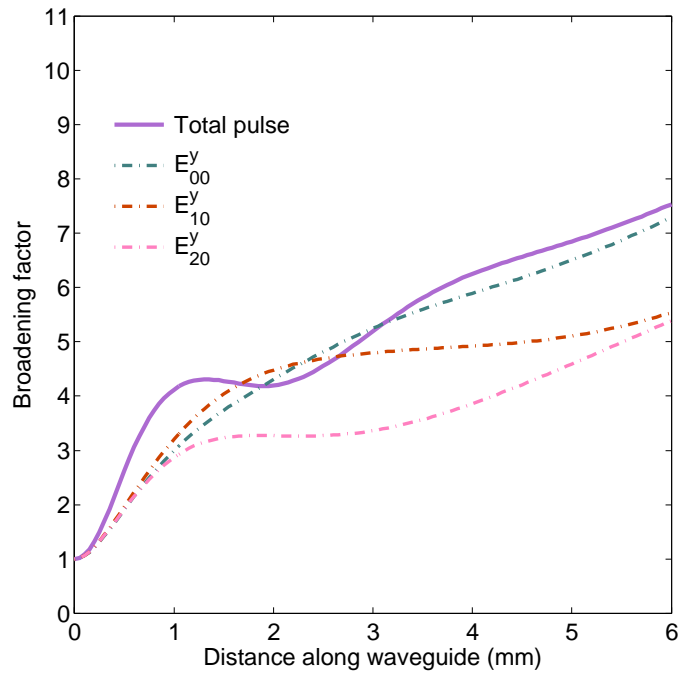
To clarify the effect of contributing modes on the overall spectrum, the spectral widths of individual modes for two cases of three-mode mixing (green and violet curves in Figure 6.10) are also displayed in Figure 6.11. In the first millimetre, the spectrum of each individual mode grows rapidly, in addition to the total spectrum. Note that the widths of the curves E_{10}^y and E_{20}^y are greater than that of E_{00}^y at this stage, because XPM in the higher-order modes is fed by the main contribution of intensity from the E_{00}^y mode. Conversely, only a fraction of 30% is contributing to E_{00}^y XPM from these higher order modes. At the distance when E_{20}^y starts to separate, the decrease in the slope of the E_{20}^y spectral growth, caused by the lack of E_{00}^y and E_{10}^y in the E_{20}^y spectral

development, is apparent in both figures. After this stage, E_{20}^y grows with distance only due to its SPM. The SPM spectral growth of E_{20}^y in Figure 6.11(b) increases to higher than that in Figure 6.11(a) owing to 50% higher intensity ratio of E_{20}^y . The effect of mode separation also occurs with E_{10}^y at a later stage. The more distinct drop of the E_{10}^y curve in Figure 6.11(a) than the same curve in Figure 6.11(b) is owing to a stronger influence of the fundamental mode in spectral development at a previous distance. The amount of E_{10}^y is 70 % in the first case, whereas only 40 % for the latter. Finally, after all the mode separation, the overall spectral growth is totally governed by the SPM of the fundamental mode.

So far, propagation losses have been neglected in the modeling. However, based on the above discussion, the spectral modifications can be predicted if higher-order modes exhibit significant losses. If losses are small on the length scales where temporal walk-off occurs, no significant spectral changes are expected since all broadening due to XPM occurs at an early stage of pulse propagation. On the other hand, if higher-order mode losses are large over much shorter distances, no significant XPM takes place and the observed spectra will probably resemble the single-mode spectrum of Figure 6.7(a).



(a)



(b)

FIGURE 6.11: (a) RMS pulse width broadening factor of individual modes contributing to the three-mode pulse (green curve) in Figure 6.10 (b) Same as (a) but for three-mode pulse (violet curve) in Figure 6.10

6.5 High resolution spectral evolution

In this section, the spectral development over the subwavelength distance was observed in the experiment. The variation in such a diminutive scale length is uniquely observed only in the multimode spectra as it can only be explained by the effect of phase velocity difference.

6.5.1 Experimental data

The high-resolution details of spectral evolution were also observed, as shown in Figure 6.12. The probe was moved to the position of $z = 3$ mm. Then it scans at 100 nm intervals within the distance of $2 \mu\text{m}$, to collect the spectra. The spectra are in the wavelength scale on the x-axis and the distance along the waveguide is on the y-axis. The spectral variation is clearly observable.

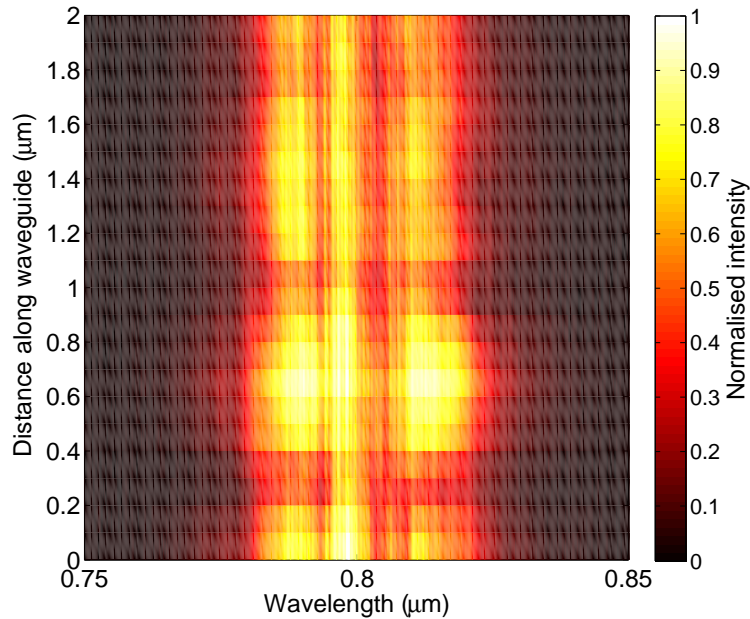


FIGURE 6.12: High-resolution spectral mapping along the waveguide over the distance of $2 \mu\text{m}$. The spectra were collected when the probe was at the position of $z = 3$ mm and moved toward the end face by the steps of 100 nm [9]

6.5.2 Simulation result

As previously demonstrated in Equation 5.20, Chapter 5, the spectrum of a multimode pulse is spatially modulated by the propagation phase difference term, $(\beta_0^{(p)} - \beta_0^{(q)})z$ and the phase difference term induced from the interplay of dispersion and nonlinearity, $(\varphi^{(p)}(z, \omega) - \varphi^{(q)}(z, \omega))$. The modulation period of the latter is much longer than the

first and therefore the wavelength-scaled spectral variation observed in Figure 6.12 can be traced back to the interference of modes owing to the propagation phase difference.

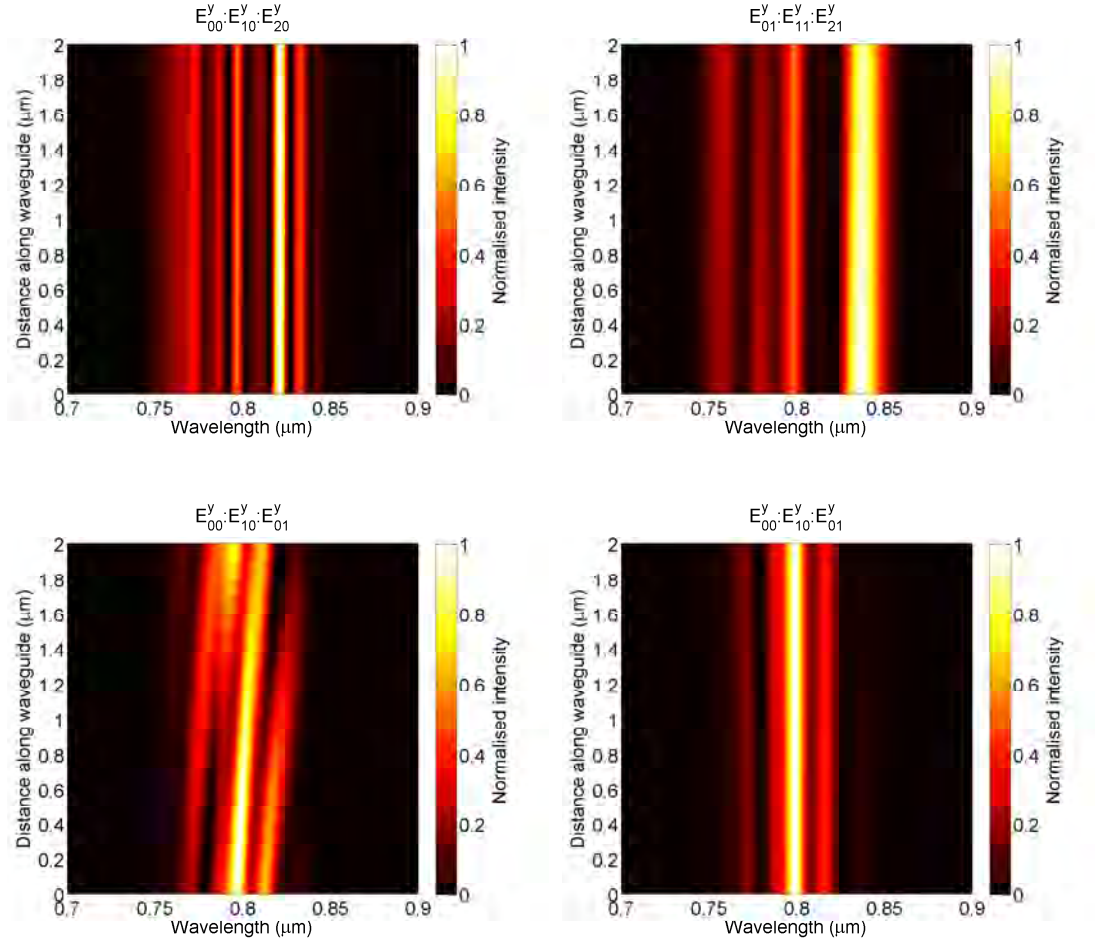


FIGURE 6.13: Simulated high-resolution spectra through the probe when the probe moves along the length of the waveguide by the distance of $2 \mu\text{m}$, starting from the position of $z = 4 \text{ mm}$. The simulation is performed with various multimode pulses in the waveguide: (a) $E_{00}^y : E_{10}^y : E_{20}^y$ (0.4:0.3:0.3) (b) $E_{01}^y : E_{11}^y : E_{21}^y$ (0.4:0.3:0.3) (c) $E_{00}^y : E_{10}^y : E_{01}^y$ (0.4:0.3:0.3) (d) Same as (c) but the phase velocity difference is neglected

The behaviour of multimode pulse propagation over the distance of few hundred nanometres can be studied by the multimode simulation by reducing the step size to the scale of 100 nm . The simulation results shown in Figure 6.13 are the spectral features observed by the probe over the distance interval $2 \mu\text{m}$ along the length of the waveguide. The probe translation starts from the position of $z = 4 \text{ mm}$ and then moves toward the end face. The spectra of three different mode mixing examples are demonstrated here. Figure 6.13(a) displays the mixing of three symmetric (the mode index $n = 0$) modes E_{00}^y , E_{10}^y and E_{20}^y by the relative intensity ratio 0.4, 0.3 and 0.3 respectively whereas Figure 6.13(b) is from the mixings of asymmetric (the mode index $n = 1$) E_{01}^y , E_{11}^y , and E_{21}^y by the same ratio. The mixing of modes from both types is in Figure 6.13(c) which consists of E_{00}^y , E_{10}^y and E_{01}^y also mixed by the same ratio.

From the numerical value of $\beta_0^{(p)}$ for various modes in Table 5.2, Chapter 5, the period of spatial variation of the term $(\beta_0^{(p)} - \beta_0^{(q)})z$ for some first lower-ordered modes is in the range of 10 μm - 100 μm if both modes p and q have the same symmetry. However the frequency of spatial modulation is decreased by one order for modes from different symmetry and therefore the period is around the order of 1 μm . This explains the absence of any spatial variation of spectra in Figure 6.13(a) and Figure 6.13(b), which are from the mixing of modes possessing the same symmetry. In contrast, the variation within the range of 2 μm appears in Figure 6.13(c) whose simulation is from the multimode mixing consisting of both symmetric (E_{00}^y and E_{10}^y) and antisymmetric modes (E_{01}^y). To confirm that the cause of variation comes from the propagation phase shift between symmetric and asymmetric modes, the term $(\beta_0^{(p)} - \beta_0^{(q)})z$ is nulled in the calculation of intensity in Equation 5.20, Chapter 5 for the simulation utilising the same set of contributing modes as Figure 6.13(c). The outcome is shown in Figure 6.13(d), where no spectral variation is noticeable.

To confirm that the variation is exactly owing to the multimode nature, Figure 6.14 displays the high resolution spectra of the individual contributing modes of the spectrum in Figure 6.13(c). Although it can be easily predicted, there is no observable variation occurring in any spectra of Figure 6.14(a)-Figure 6.14(c).

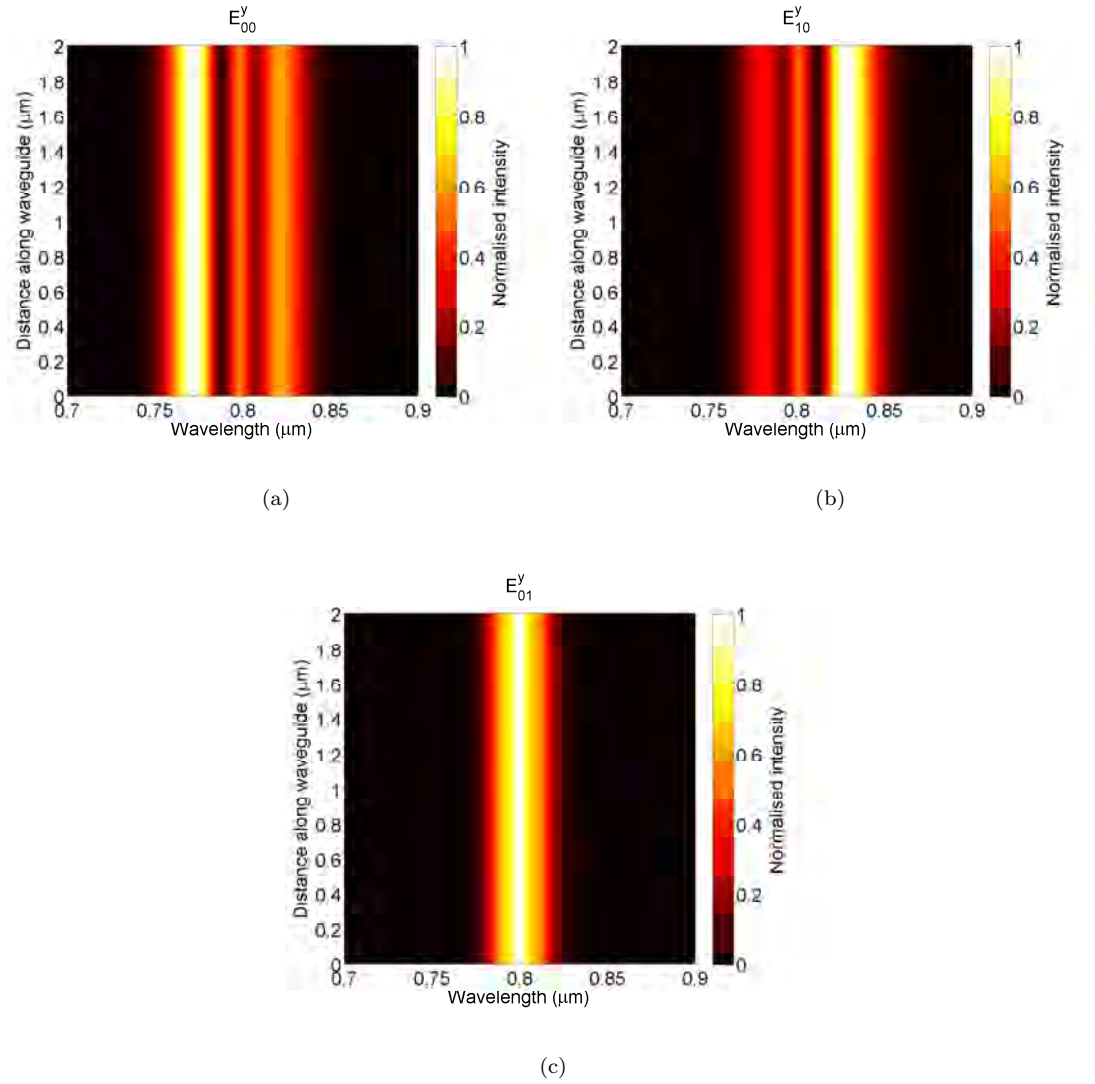


FIGURE 6.14: Same as Figure 6.13(c) but for individual contributing modes (a) E_{00}^y (b) E_{10}^y and (c) E_{01}^y

6.6 Spectral variation across the waveguide

This section discusses another unique characteristic of observed spectra, which is the spectral variation across a multimode nonlinear waveguide. The simulation can prove that the cause of the feature is the interference of modes containing various transverse field distributions, whose degree of complexity increases with the mode number.

6.6.1 Experimental data

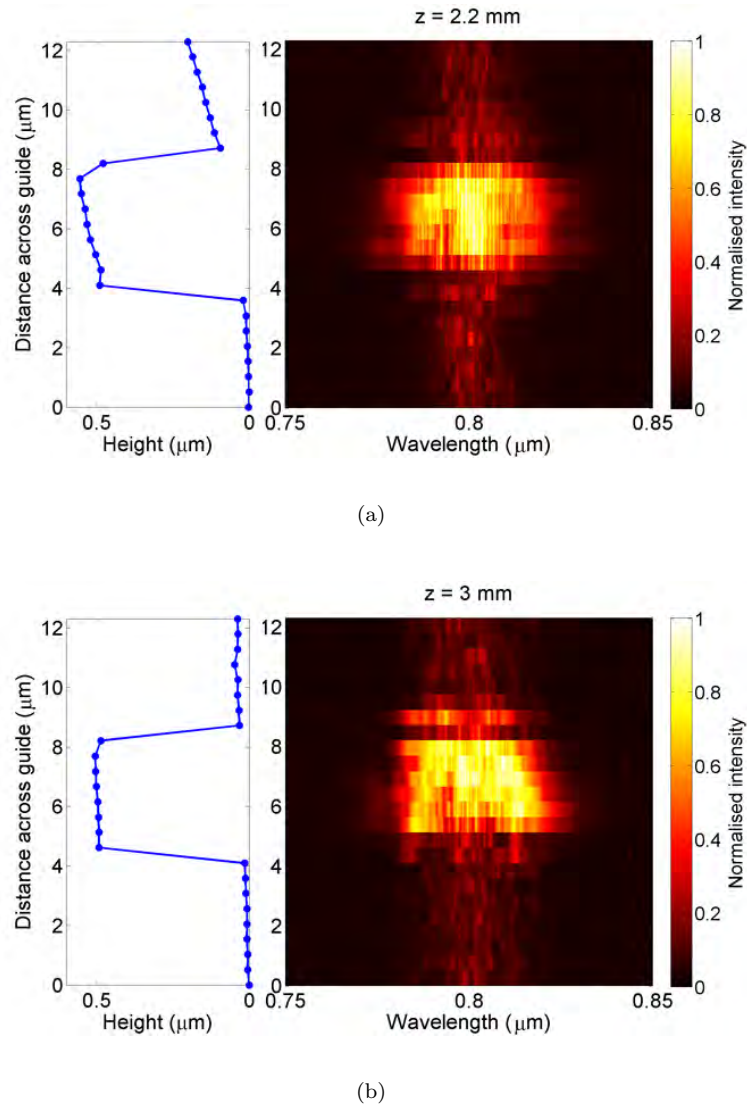
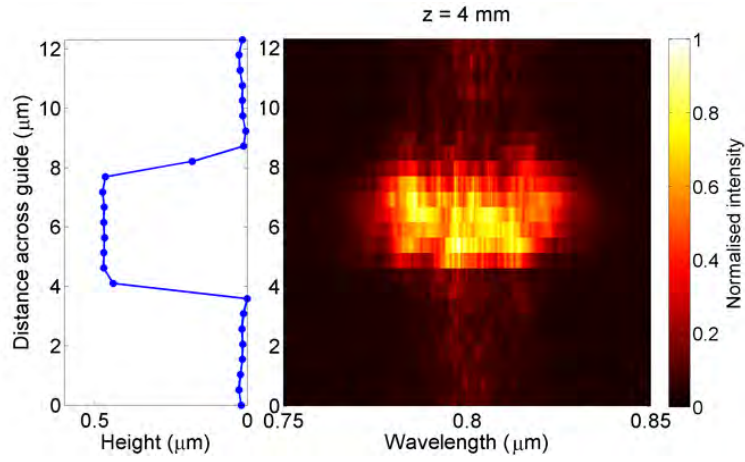
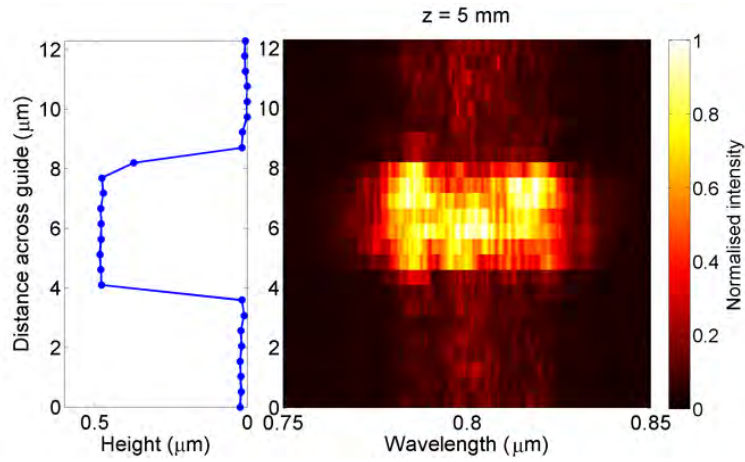


FIGURE 6.15: High-resolution spectral mappings across the waveguide over the distance of $12\ \mu\text{m}$ and their corresponding topographical image. The spectra were collected by the probe translating across the waveguide by the steps of $500\ \text{nm}$ at various positions from the entrance face: (a) $z = 2.2\ \text{mm}$ (b) $z = 3\ \text{mm}$ (c) $z = 4\ \text{mm}$ and (d) $z = 5\ \text{mm}$



(c)



(d)

FIGURE 6.15: High-resolution spectral mappings across the waveguide over the distance of 12 μm and their corresponding topographical image. The spectra were collected by the probe translating across the waveguide by the steps of 500 nm at various positions from the entrance face: (a) $z = 2.2$ mm (b) $z = 3$ mm (c) $z = 4$ mm and (d) $z = 5$ mm (cont)

Additionally, the spectral evolution across the waveguide was studied and the result is shown in Figure 6.15. The SNOM probe was placed at various locations from the front end and then moved across the waveguide to collect the spectra at 500 nm intervals and thus the topography was obtained simultaneously. The spectral image is displayed in wavelength scale over the distance of 12 microns across the waveguide. The left graph shows the corresponding surface topography of the waveguide given by the SNOM probe's height control system. It spatially maps with the spectral image on the right-hand side and the points in the graph are the locations where the probe collected the spectra.

As already described in Section 3.3, Chapter 3, in the shear force technique, the probe is attached with a laterally dithering tuning fork. The oscillation of the probe may affect the spatial accuracy for such high resolution spectral maps, as shown in Figure 6.12 or Figure 6.15. Since the typical oscillating amplitude is of the order of 10^{-1} nm - 1 nm [16–18], the error of the position on the waveguide where the probe collects the spectra can be approximated to be in the range of 0.1% - 1%.

Furthermore, the end faces of the waveguide form a Fabry-Perot etalon which can behave as a spectral filter. The filtering performance might not be effective owing to the low reflectivity of the end faces and therefore low-finesse of the etalon. However, it can have an affect on the output spectra but probably not on the measured spectra from the near-field above the waveguide.

6.6.2 Simulation result

Similarly to what was observed in the experiment, the spectral feature of the simulation result also varies noticeably across the waveguide. Three-mode mixing spectral variations across the central line ($x = 0$) of the waveguide, observed by the probe placed at the position of $z = 4$ mm, are shown in Figure 6.16. The intensity ratio of contributing modes is the same ratio as in Figure 6.7(a). The individual modal contributions are also displayed in Figure 6.16(b) - Figure 6.16(d). Modal interference is clearly apparent in Figure 6.16(a), where both the shift in wavelength range and fluctuations of the spectral linewidth can be observed. In contrast, the spectral features across the guide for the individual modes, as seen in Figure 6.16(b) - Figure 6.16(d), appear to be constant. Note that the shift of individual modal spectra from the central frequency is due to their phase shift in time, relative to the fundamental mode pulse.

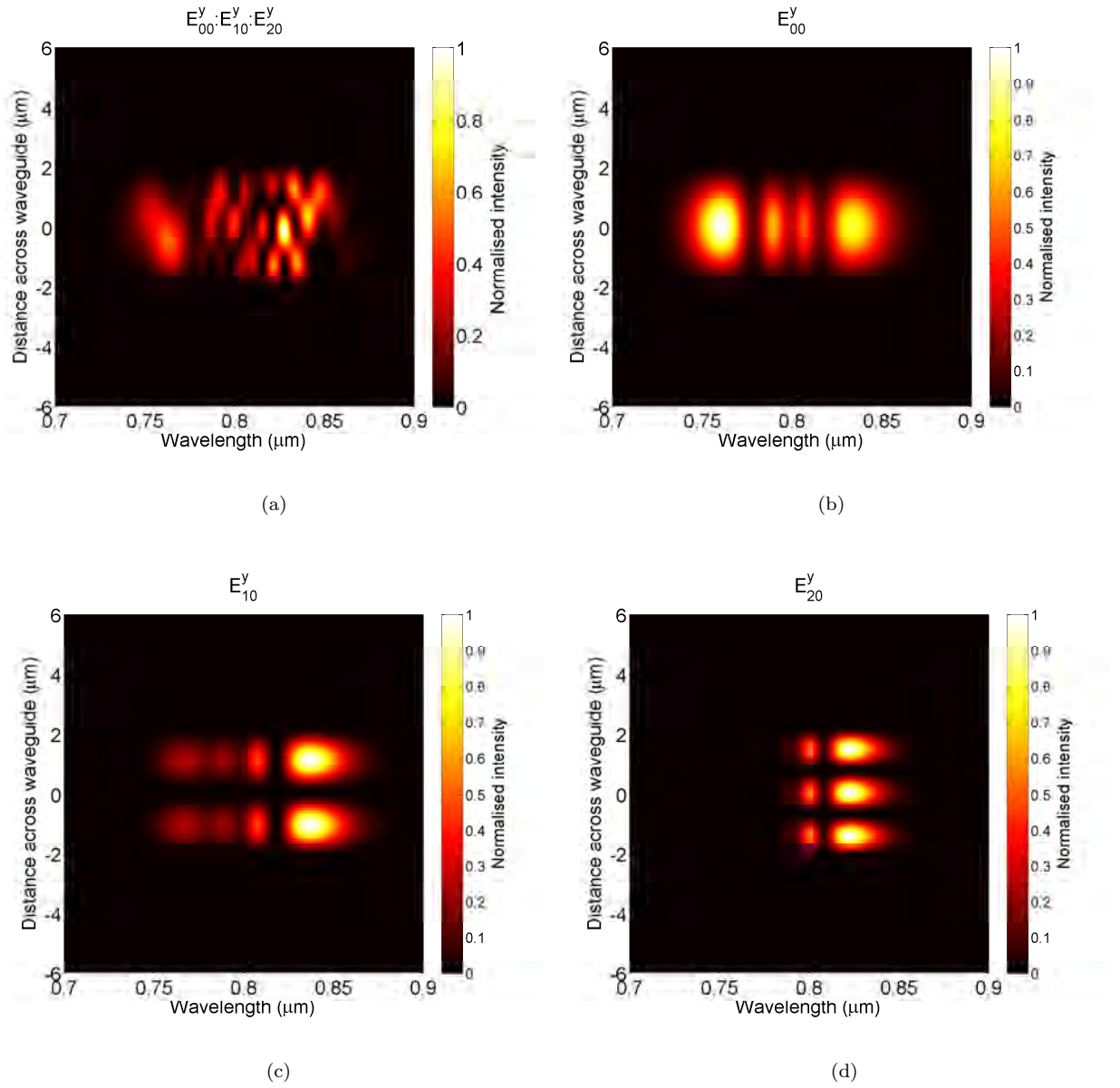


FIGURE 6.16: (a) Simulated spectra through the probe across the centre of the waveguide for the propagation of three-mode pulses. The position of the probe is $z = 4$ mm. The similar spectra of the contributing modes are also displayed: (b) E_{00}^y (c) E_{10}^y (d) E_{20}^y

6.7 Chapter conclusion

The nonlinear spectral development has been investigated in this chapter both experimentally and numerically. The SNOM system was set up and the probe was placed in the near-field region above the waveguide to detect the evanescent field. The field was delivered to the spectrum analyser. The spectral data exhibits irregular features which were not previously observed in a single mode waveguide. Therefore the numerical model, which was developed in the previous chapter, was adopted to simulate the propagation of a multimode pulse. By studying the simulated spectra, the knowledge of the physical background of spectral broadening is attainable.

The experimental results of the spectral evolution along the waveguide displayed complicated spectral features which suggest the effect of the interference of various modes. The assumption is confirmed by the simulation of the propagation of single mode and multimode pulses with various types of mixing. The spectra of the single-mode case contain only the plain feature of the self-phase modulation which is the only nonlinear effect included in the simulation. On the contrary, the multimode spectra possess fine details caused by the interference of various modes with different amounts of linear and nonlinear phase shift. Consequently its degree of complexity rises with the number of modes in the simulation.

The growth of the spectral broadening was also studied. The nonlinear broadening observed in the experiment grows with the propagation distance and the input pulse energy, similarly to that generally observed in the single mode case, as described in many research papers. However, the rate of the growth is retarded at the later distances of the waveguide propagation. By studying the simulated spectral growth along the length of the waveguide, the spectral development at the beginning is shown to be strengthened by the cross-phase modulation, in addition to the self-phase modulation. The phenomenon is decelerated by the modal separation, which is due to the group velocity difference of various modes. At a later distance of propagation, which is long enough to separate the higher-ordered modes from the main pulse in time domain, the remaining lower-ordered modes mainly contribute to the spectral broadening. The degree of this setback depends on the energy distribution among modes. Less significant changes appear in the curve of spectral growth along the distance of propagation when more energy is concentrated in the fundamental mode.

The high resolution of the spectral development was also measured and the data exhibited the subwavelength-scaled variation. The scale of the variation is much smaller than the dispersion and nonlinear lengths and therefore the variation cannot be induced by both effects. Studying the simulation results from multimode pulses containing various categories of modes can link the variation to the effect of phase velocity difference between the symmetric and asymmetric modes.

Finally, the experimental results displayed that the observed spectra also fluctuate across the waveguide. The multimode simulation affirmed the cause of modal interference. The simulated spectra, from the superposition of modes with different transverse modal fields and relatively complicated phase shifts, built up by the linear and nonlinear effects, keep changing at different locations along the width dimension of the waveguide. However, this feature does not appear in the spectra of individual modes.

6.8 References

- [1] R. R. Alfano and S. L. Shapiro. Observation of self-phase modulation and small-scale filaments in crystals and glasses. *Physical Review Letters*, 24(11):592–594, 1970.
- [2] W. L. Smith, P. Liu, and N. Bloembergen. Superbroadening in H_2O and D_2O by self-focused picosecond pulses from a YAlG:Nd laser. *Physical Review A (General Physics)*, 15(6):2396–2403, 1977.
- [3] P. B. Corkum, C. Rolland, and T. Srinivasan-Rao. Supercontinuum generation in gases. *Physical Review Letters*, 57(18):2268–2271, 1986.
- [4] P. B. Corkum, P. P. Ho, R. R. Alfano, and J. T. Manassah. Generation of infrared supercontinuum covering 3–14 μm in dielectrics and semiconductors. *Optics Letters*, 10(12):624–626, 1985.
- [5] J. K. Ranka, R. S. Windeler, and A. J. Stentz. Visible continuum generation in air silica microstructure optical fibers with anomalous dispersion at 800 nm. *Optics Letters*, 25(1):25–27, 2000.
- [6] T. A. Birks, W. J. Wadsworth, and P. St. J. Russell. Supercontinuum generation in tapered fibers. *Optics Letters*, 25(19):1415–1417, 2000.
- [7] C.-Y. Tai, J. S. Wilkinson, N. M. B. Perney, M. C. Netti, F. Cattaneo, C. E. Finlayson, and J. J. Baumberg. Determination of nonlinear refractive index in a Ta_2O_5 rib waveguide using self-phase modulation. *Optics Express*, 12(21):5110–5116, 2004.
- [8] J. D. Mills. *An investigation of phase-mask diffraction patterns and fibre Bragg gratings with scanning near-field optical microscopy*. PhD thesis, University of Southampton, 2001.
- [9] J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, M. E. Zoorob, C. Netti, and J. J. Baumberg. Observation of the developing optical continuum along a nonlinear waveguide. *Optics Letters*, 31(16):2459–2461, 2006.

- [10] J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, C. Netti, M. E. Zoorob, and J. J. Baumberg. Group velocity measurement using spectral interference in near-field scanning optical microscopy. *Applied Physics Letters*, 89(5):051101–1 – 051101–3, 2006.
- [11] M. L. M. Balistreri, A. Driessen, J. P. Korterik, L. Kuipers, and N. F. Van Hulst. Quasi interference of perpendicularly polarized guided modes observed with a photon scanning tunneling microscope. *Optics Letters*, 25(9):637–639, 2000.
- [12] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. John Wiley & Sons Inc., New York.
- [13] T. Chaipiboonwong, P. Horak, J. D. Mills, and W. S. Brocklesby. Numerical study of nonlinear interactions in a multimode waveguide. *Optics Express*, 15(14):9040–9047, 2007.
- [14] W. H. Cheng, S. F. Chi, and A. K. Chu. Effect of thermal stresses on temperature dependence of refractive index for Ta₂O₅ dielectric films. *Thin Solid Films*, 347(1-2):233–237, 1999.
- [15] G. P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, CA, USA, 3rd edition, 2001.
- [16] K. Karrai and R. D. Grober. Piezoelectric tip-sample distance control for near field optical microscopes. *Applied Physics Letters*, 66(14):1842–1844, 1995.
- [17] J. U. Schmidt, H. Bergander, and L. M. Eng. Shear force interaction in the viscous damping regime studied at 100 pN force resolution. *Journal of Applied Physics*, 87(6):3108–3112, 2000.
- [18] J. U. Schmidt, H. Bergander, and L. M. Eng. Experimental and theoretical analysis of shear-force interaction in the non-contact regime with 100 pN force resolution. *Applied Surface Science*, 157(4):295–301, 2000.

Chapter 7

Modal dispersion analysis and other aspects relating to the accuracy of the simulation

7.1 Chapter introduction

In this chapter, the novel technique developed for measuring the group velocity difference will be explained in Section 7.2. The temporal profile of the multimode pulse can be attained from the SNOM-acquired spectral data (which has already been presented in Chapter 6) via the Fourier transform. From the spectral information obtained from various locations along the length of the waveguide, the increasing temporal separation of higher-ordered modes from the fundamental mode becomes apparent. Since the extent of this separation is proportional to the group velocity of the modes, relative to that of the fundamental mode, the relation of temporal separation and propagation distance directly provides the group index difference of the higher-ordered modes and the fundamental mode. The same criteria are also applied to the simulation results and the modal separation is shown to exist in the time domain. As expected, the determination of the group index difference from the simulated time profile gives the group index difference of the contributing modes to be nearly the same value as the theoretical values that were calculated directly from the propagation constants of the modes which have been determined in Chapter 5. The discrepancy of the technique can be caused by the uncertainty in locating the temporal position of the separated modes especially owing to the superposition between adjacent modes.

All the previous simulation results have included the dispersion effect, only up to the second-order term, and therefore the effect of third-order dispersion will be discussed in Section 7.3. By comparing the dispersion length of the third-order term with the dispersion length of the second-order term and the nonlinear length, the strength of

the third-order dispersion is shown to be minimal and any changes in the spectra are insignificant when the term is included in the modeling for the propagation distance in the range of few millimetres. The evidence for this is given by the unobservable deviation of the simulated result, with the addition of the third-order term in the dispersion operator.

Finally the effect of the wavelength dependence of the transverse field will be studied in Section 7.4. In the previous simulation, the transverse field is assumed to be constant and the field distribution at the central wavelength is used in the calculation for the total spectral field over the whole wavelength range. Since the propagation constant actually depends on the wavelength of the propagating light, the mode field distribution also varies with the wavelength. However, in comparison with the previous spectra, including the wavelength dependence in the calculation does not significantly change the spectral features but only changes the height of the peaks in the feature. The extent of the deviation is associated with the spectral bandwidth and, therefore, the deviation grows with the propagation distance when the broadening is strengthened by the dispersion and nonlinear effects.

7.2 Measurement of group velocity difference

7.2.1 Spectral interference

In a multimode waveguide, each mode propagates with its own group velocity. The temporal separation of the modes induces the interference fringes in the spectral feature owing to the superposition of modal spectra with different spectral phase. The phenomenon is demonstrated theoretically in Figure 7.1 when a Gaussian input pulse, at the wavelength 800 nm, induces only two modes E_{00}^y and E_{50}^y with the group velocities 1.70214×10^8 m/s and 1.63936×10^8 m/s respectively. On the left are the time profiles at various propagation distances which start from the beginning of the propagation, $z = 0$ mm in Figure 7.1(a), and go to the distance $z = 4$ mm in Figure 7.1(e) whereas their corresponding spectral profiles are displayed on the right. The shape of the time profile contributes to the envelope of the spectrum, whereas the temporal separation (Δt) between the modes, caused by the different group velocity, introduces the modulation by means of the function $\cos(\omega\Delta t/2)$. Since the modulation length is inversely proportional to Δt , the increasing temporal separation along the length of the waveguide causes the greater number of fringes in the spectra, as can be observed in the figure.

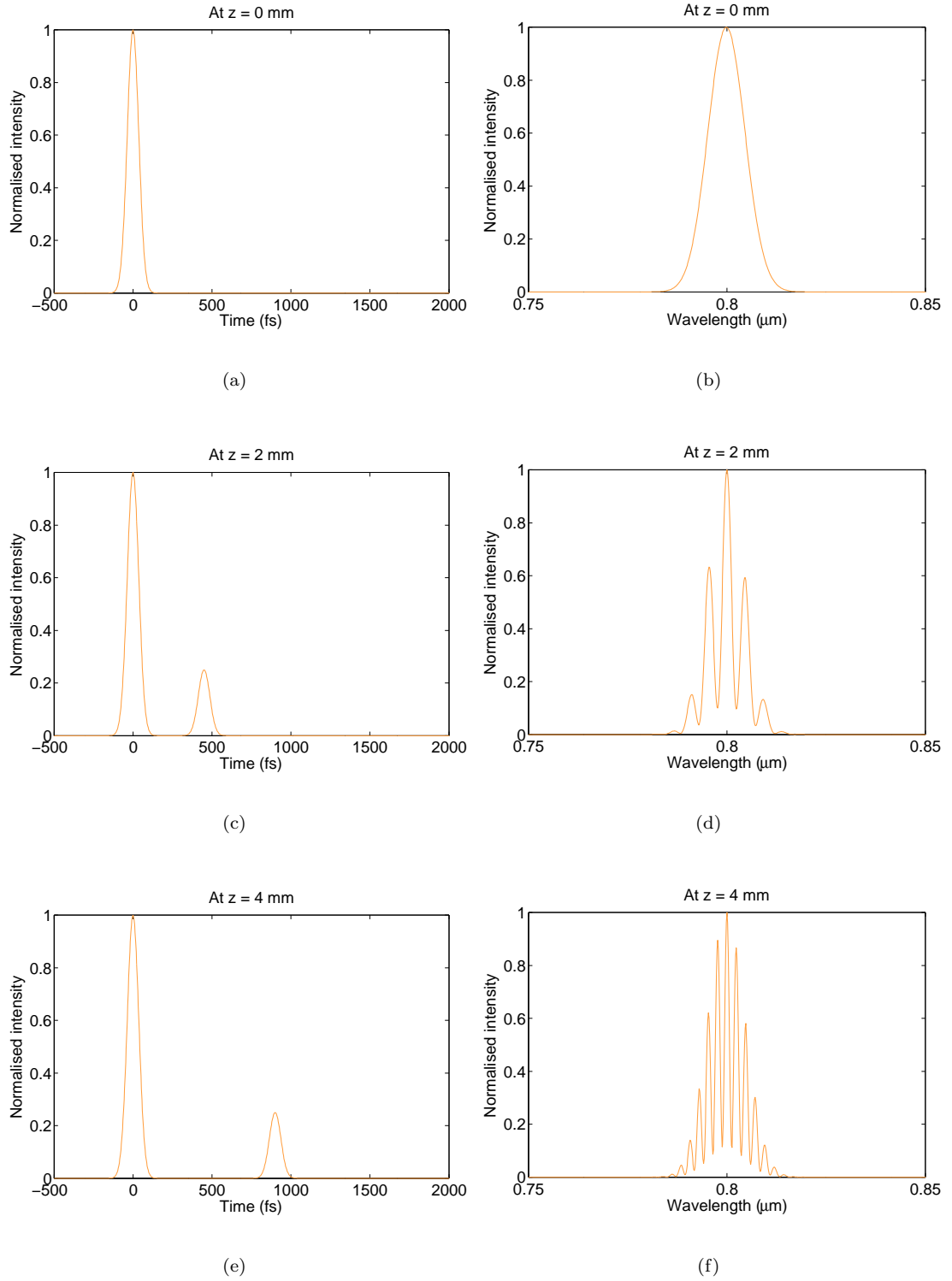


FIGURE 7.1: Spectral interference caused by 2-mode pulse propagation with different group velocities. The time profiles at various locations along the length of the waveguide are shown on the left whereas their corresponding spectra are on the right. The propagation distances are (a)-(b) 0 mm (c)-(d) 2 mm and (e)-(f) 4 mm. The number of fringes in the spectra increases with the propagation distance.

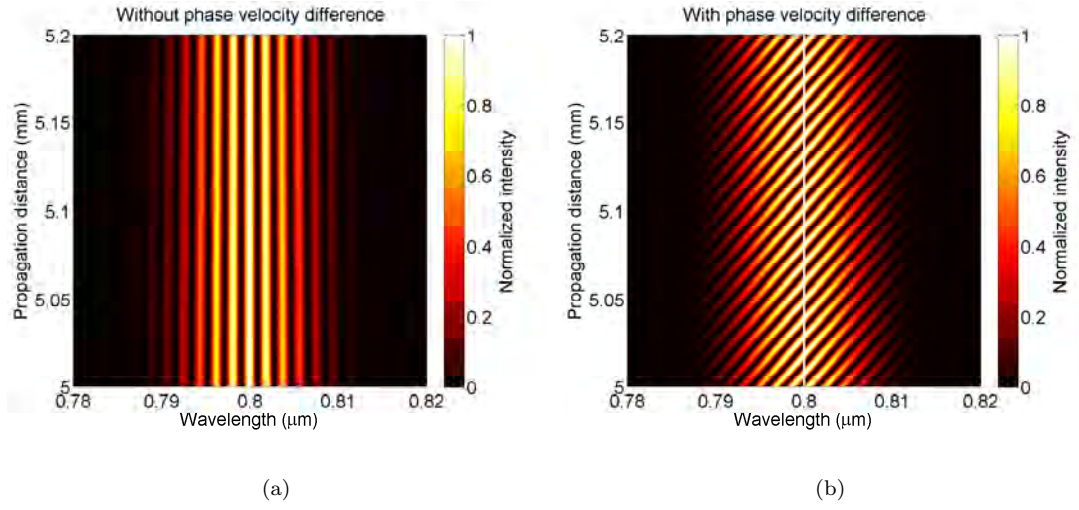


FIGURE 7.2: Spectral variation over the distance of 200 μm along the length of the waveguide after the propagation of 5 mm (a) without and (b) with the phase velocity difference

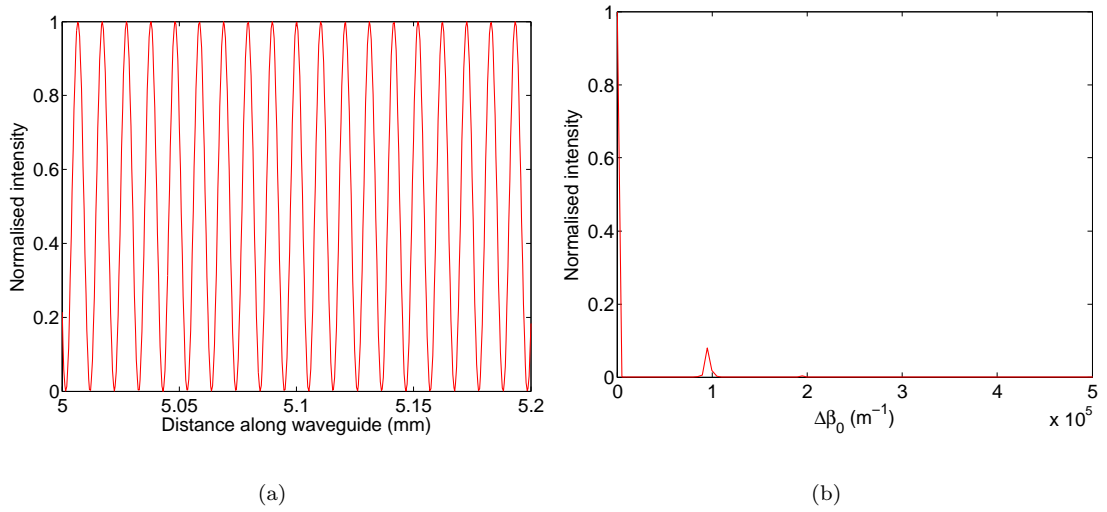


FIGURE 7.3: (a) Spectral intensity variation along the length of the waveguide at the wavelength of 800 nm, marked by the white line in Figure 7.2(b) (b) The Fourier transform of (a)

On the other hand, the phase velocity difference leads to the spatial modulation along the propagation distance. Figure 7.2 displays the spectral variations of the same 2-mode pulse as Figure 7.1 over the range of $200 \mu\text{m}$ after the propagation of 5 mm . Both spectra consists of the interference fringes originated from the group velocity difference of the contributing modes. However only in the spectrum map on the right, which includes the phase velocity difference term, is the spatial modulation observable. If a SNOM probe can be applied to collect the localised field along a planar waveguide, as in recent research [1, 2], the information of phase velocity can be extracted. For example, from the spectral map in Figure 7.2(b), the spectral intensity variation along the waveguide at the central wavelength of 800 nm , the intensity along the white line, is displayed in Figure 7.3(a) in which the modulation feature is clearly seen. On the right, Figure 7.3(b), is the Fourier transform of the variation in the spatial frequency domain. The peak in the figure, which is the modulation frequency of the spatial variation and therefore the phase velocity difference, is at the position $\frac{(\beta_0^{(0,0)} - \beta_0^{(5,0)})}{2\pi} = 9.6464 \times 10^4 \text{ m}^{-1}$.

On the contrary, in reality the waveguide contains many more modes each of which has the spectrum broadened due to the nonlinearity of the guide. As a result, the superposition of all these provides the complex interference feature and spatial variation of the spectral intensity, as already observed in the spectra collected by the SNOM in Chapter 6. The Fourier transform of the spatial variation can still give the phase velocity difference of all the modes. However the number of location points of the experimental spectra is so small that the Fourier transform cannot provide accurate results. On the contrary, the extraction of the group velocity information embedded in the localised spectra can be performed and will be explained in the next section, including the comparison with the results derived using the same criteria applied to the simulation spectra.

7.2.2 Experimental result

The manipulation of the Fourier transform on the measured spectra in Figure 6.5, Chapter 6, which displays the spectral development along the length of the waveguide, can implement the visualisation of the propagating of a pulse in time inside the waveguide at various locations. The spectral data, straightforwardly derived from the spectrometer, is in the wavelength scale (which is reciprocal to the frequency scale) whereas the frequency is directly related to the time scale via the Fourier transform. Direct transformation from the wavelength scale to the frequency scale by the relation $f = c/\lambda$, where f , c and λ are the frequency, speed of light in vacuum and wavelength, respectively, will make the spectral data unevenly distributed and cause a problem in defining the correct corresponding time scale. Therefore an interpolation process was utilised to obtain interpolated spectra with an evenly spaced frequency scale, which covers the same frequency range $3.5081 \times 10^{14} - 4.0293 \times 10^{14} \text{ Hz}$ as that of the measured data, with the

data size $N = 2^9$ and frequency step $\Delta f = 1.0200 \times 10^{11}$ Hz. The interpolation process is achieved by the MATLAB function of cubic spline interpolation. Both the experimental data (from Figure 6.5(a), Chapter 6) and the interpolated spectra, for various locations along the length of the waveguide, are displayed in Figure 7.4. The interpolation fits with the data so well that the deviation is hardly noticeable.

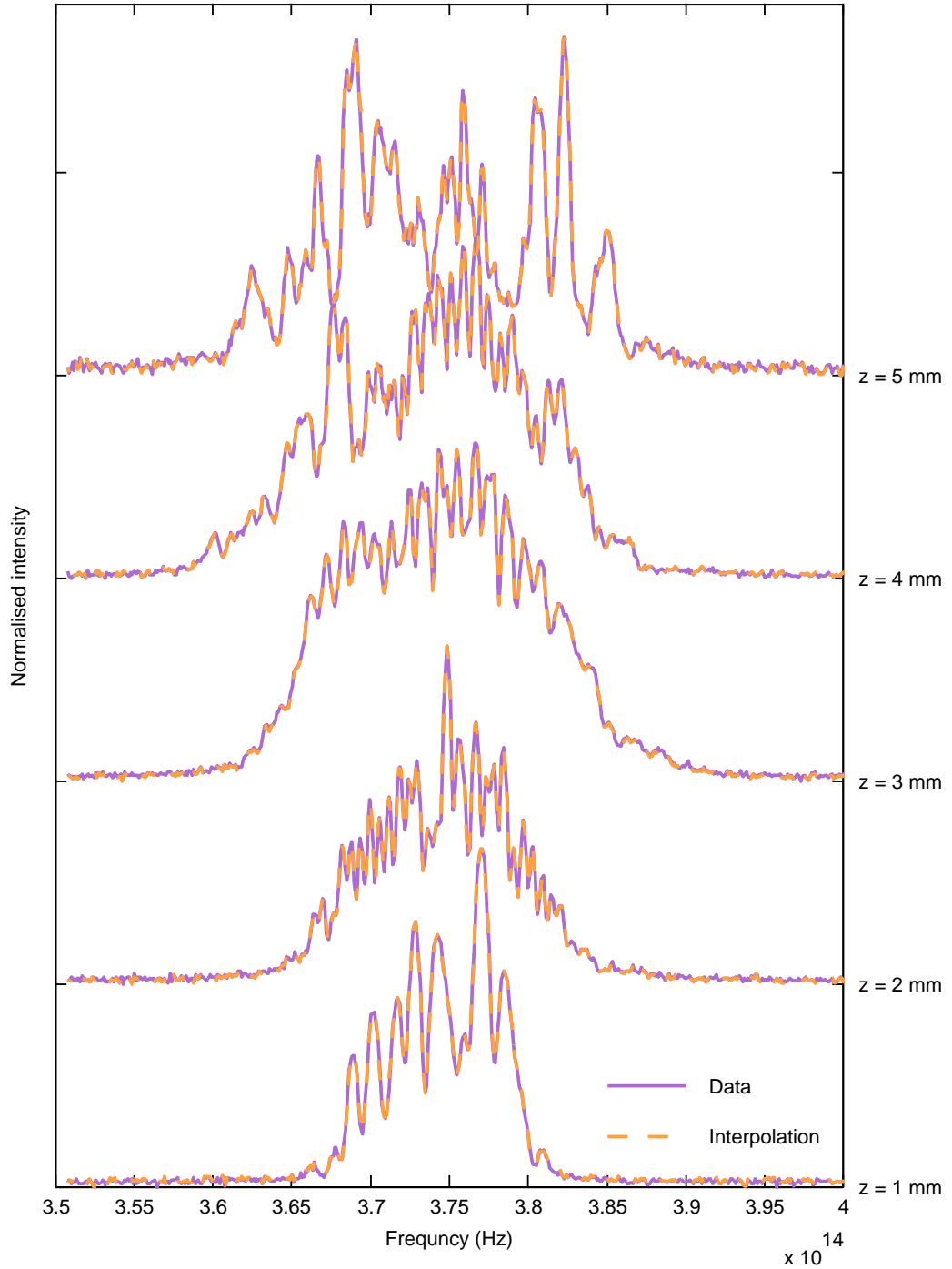


FIGURE 7.4: Measured spectral data at various locations along the length of the waveguide (solid violet curves) in comparison with the interpolated data (dashed orange curves) in the new evenly-spaced frequency scale

Then the inverse Fourier transform was applied to the interpolated spectra by the inverse discrete Fourier transform (DFT) function of MATLAB. The time scale t can be determined from

$$t = \frac{1}{\Delta f} \sum_{p=-N/2}^{+N/2-1} \frac{p}{N} \quad (7.1)$$

The outcome is the corresponding time profile with the time range from -4.9021 ps to 4.8830 ps and the time resolution $\Delta t = 19.1490$ fs. The inverse Fourier transform results, from the spectra of three different input pulse energies from 0.8 nJ to 2.1 nJ in Figure 6.5(a), Chapter 6, are displayed in Figure 7.5. The consistency of the data is evident as the minor peaks in three different curves are situated at the same proximity in time. These peaks are the higher-ordered modes and separate themselves from the main pulse due to their different group velocities. The separations of three higher-ordered modes are clearly noticeable and the dashed lines are added to mark their translation in time at various propagation distances. For future reference, these modes are labeled as (a), (b) and (c), as shown in the figure.

Obviously the pulse separation increases with the distance of propagation. Its temporal interval is proportional to the relative group velocity difference of the modes to the fundamental one, and can be written as

$$\begin{aligned} (\Delta t)_{mn} &= \frac{s}{(v_g)_{mn}} - \frac{s}{(v_g)_{00}} \\ &= \frac{(\Delta n_g)_{mn}}{c} s \end{aligned} \quad (7.2)$$

where $(\Delta t)_{mn}$ is the temporal separation of the mode E_{mn}^y from the fundamental mode E_{00}^y at the distance of propagation s , $(v_g)_{mn}$ is the group velocity of the mode E_{mn}^y , $(v_g)_{00}$ is the group velocity of the mode E_{00}^y , and $(\Delta n_g)_{mn}$ is the group index difference, which is defined as

$$(\Delta n_g)_{mn} = c \left(\frac{1}{(v_g)_{mn}} - \frac{1}{(v_g)_{00}} \right) \quad (7.3)$$

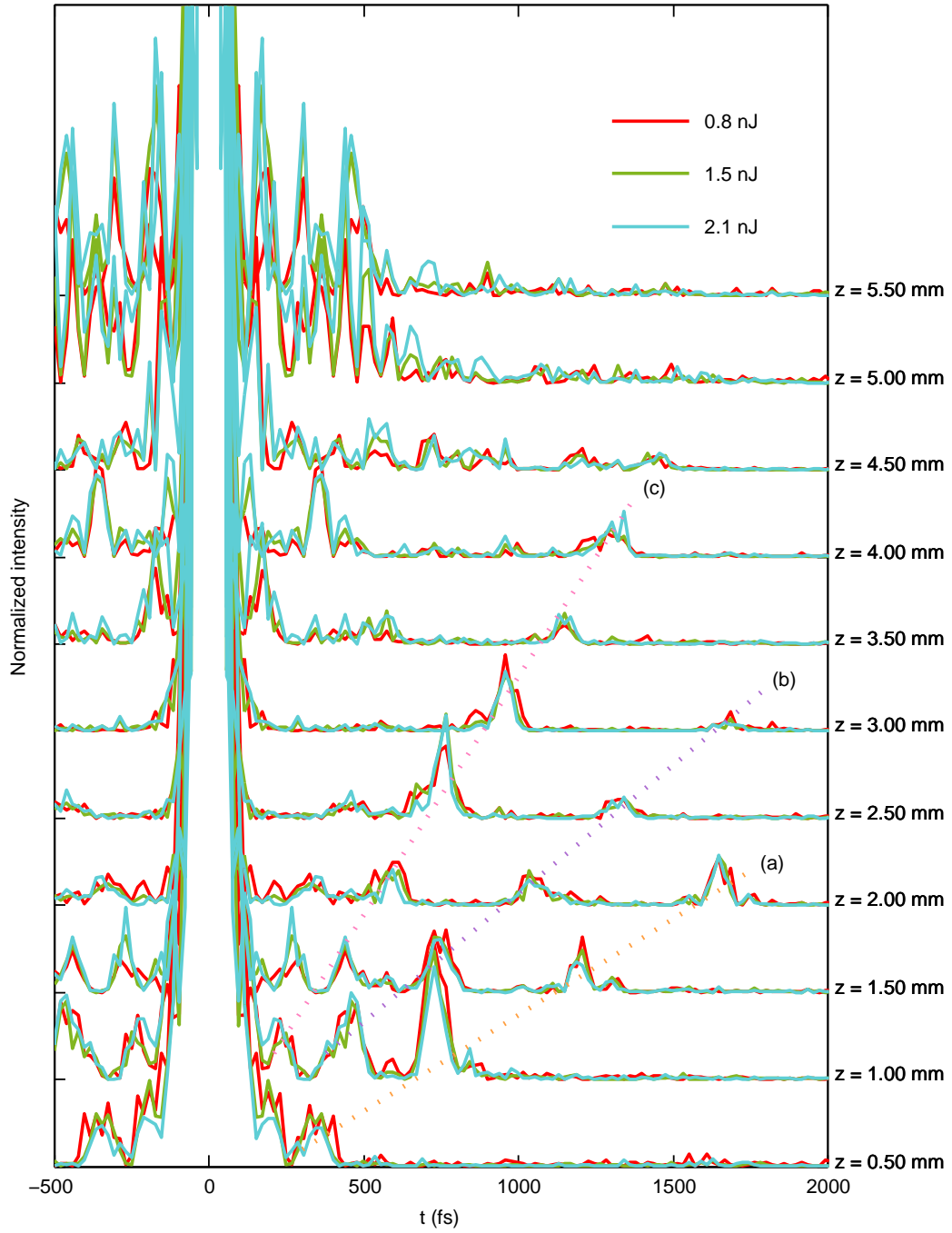


FIGURE 7.5: Fourier transform of spectra along the waveguide at different input energies of 0.8 nJ (red curves), 1.5 nJ (green curves) and 2.1 nJ (blue curves). The separations of higher-ordered modes can be observed with the dashed lines (a), (b) and (c) marking the temporal positions of three different modes at various locations along the waveguide. The details of temporal separations and distances of these three modes will be later shown in Figure 7.6 [3].

Therefore the measurement of time separation of the minor peaks in Figure 7.5 from the central one, at various propagation distances, can give the information of the group index difference of various modes. Figure 7.6 shows the values of time separation of the modes (a), (b) and (c) plotted against the distance of propagation. The polynomial curve fitting, based on the least square algorithm, was applied to determine the line equation in which the slope provides the value of $(\Delta n_g)_{mn}/c$ from the relation of s and $(\Delta t)_{mn}$ in Equation 7.2. These interpolating lines are also included in the figure and the value of Δn_g and the coefficient of determination (R^2) of all three lines are summarised in Table 7.1. The interpolation fits well with the experimental data, as can be confirmed by the value of R^2 close to 1.

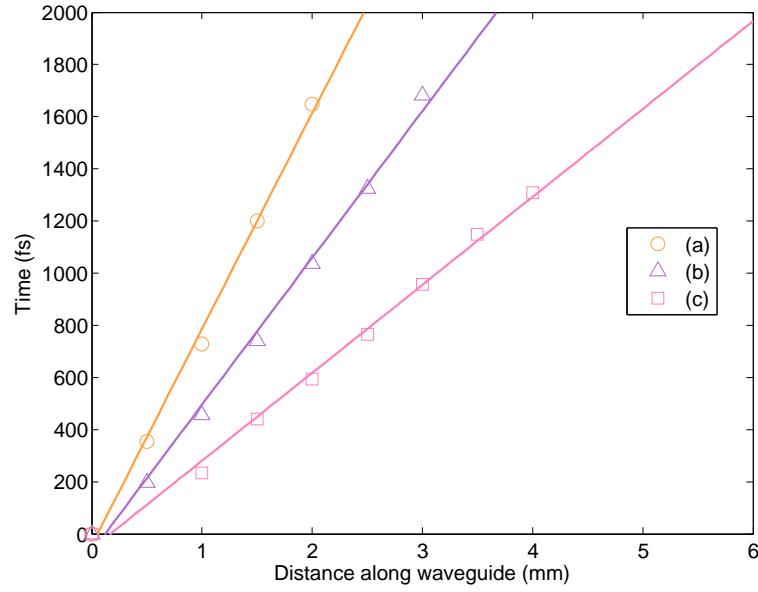


FIGURE 7.6: The separation in time of the mode (a), (b) and (c) in Figure 7.5 at various propagation distances

Mode	Curve	Δn_g (m/s)	R^2
(a)	—	0.2484	0.9962
(b)	—	0.1689	0.9946
(c)	—	0.1012	0.9948

TABLE 7.1: Group index difference (Δn_g) of some higher-ordered E^y modes whose separation from the main pulse can be observed in Figure 7.5

On the other hand, the group index difference of the calculated propagation modes in Section 5.3, Chapter 5, can be determined from

$$(\Delta n_g)_{mn} = c((\beta_1)_{mn} - (\beta_1)_{00}) \quad (7.4)$$

where $(\beta_1)_{mn}$ is the first-order term of the Taylor expansion of the propagation constant, at the wavelength 800 nm, and reciprocal to the group velocity $(v_g)_{mn}$. The calculated value of Δn_g for some higher-ordered E^y modes is displayed in Table 7.2.

Mode	Δn_g	Mode	Δn_g
E_{10}^y	0.00549	E_{01}^y	0.07661
E_{20}^y	0.01476	E_{11}^y	0.08487
E_{30}^y	0.02798	E_{21}^y	0.09891
E_{40}^y	0.04544	E_{31}^y	0.11913
E_{50}^y	0.06749	E_{41}^y	0.14616
E_{60}^y	0.09465	E_{51}^y	0.18091
E_{70}^y	0.12755	E_{61}^y	0.22464
E_{80}^y	0.16706	E_{71}^y	0.27912
E_{90}^y	0.21434	E_{81}^y	0.37346
$E_{10,0}^y$	0.27101		
$E_{11,0}^y$	0.42750		

TABLE 7.2: Group index difference (Δn_g) of some higher-ordered E^y modes calculated from the dispersion curves that are determined by the effective index method

By inspecting the results from the calculation and the experimental measurement, the value of Δn_g for the modes (a), (b), (c) in Table 7.1 are seen to be close to those of the modes E_{61}^y , E_{80}^y , E_{21}^y with percentage differences of 10.5769%, 1.1014%, and 2.3152% respectively. The origin of the error can be attributed to the uncertainty in determining the time position of the high-ordered mode pulses. Firstly, the time resolution limits the precision by the error of 1% - 3%. This error can be reduced by increasing the time data size. However, the uncertainty in determining the temporal position of the separated higher-ordered mode pulses is mainly caused by the overlapping of adjacent modes that have a close value of group velocity. Even though their group velocity difference, relative to the fundamental mode, is significant enough that they are absolutely separated from the main pulse at some distance of propagation, their group velocity difference with respect to each other is still small such that only the superposition of them is present. From Table 7.2, one pair of such modes is E_{61}^y and E_{90}^y . At the propagation distance of 2 mm, their group velocity difference gives the temporal separation between two modes of 68.6667 fs. Such a range of separation, which is less than the pulse width, cannot allow the modes to be resolvable. Therefore the identification of these modes in time can be ambiguous. This probably results in the large percentage difference of the group index difference from the mode (a) in Figure 7.5, which is determined within the range of only the first few millimetres of propagation distance.

7.2.3 Simulation result

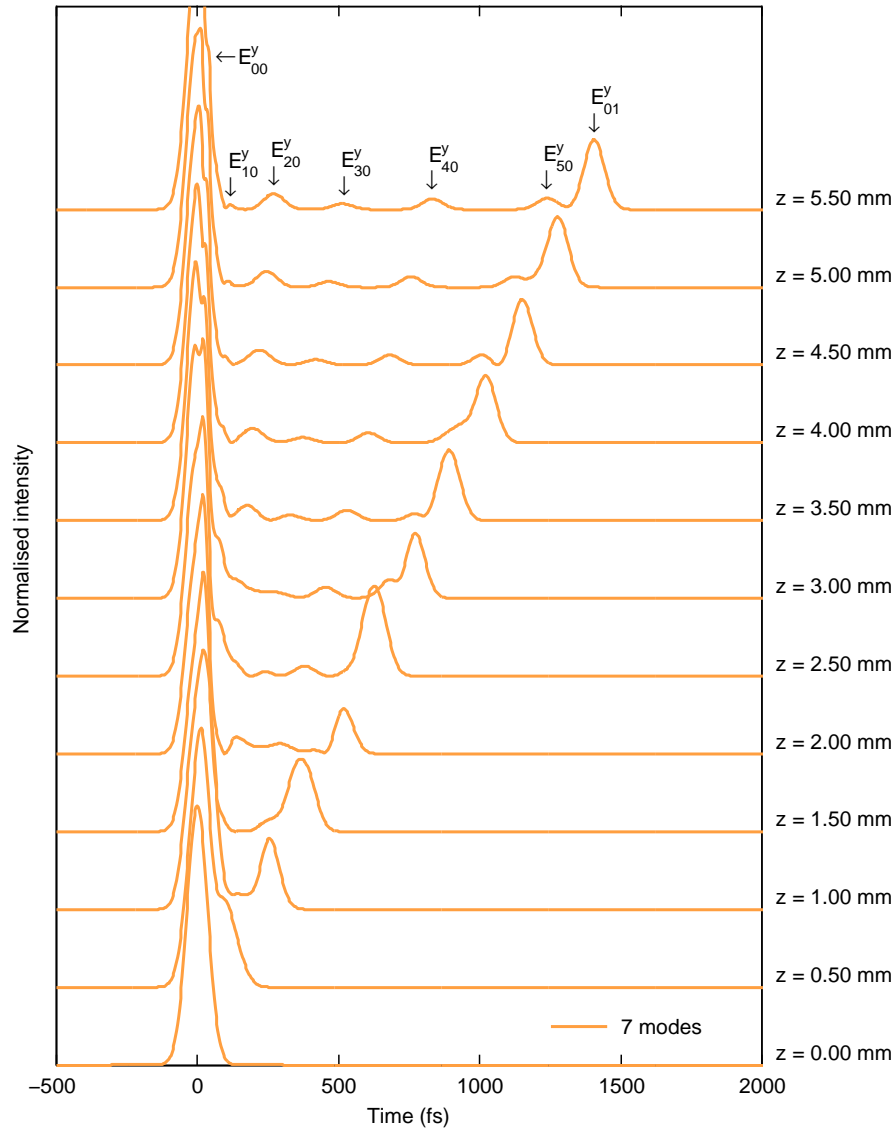


FIGURE 7.7: Simulated time profile at various propagation distances for the spectrum of the 7-mode pulse in Figure 6.7(a), Chapter 6 [4]

Similarly to that done with the experimental spectra, the inverse Fourier transform is applied to the simulation result in Figure 6.7(a), Chapter 6. The simulated spectrum of the 7-mode pulse was collected by a SNOM probe placed at the location 20 nm above the top surface and the lateral position shifted by 200 nm from the vertical central line of the waveguide. The corresponding temporal evolution of the pulse for 7-mode mixing is shown in Figure 7.7. The time profile in this example shows the pulse separating into its various modes along the length of the waveguide, which is similar to what has already been seen in Figure 7.5. Without taking account of absorption loss in the waveguide, the peak height of these higher-ordered mode pulses is expected to be constant and its value should be related to the relative intensity ratio of the mode mixing. On the

contrary, the fluctuation of peak height is clearly noticeable and such a fluctuation also appears in the experimental result in Figure 7.5. This feature is similar to what can be observed as the result of light interference where the intensity varies with the distance. The vast change of intensity occurs in the farthest peak in the first half length of the waveguide. This is owing to the interference of E_{50}^y and E_{01}^y . However, when E_{01}^y starts to separate, at the distance of 3 mm, the effect of interference is subdued and the peak height becomes stable.

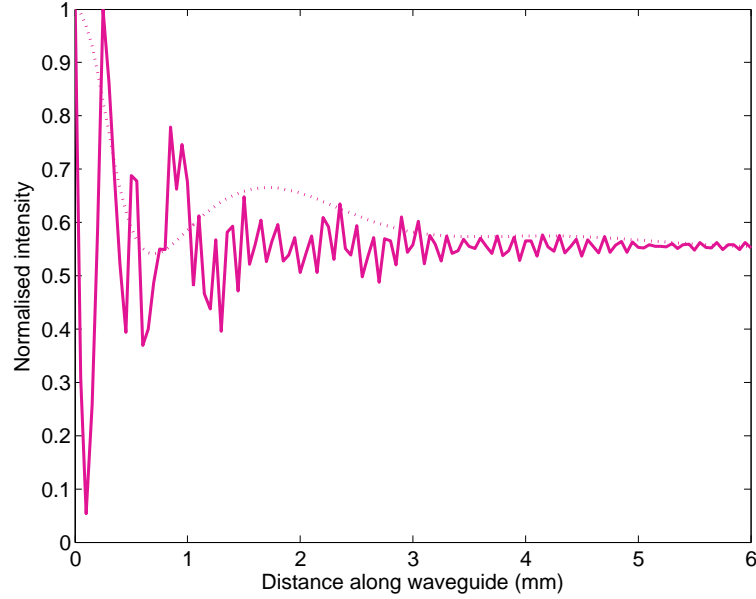


FIGURE 7.8: Relative integrated intensity collected by the SNOM probe in the simulation. The dashed curve displays the envelope of the variation when the phase velocity difference term is removed.

The effect of interference also results in the alternating intensity collected by the SNOM probe, as shown in Figure 7.8. The solid curve is the simulated data of the integrated intensity over time, collected by the probe at different locations along the length of the waveguide. The huge intensity variation, especially within the first few millimetres of the propagation, is caused by the mode field interference, which becomes less predominant when a greater number of higher-order modes separate themselves from the main pulse. Consequently, the curve tends to settle towards the end of the guide. The rapid spatial variation is due to the phase modulation caused by phase velocity difference among the modes. The dashed curve shows the envelope of the variation when the effect of phase velocity difference is neglected.

In the same way as previously demonstrated with the time profile of the experimental data, the time separation of various modes from the central peak is measured and plotted against the distance of propagation, as shown in Figure 7.9. The group index difference is determined from the slope of the linear least-square fitting, which is also displayed as lines in the figure. The values of Δn_g and R^2 are shown in Table 7.3. Compared with

those calculated values in Table 7.2, the determination of the group index difference from the time profile provides the percentage error in the range of 0.0880% - 0.4065%. Obviously the degree of error is much smaller than that from the experimental result. This is owing to the fact that the simulation result is not subject to experimental noise and the number of contributing modes is much less than the real waveguide. As a result, the modal separation can be clearly identified.

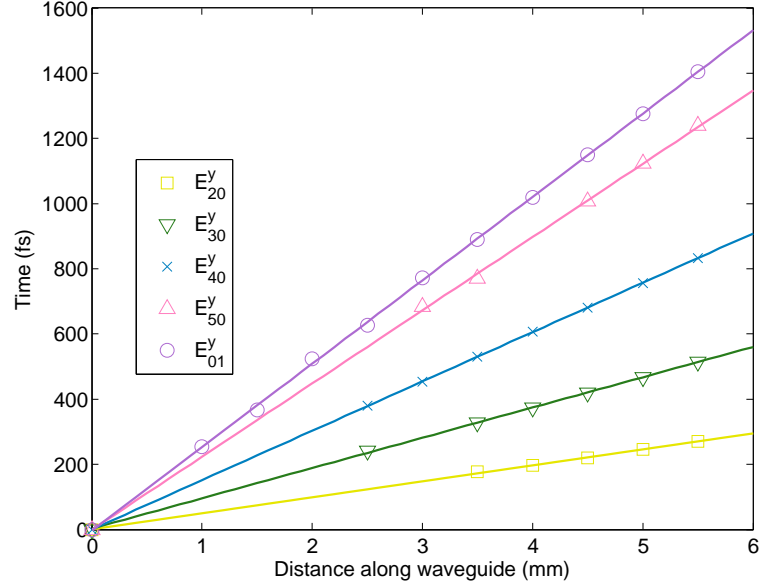


FIGURE 7.9: Time separation from the fundamental mode of various higher-ordered modes in Figure 7.7 against the distance of propagation

The extraction of the group velocity information from the simulation results not only testifies to the correctness of the simulation but also describes the intensity variation of separated higher-ordered mode peaks and the inaccuracy of locating their temporal positions.






Mode	Line	Δn_g (m/s)	R^2
E_{20}^y		0.0147	0.9996
E_{30}^y		0.0279	0.9997
E_{40}^y		0.0454	1.0000
E_{50}^y		0.0674	0.9996
E_{01}^y		0.0768	0.9997

TABLE 7.3: Group index difference Δn_g of E_{20}^y , E_{30}^y , E_{40}^y , E_{50}^y and E_{01}^y whose separation from the main pulse can be observed in Figure 7.7

7.2.4 The effect of nonlinearity on the measurement

The nonlinearity is hardly expected to play any important roles in the determination of the group velocity difference. As has already been shown in Figure 7.5, the locations of separated higher-ordered mode peaks are at the same positions at different energies. Since the nonlinearity directly depends on the input power, it can be assumed that the effect does not cause any significant changes in the measurement of group velocity difference. The assumption is confirmed by the comparison between the simulation performed when both dispersion and nonlinear effects are included in the numerical model and when the nonlinear effect is nullified, as shown in Figure 7.10. The simulations are the propagation of a 7-mode pulse with the same contributing modes and relative intensity ratio as that in Figure 6.7, Chapter 6. The probe in the simulation is at the position of $z = 5.5$ mm. Although in Figure 7.10(a) the spectra, observed by the probe, from both cases consist of the complex interference feature, an enormous difference between those is clearly seen and only large broadening is present in the case of the inclusion of the nonlinearity. However the time profiles, shown in Figure 7.10(b), display insignificant deviation and the separated higher-ordered peaks from both cases locate at the same position.

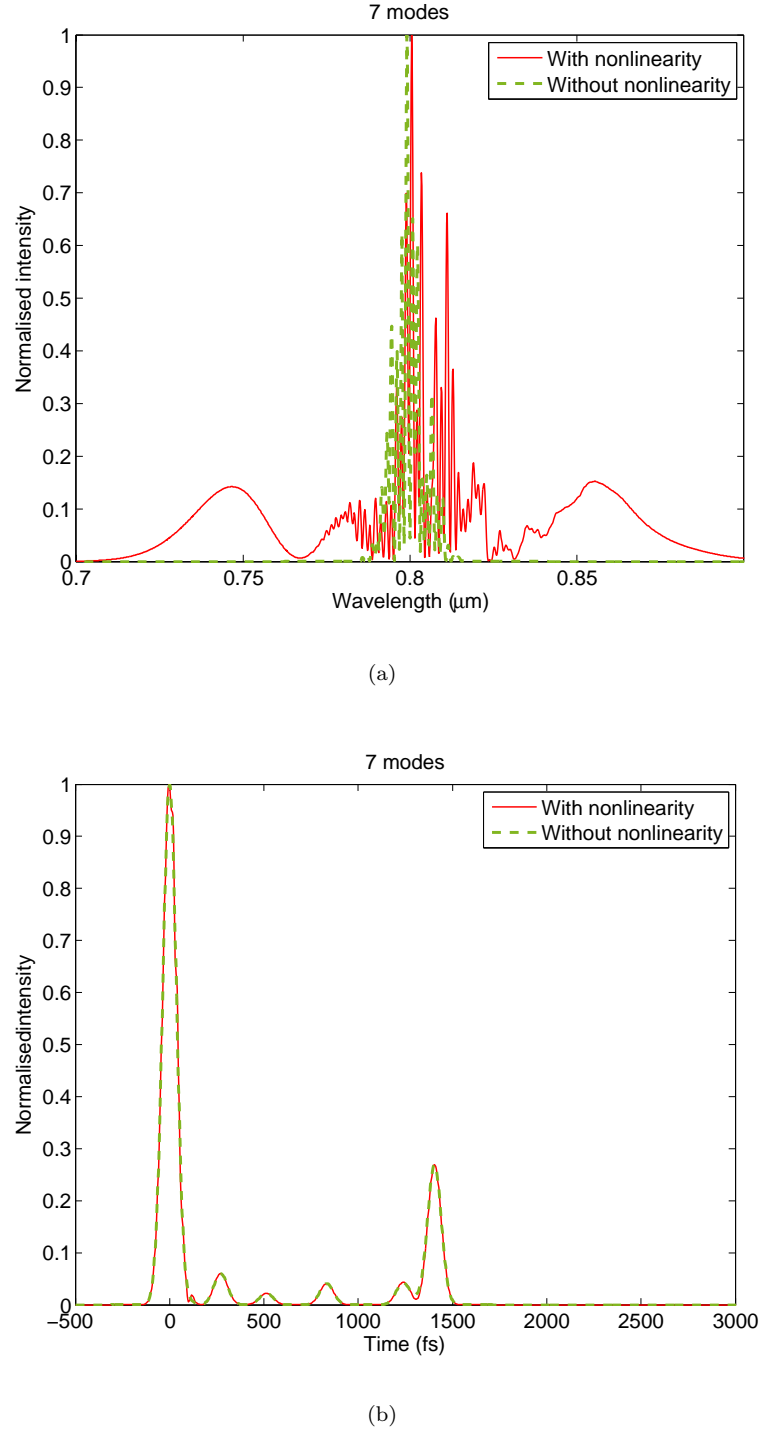


FIGURE 7.10: Comparison between the simulations of a 7-mode pulse propagation with the same contributing modes and relative intensity ratio as that in Figure 6.7, Chapter 6. Two cases of the simulations are that one includes both dispersion and nonlinear effects (solid red curves) and the other excludes the nonlinear term (dashed green curves): (a) simulated spectra observed by the probe which is placed at the position of $z = 5.5$ mm and (b) their corresponding time profiles.

7.3 The effect of third-order dispersion

All the simulation results, which have been discussed so far, are generated by the model containing the dispersion terms up to the second-order. The accuracy of the simulation can be enhanced by adding the third-order term β_3 in the dispersion. The input pulse in this research is in the femtosecond regime, which makes the fraction $\Delta\omega/\omega_0$ become more significant. In such a case, neglecting the higher-order terms further reduces the accuracy of the approximation and the next higher-order term, namely the third-order term, is probably required to be included in Equation 5.16, Chapter 5. Moreover, some low-ordered E^y modes of the waveguide have a zero-dispersion wavelength in the proximity of the pump wavelength, 800 nm, as can be observed in Figure 5.10, Chapter 5. In the absence of GVD, third-order terms become predominant for the dispersion effect. However this can be compensated by a much higher strength of nonlinear effect.

The efficiency of these contributing effects on the nonlinear pulse propagation can be quantified in terms of the length scale over which the dispersion and nonlinear effects become predominant. By inspecting the NLS equation, the strength of the dispersion term is proportional to the dispersion parameter and the derivatives of the field with respect to time. Therefore the dispersion length is defined as [5]

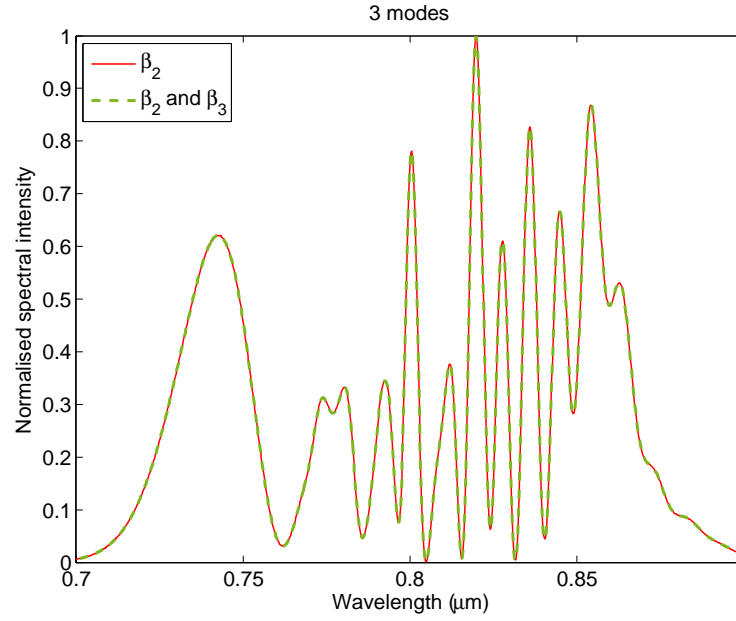
$$L_D = \frac{T_0^2}{\beta_2} \quad (7.5)$$

$$L'_D = \frac{T_0^3}{\beta_3} \quad (7.6)$$

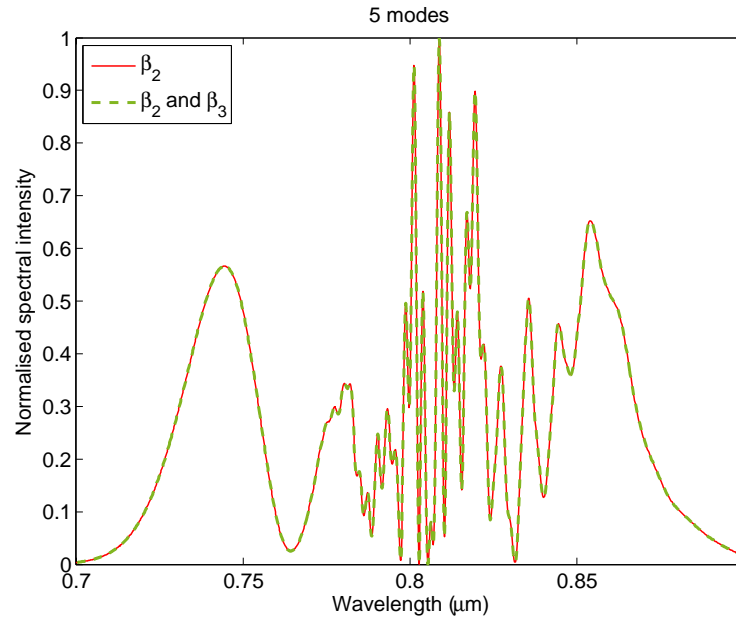
where T_0 is the half-width of the pulse at the intensity $1/e$ of the maximum value, β_2 and β_3 are the second-order and third-order dispersion constants respectively. Since the pulse width, Δt_{FWHM} , of the simulated input pulse is 85.707 fs, the value of $T_0 = \Delta t_{\text{FWHM}}/2(\ln 2)^{1/2}$ is 51.47 fs. For the contributing modes of the previous simulated spectra in Figure 6.7, Chapter 6, the value of $|\beta_2|$ is of the order of $10^{-27} - 10^{-24}$, whereas $|\beta_3|$ is around $10^{-40} - 10^{-39}$ and therefore the dispersion length $L_D \sim 1$ mm to 1 m and $L'_D \sim 1$ cm to 10 cm. On the other hand, the nonlinear term in the NLS equation grows with the nonlinear parameter and the optical power and as a result the nonlinear length is defined by

$$L_{NL} = \frac{1}{\gamma P_0} \quad (7.7)$$

where P_0 is the peak power of the input pulse and γ is the nonlinear parameter. The value of nonlinear parameter γ , shown in Table 5.4, Chapter 5, and the peak power $P_0 \sim 0.7$ kW will result in the value of the nonlinear length $L_{NL} \sim 0.2$ mm to 0.7 mm.



(a)



(b)

FIGURE 7.11: The simulated spectra observed by the probe which is placed at the position of $z = 5.5$ mm for (a) 3 modes (b) 5 modes. The result when the dispersion consists only of the second-order term is shown by the solid red curve whereas the dashed green curve is for including the third-order term. The intensity ratio is the same as those in Figure 6.7, Chapter 6. Both curves are totally superimposed and their deviation cannot be observed.

For the length of the waveguide of the order of 6 mm, the nonlinear effect is stronger than the others by at least one order of the length scale. Therefore the additional third-order term in the dispersion effect is not expected to make any noticeable changes. This assumption can be certified by the comparison between the simulations, including the dispersion up to the second-order and third-order terms.

Figure 7.11 displays the simulated spectra observed by the probe, placed at the position of $z = 5.5$ mm, for the propagation of 3-mode and 5-mode pulses. The details of the contributing modes are the same as those in Figure 6.7, Chapter 6. The solid red curves represent the results when the dispersion effect is taken into account up to the second-order term, β_2 , whereas the dashed green curves are the simulation results when the dispersion effect includes the additional third-order term β_3 . Both curves are totally superimposed and their deviation cannot be observed.

7.4 The effect of wavelength-dependent transverse field distribution

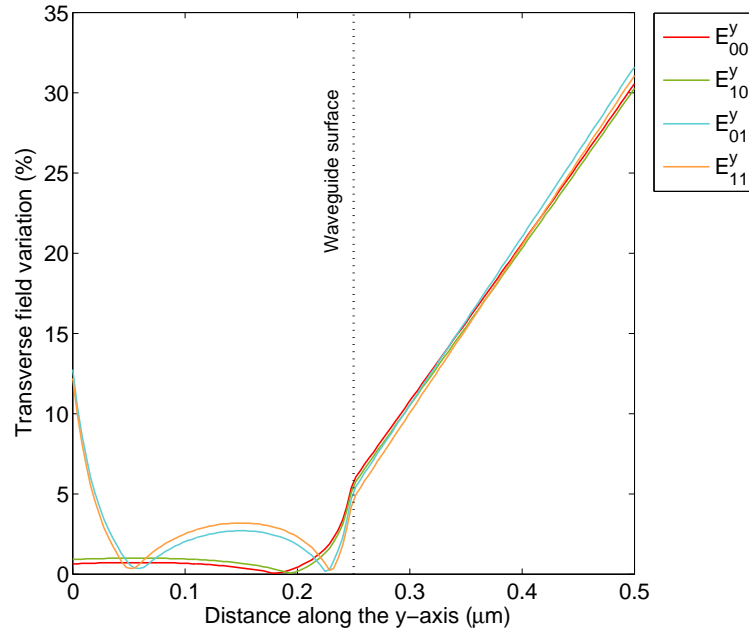


FIGURE 7.12: Variation of transverse field over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{01}^y (blue), and E_{11}^y (orange) at different distances from the centre of the waveguide along the line $x = 200$ nm parallel to the height dimension

Transverse field distributions of various modes are actually wavelength-dependent and the effect of this dependency will be discussed in this section. Figure 7.12 displays the percentage field variation of the transverse field for various modes, over the wavelength range 700 - 900 nm, which covers the wavelength range of the study spectra. The average field variation, from that of the central wavelength 800 nm, is calculated by

$$\bar{\epsilon}(x, y) = \frac{1}{N} \sum_{p=1}^N \left| \frac{U(x, y, \lambda_p) - U(x, y, \lambda_0)}{U(x, y, \lambda_0)} \right| \quad (7.8)$$

where N is the wavelength data size and $U(x, y, \lambda_p)$ and $U(x, y, \lambda_0)$ are the transverse fields for the wavelength λ_p and the central wavelength λ_0 . The variation is plotted against the distance along the vertical line $x = 0.2 \mu\text{m}$ and starts from the centre of the waveguide ($y = 0$) and extends to the evanescent region ($y > +0.25 \mu\text{m}$). The horizontal shift of the vertical axis by 200 nm is the same as the position where the simulated SNOM probe was placed for the results displayed in Figure 6.7(a), Chapter 6. The features of the variation can be categorised into two types according to the mode symmetry around the central line $y = 0$. Within the waveguide $y = 0$ to $y = +0.25 \mu\text{m}$, the variation is less than 1% for symmetric modes. More fluctuation of the field, with a value in the range of less than 10%, can be observed in asymmetric modes which consist of one node in their field distribution along the y -direction and have the studied wavelength range close to their cut-off. The variation turns to increase abruptly in the area near the surface of the waveguide and then grows linearly in the evanescent region for all types of modes. At the position $y = +0.27 \mu\text{m}$, which is the position of the SNOM probe in the previous simulation results, the variation of the field is around 7%.

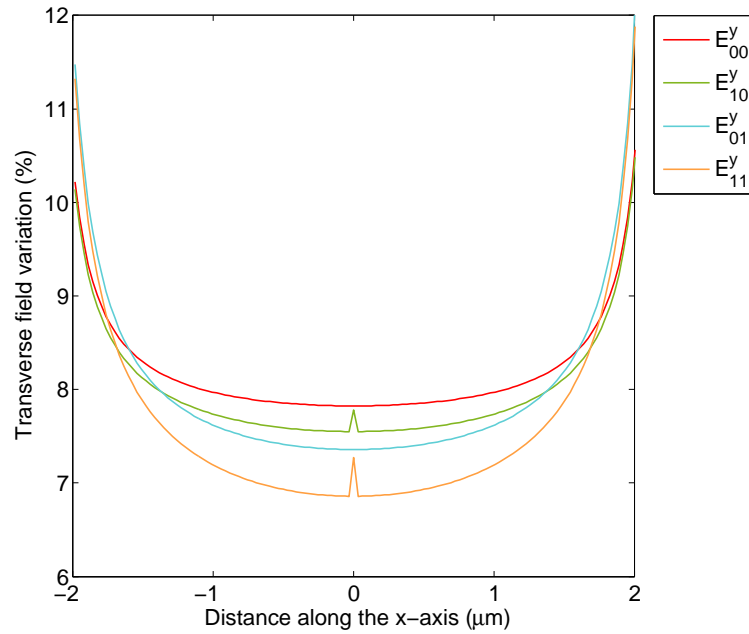


FIGURE 7.13: Variation of transverse field over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{01}^y (blue), and E_{11}^y (orange) at different distances along the $y = 0.27 \mu\text{m}$ parallel to the horizontal dimension of the waveguide

On the other hand, Figure 7.13 displays the transverse field variation along the horizontal line $y = +0.27 \mu\text{m}$, where the SNOM was supposed to be situated. In contrast to the field variation along the height dimension, the wavelength dependency of the field

distribution changes much more slowly over the distance range $x = -1 \mu\text{m}$ to $x = +1 \mu\text{m}$ over which the transverse field variation is 7% - 8%. In this region the variation in higher-ordered modes is lower than the fundamental mode. The discontinuity around the centre $x = 0$ occurs in the curves of E_{10}^y and E_{11}^y owing to the node at the centre in their field distribution, where the field becomes zero.

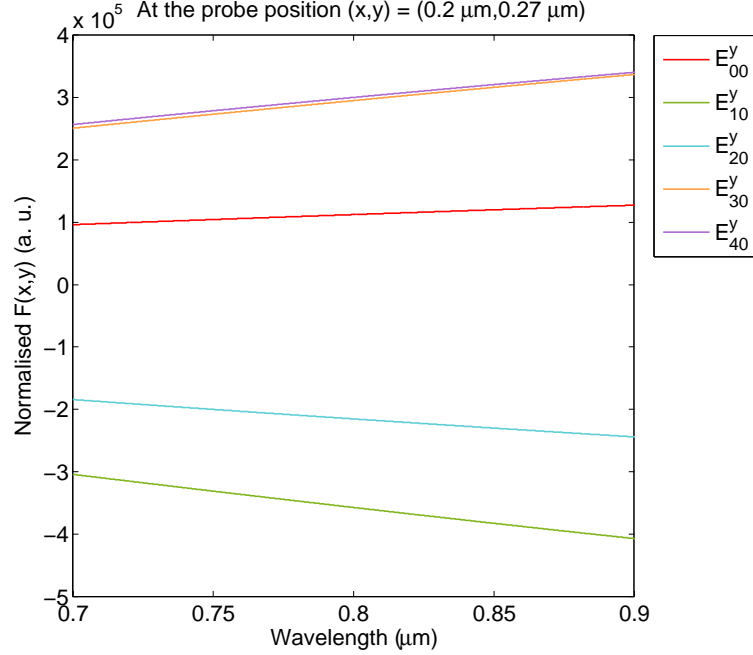
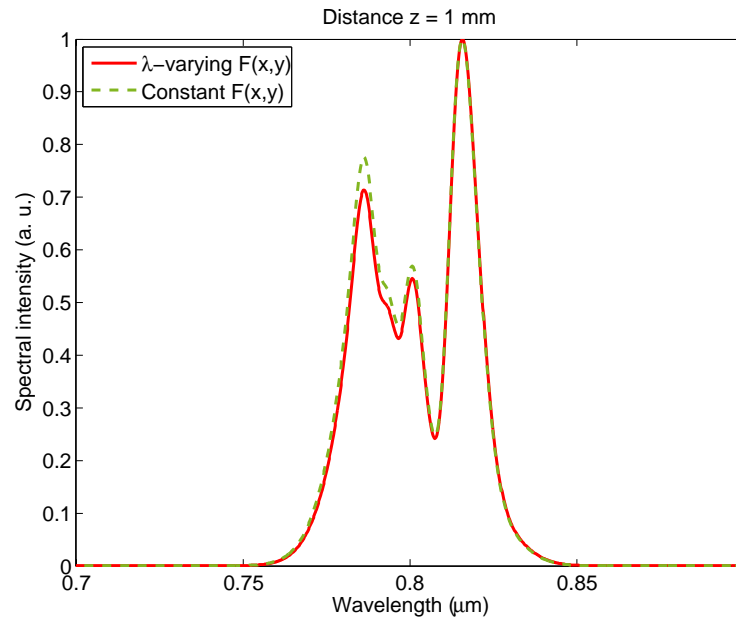
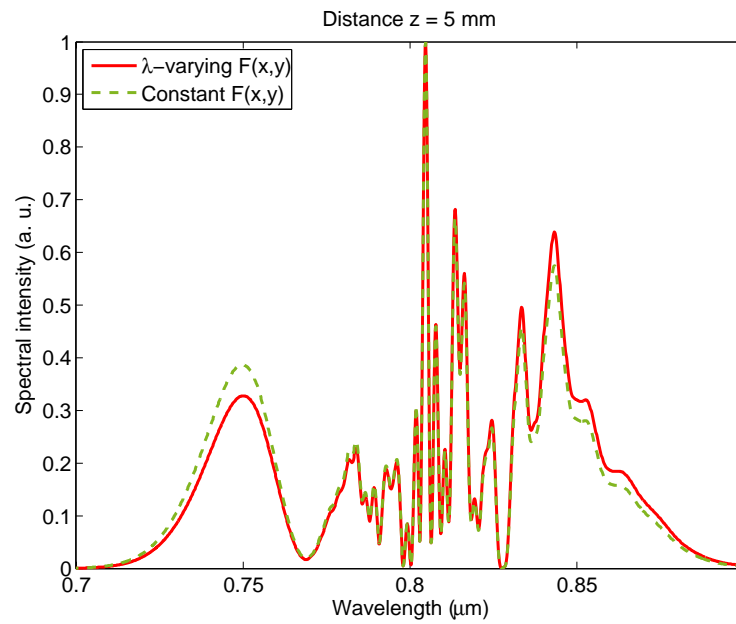


FIGURE 7.14: Normalised transverse field $U(x, y)$ over the wavelength 700 - 900 nm for E_{00}^y (red), E_{10}^y (green), E_{20}^y (blue), E_{30}^y (orange) and E_{40}^y (violet) at the position of the SNOM probe $(x, y) = (0.2 \mu\text{m}, 0.27 \mu\text{m})$

The normalised transverse field $U(x, y)$ at the position of the SNOM probe, $(x, y) = (0.2 \mu\text{m}, 0.27 \mu\text{m})$ at various wavelengths in the range of 700 nm - 900 nm, is displayed in Figure 7.14. The field varies in such a way that $|U(x, y)|$ rises with the increase of the wavelength. This is due to the deeper penetration of the field into the evanescent region for higher wavelengths. When this wavelength-dependent $U(x, y)$ is included in the calculation of the total field collected by the SNOM probe in Equation 5.18, Chapter 5, the effect causes some changes in the spectral features, as shown in Figure 7.15. The spectra are displayed with the solid red curve whereas the previous results in Figure 6.7, Chapter 6, which were calculated from the assumption of constant $U(x, y)$ over all the wavelength range of the spectra, are also included in the graphs and displayed with the dashed green curves. Figure 7.15(a) and Figure 7.15(b) are the simulated spectra from 5-mode pulse propagation. The spectra are observed by the probe while it is at the position of $z = 1 \text{ nm}$ and $z = 5 \text{ nm}$ respectively. The five modes consist of E_{00}^y , E_{10}^y , E_{20}^y , E_{30}^y , E_{40}^y and E_{50}^y mixed by the relative intensity ratio 0.7:0.05:0.05:0.05:0.05.



(a)



(b)

FIGURE 7.15: Comparison of 5-mode mixing simulated spectra between the case of constant (solid red curve) and wavelength-dependent (dashed green curve) transverse field. The simulated spectra are observed by a probe which is placed at the position of
(a) $z = 1$ mm and (b) $z = 5$ mm

As can be seen from the graphs, both solid red and dashed green curves contains similar spectral features. The effect of introducing the wavelength dependence of $U(x, y)$ in the field calculation causes only a minor alteration in spectral energy distribution, in such a way that the energy in the low wavelength region is transferred into the high wavelength region. By comparing the spectral features at 1 mm and 5 mm, it seems that the effect of wavelength dependence becomes more evident when the pulse propagates longer distances. The statement is confirmed by the calculation of root mean squared deviation (RMSD) between both types of spectra over all the wavelength range at different propagation distances. For two sets of data $I_1(\lambda)$ and $I_2(\lambda)$, with the data size N , RMSD can be determined from

$$\text{RMSD} = \sqrt{\frac{\sum_{p=1}^N (I_1(\lambda_p) - I_2(\lambda_p))^2}{N}} \quad (7.9)$$

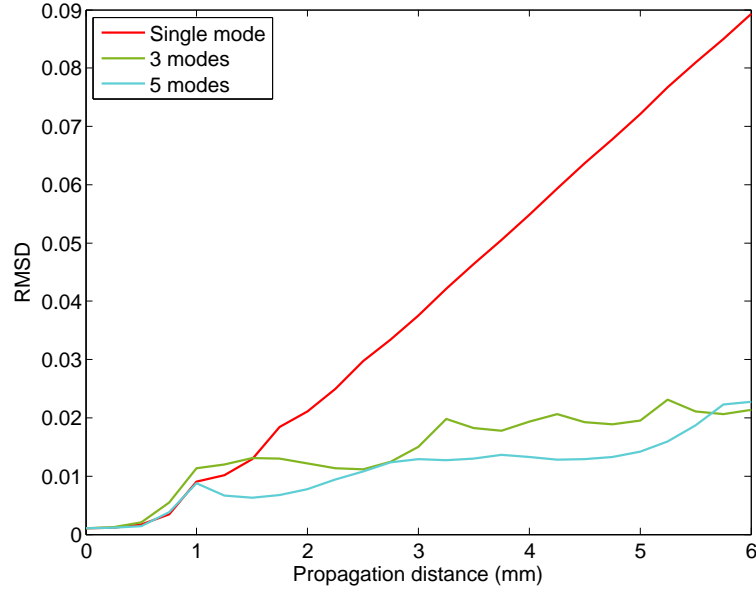


FIGURE 7.16: The root mean squared deviation (RMSD), calculated by Equation 7.9, between the spectra simulated when $U(x, y)$ is constant and wavelength-dependent. The deviations are from the spectra of single mode (red curve), three-mode (green curve) and five-mode pulses (blue curve). The detail of contributing modes and the relative intensity ratio is the same as already described in Figure 6.7, Chapter 6.

Figure 7.16 displays RMSD between the spectra from the calculation, with and without taking account of the wavelength dependency of the transverse field from single-mode (red curve), three-mode (green curve) and five-mode (blue curve) pulse propagation, with the details of the contributing mode and their relative intensity ratio the same as that in Figure 6.7, Chapter 6. For all the three curves, the deviation increases with the length of the propagation distance. The deviation in the single mode case is distinctly higher than in the multimode case and grows almost linearly along the distance of propagation. This is owing to the growth of the spectral bandwidth, which raises the degree of field

variation and therefore results in higher values of the deviation from the spectra with the constant transverse field. As already discussed in Section 6.4, Chapter 6, at the beginning of the propagation the spectral broadening of the multimode pulse is stronger than the single-mode and therefore results in a higher deviation of the three-mode and five-mode curves over the single-mode in Figure 7.16. The retardation of the spectral development caused by the walk-off of higher-ordered modes leads to the small increase of the deviation at the later stage compared with the steep rise of the deviation in the single-mode case.

7.5 Chapter conclusion

The temporal characteristic of the multimode pulse in the waveguide was derived by the application of the Fourier transform on the SNOM-acquired spectra developing along the length of the waveguide. The temporal profile reveals the modal separation owing to different group velocities of the propagation modes inside the waveguide. The group index difference of higher-ordered modes, relative to the fundamental mode, can be directly determined from the relation of temporal separation and propagation distance. Such a phenomenon was also visualised in the Fourier transform of the simulated spectra. However, the inaccuracy of measuring the mode index difference by this approach can be caused by the obscurity in defining the positions of separated higher-ordered mode pulses in time, especially when the pulses are superposed on one another.

In the previous chapter, the simulated spectra are based on the modeling whose dispersion effect is approximated by only up to the second-ordered term of the Taylor expansion and the adequacy of the approximation can be questionable. Therefore the strength of the third-ordered term in comparison with the second-ordered term and the nonlinear effect was investigated. With the order of the value of the dispersion and nonlinear parameters in this study, no significant deviation can be anticipated in the extension of the third-order term. Additionally the sufficiency of the second-order term is further confirmed by the unobservable alteration in the spectral results generated by the simulation including the third-order dispersion term.

Finally the aspect of the wavelength-dependency of the transverse field was discussed. The modeling was performed with the transverse field varying at various wavelengths and the results were compared with the previous simulation which took the transverse field as a constant value for all the range of wavelengths. Both kinds of results contain the same spectral feature and the only discrepancy is the redistribution of the intensity among the wavelengths.

7.6 References

- [1] M. L. M. Balistreri, J. P. Korterik, L. Knipers, and N. F. van Hulst. Local observations of phase singularities in optical fields in waveguide structures. *Physical Review Letters*, 85(2):294–297, 2000.
- [2] M. L. M. Balistreri, J. P. Korterik, L. Kiupers, and N. F. van Hulst. Phase mapping of optical fields in integrated optical waveguide structures. *Journal of Lightwave Technology*, 19(8):1169–1176, 2001.
- [3] J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, C. Netti, M. E. Zoorob, and J. J. Baumberg. Group velocity measurement using spectral interference in near-field scanning optical microscopy. *Applied Physics Letters*, 89(5):051101–1 – 051101–3, 2006.
- [4] T. Chaipiboonwong, P. Horak, J. D. Mills, and W. S. Brocklesby. Numerical study of nonlinear interactions in a multimode waveguide. *Optics Express*, 15(14):9040–9047, 2007.
- [5] G. P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, CA, USA, 3rd edition, 2001.

Chapter 8

Conclusions and suggested further work

8.1 Conclusions

Scanning near-field optical microscopy has been applied to study the dispersion and nonlinear phenomena in a multimode rectangular waveguide. The research consists of both experimental and numerical studies. The basic theory of SNOM was fully explained in Chapter 2. In contrast to conventional optical detectors, the SNOM technique can collect the localised evanescent field which contains the higher spatial frequency part of the scattered field. Therefore the technique provides subwavelength resolution which is not bound by diffraction. A SNOM system basically contains a subwavelength optical probe and a height control system which maintains the probe in the near-field range, which is around a few tens of nanometres from the surface of the sample. However, the particular details of the system configuration vary depending on the nature of the data required and the samples. The technical specification of the SNOM system in this research was given in Chapter 3. The probe was fabricated from a tapered optical fibre and its height was controlled by a shear-force technique, in which a probe is glued to a dithering tuning fork. The deviation of the amplitude of oscillation, caused by the varying shear force interaction while the probe is brought down toward the surface of the waveguide, was monitored by an electronic feedback system retaining the probe in the near-field region. The SNOM tip was positioned on a piezoelectric stage and translated over the surface of the waveguide to collect the evanescent field. The scattered light passed to a spectrometer connected at the other end of the fibre probe.

The theoretical background of light confinement in a rectangular waveguide was explained in Chapter 4. The mode determination was approximated by the effective index method (EIM) which is based on the assumption of a variable-separable field function. Therefore the two-variable problem is converted to the determination of propagation

modes on two planar waveguides with the index profiles orthogonal to each other. Both waveguides are coupled by the fact that the propagation constant of the first waveguide renders the effective index for the core of the second waveguide. The perturbation theory was also applied to enhance the accuracy of the EIM. The light pulse of each mode propagating inside the waveguide experiences both dispersion and nonlinear effects. The evolution of the pulse can be formulated as the nonlinear Schrödinger equation which is based on Maxwell's equations relating the interaction of the electric field and the polarisation of the waveguide medium.

The technical aspect of the numerical modeling was explained in Chapter 5. The algorithm utilised in the research to numerically model the nonlinear pulse propagation was the split-step Fourier method (SSFM). As its name suggests, the method divides the waveguide span into many small sections. In each section, the nonlinear and dispersion operators are assumed to interact independently. The former has its analytical solution in the time domain whereas the latter is in the frequency domain with the assistance of the Fourier transform. The accuracy of the algorithm is achieved owing to the non-commutability of the operators. The basic scheme, which treats both types of effects over the full step, is the first-order approximation whereas the second-order procedure is achieved by breaking up the operation of dispersion to half a step. Further improvement of the algorithm to the higher-order can be achieved by the application of a Runge-Kutta-like method. Alternatively the adaptive step scheme can be utilised to keep controlling localised error. However, all these come at the expense of the extra calculation cost. The dispersion operator is determined from the dispersion curves of the propagation constant and its higher derivatives. Unlike the single mode case, the intensity-dependent nonlinear parameter for multimode pulses are related to not only the integral of the transverse field distribution of the mode itself but also the overlapping integral of all the contributing modes.

This numerical modeling was applied to study the nonlinear pulse propagation in a Ta_2O_5 multimode waveguide fabricated on a silicon chip. The silicon-based compatibility can provide a future prospect for the waveguide to be applied to optical integrated circuits. The evolution of a pulse along the length of propagation was presented in Chapter 6. Both experimental and numerical results displayed similar qualitative features of the pulse spectra. The very fine complicated features are observable in both experimental and calculated spectra. Since such features are not present in the single-mode case, which retains only the self-phase modulation for the system's nonlinearity, the fine, unique features can be attributed to the interference of the contributing modes. This is also certified by the increase of the complexity with the number of modes. Moreover, the spectral growth reduces at a later distance of propagation even though the broadening process is rapid at the beginning. This cannot be fitted with the single-mode model in which the spectra broadens uniformly over the waveguide's full length. By analysing the simulated spectral broadening of contributing modes individually, it was shown that at

the beginning of the guide, the broadening grows quickly owing to additional cross-phase modulation. However, the broadening retarded later on when higher-order modes with different group velocities started to separate themselves from the main pulse.

The SNOM technique also provides the high-resolution observation of pulse development on a few hundred microns scale. As demonstrated by the simulated spectra, the variation of spectral features along the length of propagation over such a length scale is attributed to the phase velocity difference of contributing modes from different modal field symmetries. On the other hand, the variation across the waveguide over a similar range of scale is owing to the difference of phase and transverse field distribution of the contributing modes.

By applying the Fourier transform to the experimentally acquired spectral data, the corresponding time profile was available. In the time domain, the pulse exhibits mode separation according to the different group velocities at various distances along the length of the waveguide. As presented in Chapter 7, this can be utilised to derive the group index information of the waveguide modes. As demonstrated by the numerical result, this technique of acquiring the mode's group velocities is not affected by the nonlinearity of the waveguide. However, the accuracy of the method is bound by the precision of locating the time position of the separated higher-order mode peaks which can be obscured by the adjacent modes.

Other effects which were omitted in the numerical simulation were also discussed. For the scale length of the waveguide in this research, the higher-order dispersion was not expected to play a significant role in the pulse evolution. This claim was supported by the result from the simulation with the inclusion of the third-order dispersion. The final issue is the effect of wavelength dependence. For the sake of computational simplicity, the transverse field had been assumed to be constant over all the wavelength range of the observed spectra. The validity of the assumption was investigated by comparing the simulated spectra from the modeling by applying both the constant and wavelength-varying transverse field. The deviation, which grows with propagation distance, is however negligible, since the spectral features do not change because only the relative intensity of the peaks in the spectra is redistributed.

8.2 Suggested further work

More measurements of the characteristics of the waveguide are required in order to enable quantitative comparison between the measured and simulation data. One of the crucial parameters which is absent from the experimental work of this research is the relative ratio among propagation modes excited in the waveguide. This can be acquired by measuring the intensity distribution from the end face of the waveguide and comparing it with the calculated intensity profile of various propagation modes.

Moreover, controlling mode excitation in the waveguide would improve the investigation of the multimode effect on the spectral broadening. The task can be achieved by utilising a prism coupler instead of an objective lens to control the angle incidence of propagating light inside the waveguide.[1]. The further enhancing the accuracy of the study can be fulfilled by including the propagation loss of waveguide modes. The propagation loss can be measured by the nondestructive Fabry-Perot resonance technique. The structure of the waveguide can be considered as a Fabry-Perot resonator with two end faces as the resonator's reflecting planes and the loss of light propagating inside can be determined by the ratio between the maximum and minimum intensity outputs of the resonator [2]. Furthermore, the loss measurement for individual modes of multimode waveguides can be accomplished by implementing a prism as the output coupler to separate the waveguide modes [3].

So far in this work, the acquisition of the optical information by the SNOM technique has provided only the spectral intensity, whereas the spectral phase has been neglected. The localised spectral phase could possibly be gained by the simultaneous use of SNOM with other ultrashort pulse characterising methods such as spectral phase interferometric techniques [4–6] or frequency-resolved optical gating (FROG) techniques [7–9]. The complete information would allow a full understanding of the temporal and spectral pulse shaping induced by dispersion and optical nonlinearities and could assist in the optimisation of continuum generation sources. Furthermore, the numerical model can be developed to include other nonlinear effects such as Raman scattering or four-wave mixing and utilised to model other types of the waveguide.

Finally, the cooperation of both models for waveguide mode calculation and nonlinear pulse propagation may provide another waveguide design tool for optimising continuum generation or any desired nonlinear performance.

8.3 References

- [1] R. G. Hunsperger. *Integrated Optics: Theory and Technology*. Springer, Berlin, Germany, 5th edition, 2002.
- [2] G. T. Reed and A. P. Knights. *Silicon Photonics: An Introduction*. John Wiley & Sons, Chichester, England, 2004.
- [3] C.-T. Lee. Nondestructive measurement of separated propagation loss for multimode waveguides. *Applied Physics Letters*, 73(2):133–135, 1998.
- [4] V. Wong and I. A. Walmsley. Analysis of ultrashort pulse-shape measurement using linear interferometers. *Optics Letters*, 19(4):287–289, 1994.
- [5] I. A. Walmsley and V. Wong. Characterization of the electric field of ultrashort

- optical pulses. *Journal of the Optical Society of America B (Optical Physics)*, 13(11):2453–2463, 1996.
- [6] C. Iaconis and I. A. Walmsley. Spectral phase interferometry for direct electric-field reconstruction of ultrashort optical pulses. *Optics Letters*, 23(10):792–794, 1998.
- [7] D.J. Kane and R. Trebino. Characterization of arbitrary femtosecond pulses using frequency-resolved optical gating. *IEEE Journal of Quantum Electronics*, 29(2):571–579, 1993.
- [8] K. W. DeLong, R. Trebino, J. Hunter, and W. E. White. Frequency-resolved optical gating using second-harmonic generation. *Journal of the Optical Society of America B (Optical Physics)*, 11(11):2206–2215, 1994.
- [9] R. Trebino, K. W. DeLong, D. N. Fittinghoff, J. N. Sweetser, M. A. Krumbugel, B. A. Richman, and D. J. Kane. Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. *Review of Scientific Instruments*, 68(9):3277–3295, 1997.

Appendix A

Publications

T. Chaipiboonwong, P. Horak, J. D. Mills, and W. S. Brocklesby. Numerical study of nonlinear interactions in a multimode waveguide. *Optics Express*, 15(14):9040-9047, 2007.

T. Chaipiboonwong, P. Horak, J. D. Mills, and W. S. Brocklesby. Numerical simulation of continuum generation in a multimode nonlinear waveguide. In *European Conference on Lasers and Electro-Optics and the International Quantum Electronics Conference (CLEO/Europe - IQEC 2007)*, page 350, Munich, Germany, 2007.

J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, J. J. Baumberg, M. D. B. Charlton, M. E. Zoorob, and M. C. Netti. Spectral evolution of femtosecond pulses in nonlinear waveguides: measuring continuum generation and group velocity with NSOM. In *9th International conference on near-field optics, nanophotonics and related techniques*, Lausanne, Switzerland, 2006.

J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, C. Netti, M. E. Zoorob, and J. J. Baumberg. Group velocity measurement using spectral interference in near-field scanning optical microscopy. *Applied Physics Letters*, 89 (5):051101/1-051101/3, 2006.

J. D. Mills, T. Chaipiboonwong, W. S. Brocklesby, M. D. B. Charlton, M. E. Zoorob, C. Netti, and J. J. Baumberg. Observation of the developing optical continuum along a nonlinear waveguide. *Optics Letters*, 31(16):2459-2461, 2006.

J. D. Mills, T. Chaipiboonwong, M. D. B. Charlton, M. E. Zoorob, M. C. Netti, J. J. Baumberg, and W. S. Brocklesby. Waveguide group velocity determination by spectral interference measurements in near-field optical scanning microscopy. In *Conference on Lasers and Electro-Optics/Quantum Electronics and Lasers Science Conference (CLEO/QELS 2006)*, Long Beach, CA, USA, 2006.

T. Chaipiboonwong, J. D. Mills, M. D. B. Charlton, M. E. Zoorob, M. C. Netti, J. J.

Baumberg, and W. S. Brocklesby. Visualization of optical continuum evolution along a nonlinear waveguide. In *Conference on Lasers and Electro-Optics/Quantum Electronics and Lasers Science Conference (CLEO/QELS 2006)*, Long Beach, CA, USA, 2006.

Appendix B

Operation parameters of P-2000 Micropipette Puller

The details of five parameters required to be set for the operation of the micropipette puller are described below. All the information is from the instrument manual and private communication with engineers of the Sutter company [1–3]:

1. HEAT (Range 0-999)

The **HEAT** parameter specifies the power of the laser with the maximum value of 10W. The suggested value for the optical fibre is in the range of 280-350. The unit of the parameter is calibrated by two points: 190 units \rightarrow 1 W and 820 units \rightarrow 8 W. As a result, the relation between the unit and the laser power can be determined from

$$\text{Power} = 0.0111 \times \text{Unit} - 1.1111 \quad (\text{B.1})$$

2. FILAMENT (Range 0-15)

The **FILAMENT** parameter specifies the length of the region scanned by the laser beam. The lengths for the first six units are shown in Table B.1. For the optical fibre, **FILAMENT** should always be set at 0.

FILAMENT	Scan length (mm)
0	1
1	1.5
2	1.9
3	4.5
4	6.5
5	8

TABLE B.1: Corresponding scan length for the unit of FILAMENT [1]

3. VELOCITY (Range 0-255)

The **VELOCITY** parameter sets the velocity of the carriage system measured before the hard pull is applied. The velocity monitoring system consists of a velocity transducer circuit. One unit of the parameter is comparable to a change in the transducer's output voltage by one or more millivolts, depending on the transducer being used. The recommended value for the optical fibre is in the range of 18-20. Typically the circuit is calibrated in such a way that 120 mV corresponds to a velocity of 1 inch/s.

4. PULL (Range 0-255)

The **PULL** parameter specifies the current passing through the pull solenoid and therefore is proportional to the force of the hard pull. The maximum unit of 255 is set to produce the solenoid current of 2.5 A. The suggested value of 150 is for the optical fibre which gives the solenoid current of around 1.47 A. The pull force is proportional to the solenoid current and the relative position of the solenoid and the plunger. The latter factor depends on the cable adjustment and therefore the resultant force varies from one puller to another.

5. DELAY (Range 0-255)

The **DELAY** parameter defines the time delay (Δt) for the start of the hard pull after the laser turn-off. One unit is equivalent to 1 ms and the unit of 128 is set for $\Delta t = 0$, that is, the hard pull starts just as the heat turns off. Therefore, Δt is determined by

$$\Delta t = (\text{DELAY} - 128) \times 10^{-3} \quad (\text{B.2})$$

If $\text{DELAY} > 128$ then the hard pull starts at Δt before the laser turns off and vice versa for $\text{DELAY} < 128$. For example, the unit of 126, which is the recommended

value for the optical fibre, will set the system activating the hard pull 2 ms before the laser heating is deactivated whereas for the unit of 130, the hard pull will start after the laser turns off by 2 ms.

B.1 References

- [1] Sutter Instrument Company. Operation manual: P-2000 micropipette puller.
- [2] Stelea, S. Re: P-2000 micropipette puller. Private communication, 25 July 2008.
- [3] Belgum, J. P2000. Private communication, 11 September 2008.