SATURABLE ABSORPTION IN SINGLE-WALLED CARBON NANOTUBES AT 1550 nm

by

ROBERT TAK FAI MA

B.A.Sc., The University of British Columbia, 2003

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

April 2006

ABSTRACT

Single-walled carbon nanotubes (SWNTs) provide an alternative to multi-quantum well (MQW) based saturable absorbers at optical communication wavelengths. This work is aimed at measuring the nonlinear absorption coefficient of SWNTs at 1550 nm and to compare the performance of SWNT based saturable absorbers with MQW based saturable absorbers.

In this work, saturable absorption and optical bleaching in a SWNT sample at 1550 nm were observed by the Power-scan and the two-color pump-probe measurement methods. A measurement system capable of performing Z-scan, P-scan, and pump-probe measurements was developed. The filtration method was used to fabricate a thin film of SWNTs grown by the HiPco process onto a sapphire substrate. The experimental results obtained from the P-scan measurements were analyzed to estimate the nonlinear absorption coefficient of the SWNT sample. The estimated values are large compared to those of other nonlinear optical materials. Optical bleaching was observed in the pump-probe measurement. A figure of merit to compare the performance of saturable absorbers based on their modulation depth, saturation intensity, and recovery time is proposed. Using it, the figures of merit of the SWNT sample used in this work, those of several SWNT based saturable absorbers reported in the literature, and those of several MQW based saturable absorbers, also reported in the literature, were calculated. The calculations showed that the performance of SWNT based saturable absorbers is comparable to MQW based saturable absorbers. The optical properties of SWNTs, combined with the low cost of fabrication, demonstrate the potential of SWNT based devices to find a large number of optical applications.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

CHAPTER 1

INTRODUCTION

Saturable absorbers are used in optical applications such as mode-locking [1, 2] and passive optical regeneration [3-5]. The fast recovery times of saturable absorbers and their ability to attenuate noise to increase signal-to-noise ratio promote higher bit-rates and extended error-free transmission distances of periodically amplified optical transmission lines [3-5].

A saturable absorber is an intensity dependent nonlinear optical device. At low intensity, the saturable absorber will absorb incoming light; whereas at high intensity, the ground state of the absorber is depleted and the absorption bleached out. Currently, the most commonly used saturable absorbers are semiconductor-based multi-quantum-well (MQW) devices [3-5]. Fabrications of such MQW based devices are complicated and costly, and these devices generally possess slow recovery times.

Carbon nanotubes were first discovered by Iijima in 1991 [6]. They are unique nanosturctures with interesting electrical, mechanical, chemical and optical properties [7]. In recent years, two groups [8, 9] have reported saturable absorption with ultrafast sub-picosecond recovery time in single-walled carbon nanotubes (SWNTs) at optical communication wavelengths. Mode-locking using SWNT based saturable absorbers has also been demonstrated in fiber ring lasers [10, 11]. The optical properties of SWNTs depend on their tube geometry, which implies that devices based on them can be tuned by selecting the appropriate tube geometry. Moreover, the fabrication cost and complexity of SWNT based devices are potentially much lower than those of MQW based devices.

All of the above indicate that SWNTs provide an alternative to MQW based devices at optical communication wavelengths.

So far, not much work has yet been put into quantitatively measuring the nonlinear absorption coefficient of SWNTs at optical communication wavelengths nor into comparing the performance of SWNT based saturable absorbers to those of MQW based saturable absorbers; therefore, it is the aim of this work to measure the nonlinear absorption coefficient of SWNTs at 1550 nm and to compare the performance of SWNT based saturable absorbers to those of MQW based saturable absorbers.

Chapter two provides an introduction to SWNTs. The structural parameters of SWNTs are determined from those of graphene. From the calculations of the electronic dispersion relations of SWNTs, it will be shown that the optical properties of SWNT based devices can be tuned by selecting appropriate tube geometries.

In Chapter three, the experimental techniques used in this work are presented. Overviews of the Power-scan and the pump-probe measurement methods are provided. The design of the measurement system to perform the described measurements is shown. The fabrication of the SWNT sample is also described and the quality of the sample evaluated.

In Chapter four, the obtained experimental results are analyzed to estimate the nonlinear absorption coefficient of SWNTs. A figure of merit to evaluate the performance of saturable absorbers based on their modulation depth, saturation intensity, and recovery time is proposed. The calculated figure of merits for the fabricated SWNT samples and several SWNT based saturable absorbers reported in the literature [8, 9, 12]

are compared with those of MQW based saturable absorbers, also reported in the literature [13-15], to show the potential of SWNT based devices.

Finally, in chapter five, the obtained results are summarized. Recommendations for future work are also provided.

CHAPTER 2

SINGLE-WALLED CARBON NANOTUBE

2.1 Introduction

The properties of SWNTs are determined by their chirality, which is unique for different tube geometries. Due to the similarity in crystal structure of graphene and SWNTs, graphene provides the platform for obtaining the basic properties of SWNTs. In this chapter, the basic properties of SWNTs are introduced. The electronic structure of SWNTs is obtained by applying the zone-folding approximation [7, 16] to the energy dispersion relations of 2D graphite. The optical properties of SWNTs are also described. Much of the discussion of SWNTs given here follows references [16] and [17].

2.2 Structures of Single-Walled Carbon Nanotubes

A SWNT surface lattice has the same crystal structure as graphene: a honeycomb lattice of carbon hexagons; therefore, a SWNT can be regarded as a single graphene sheet wrapped into a seamless cylinder. Unwrapping the cylinder reveals the unit cell of SWNTs defined by two vectors perpendicular to each other: the chiral vector C_h and the translational vector T. The ends of C_h coincide when the graphene is rolled up. The unit cell of a particular SWNT is illustrated in figure 2.1a. C_h defines the direction in which the graphene is rolled into a tube and is unique for different tubes. C_h is specified by a pair of integers (n,m), which relates C_h to the real space unit vectors a_1 and a_2 as

$$C_h = na_1 + na_2 \qquad (2.1)$$

where $|a_1| = |a_2| = 1.44$ Å $\times \sqrt{3} = 2.49$ Å. The two real space vectors are shown in figure 2.1a. SWNTs are denoted by the integers (n,m).

SWNTs can be classified as either being achiral or chiral based on their symmetry. An achiral SWNT has an identical mirror image about C_h while a chiral SWNT does not. There are two forms of achiral SWNTs: zigzag and armchair. For zigzag SWNTs, m = 0 and C_h goes through the zigzag edges of the hexagons. For armchair SWNTs, n = m and C_h goes through the diagonal lines of the hexagons. See figures 2.1b and c for illustrations of the directions of C_h of zigzag and armchair tubes. For chiral SWNTs, C_h makes a chiral angle θ_c to a_1. θ_c defines the direction of C_h and can be calculated by taking the inner product of C_h and a_1, yielding the expression

$$\cos\theta_c = \frac{a_1 \cdot C_h}{|a_1| \cdot |C_h|} = \frac{2n + m}{2\sqrt{n^2 + m^2 + nm}} \tag{2.2}$$

θ_c also denotes the tilt angle of the hexagons with respect to the direction of the tube axis. Equation 2.2 shows that for zigzag and armchair SWNTs, $\theta_c = 0°$ and $30°$, respectively. In the case of chiral SWNTs, θ_c is restricted to $0° < \theta_c < 30°$ (or n > m > 0). This is due to the symmetry of the crystal structure. For tubes with $0° < \theta_c < 30°$, there exist an equivalent tube with $30° < \theta_c < 60°$. In general, because of the hexagonal symmetry of the honeycomb lattice, for any given C_h there exists an equivalent one with $\theta_c \leq 60°$; therefore, θ_c is restricted to $0° \leq \theta_c \leq 30°$ (or n $\geq$ m $\geq$ 0). See figure 2.1d for an illustration of the symmetry operation of θ_c. In figure 2.1d, the black and open arrows represent the directions of C_h of zigzag and armchair SWNTs, respectively, while the grey arrows represent the direction of C_h of a particular chiral SWNT. Due to the

hexagonal symmetry of the honeycomb lattice, all the grey arrows are equivalent C_h. The same applies to the zigzag and armchair C_h.

(a)

(b)

(c)

(d)
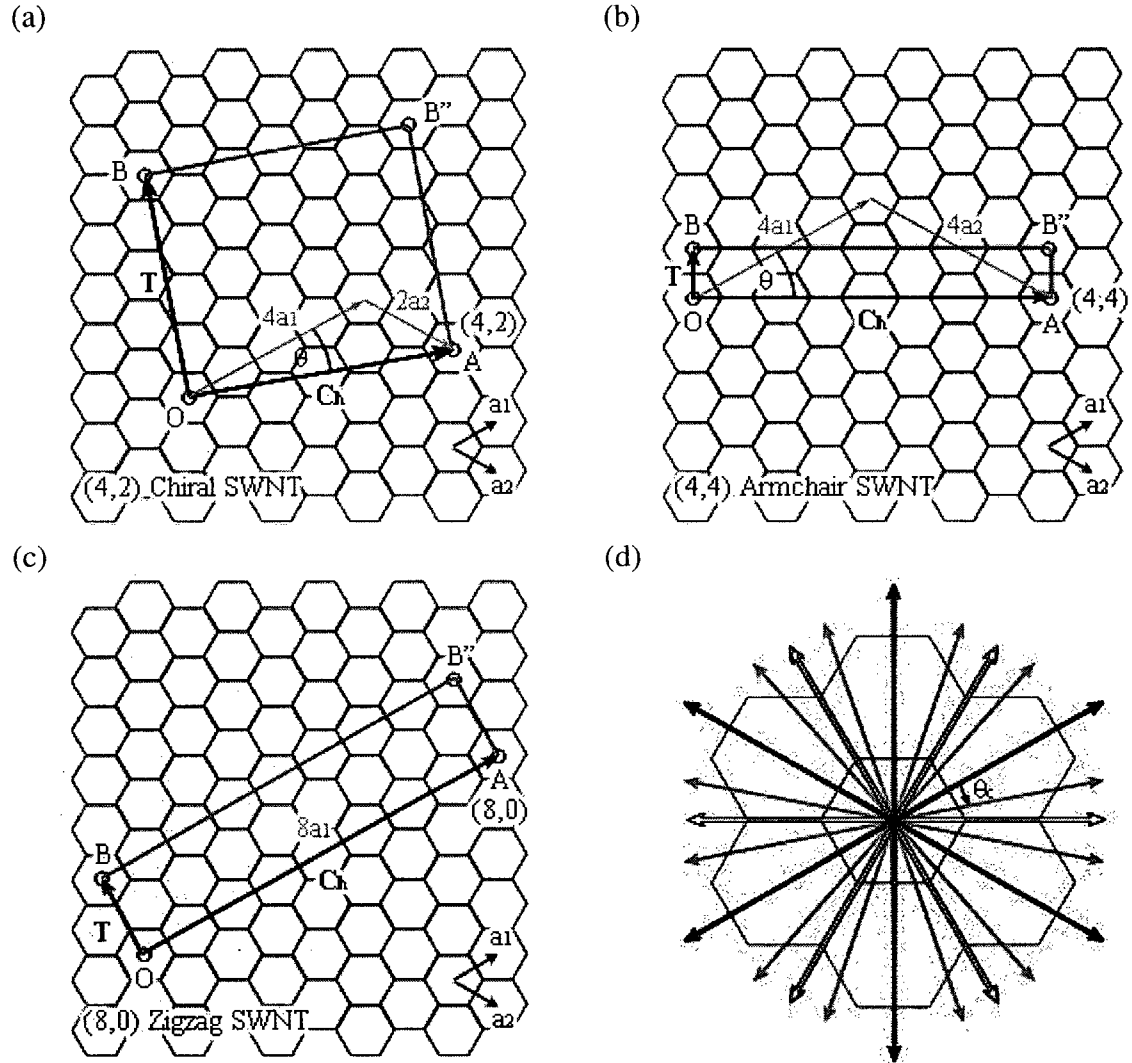
Figure 2.1: The unit cells of (a) a (4,2) chiral, (b) a (4,4) armchair, and (c) a zigzag SWNT. The unit cell is defined by OAB"B. When a SWNT is rolled up, O and A coincide. The vectors OA and OB define the chiral vector C_h and translational vector T, respectively. a_1 and a_2 are the real space unit vectors. θ_c is the chiral angle. Figure 2.1d shows the symmetry operation of θ_c. The black and open arrows represent the directions of C_h of zigzag and armchair SWNTs, respectively, while the grey arrows represent the direction of C_h of a particular chiral SWNT. Due to the hexagonal symmetry of the honeycomb lattice, all the grey arrows are equivalent C_h. The same applies to the zigzag and armchair C_h.

6

The diameter of a SWNT, d_t, is given by L/π, where L is the circumference of the SWNT which, in turn, is given by

$$L = \left| C_h \right| = a\sqrt{n^2 + m^2 + mn} \qquad (2.3)$$

The translational vector T is defined as the length of the unit cell. It is parallel to the tube axis and is normal to C_h. T is related to the real space unit vectors a_1 and a_2 as

$$T = t_1 a_1 + t_2 a_2 \qquad (2.4)$$

where t_1 and t_2 are integers given by

$$t_1 = \frac{2m + n}{d_R}, \quad t_2 = -\frac{2n + m}{d_R} \qquad (2.5)$$

where d_R is the greatest common divisor of $(2m + n)$ and $(2n + m)$. Similar to C_h the length of T is given by

$$\left| T \right| = a\sqrt{t_1^2 + t_2^2 + t_1 t_2} = \frac{\sqrt{3}L}{d_R} \qquad (2.6)$$

Dividing the area of the unit cell $|C_h \times T|$ by the area of the hexagon $|a_1 \times a_2|$ gives the number of hexagons per unit cell, N, as function of n and m,

$$N = \frac{\left| C_h \times T \right|}{\left| a_1 \times a_2 \right|} = \frac{2\left(n^2 + m^2 + nm \right)}{d_R} = \frac{2L^2}{a^2 d_R} \qquad (2.7)$$

Since there are two carbon atoms per hexagon, the total number of carbon atoms in a unit cell is 2N.

From the above equations it can be seen that the diameter, chirality as well as other parameters of SWNTs are determined by C_h and its indices (n,m).

2.3 Electronic Structure of Single-Walled Carbon Nanotubes

The density of states of a SWNT is given by the energy dispersion, which, in turn, is obtained by zone-folding the energy dispersion of 2D graphite [7]. Here only the valence π and the conduction π^* energy bands are considered. Shown in figure 2.2a and 2.2b are the unit cell and the Brillouin zone of 2D graphite shown as the rhombus and the shaded hexagon, respectively. A and B are carbon atoms in the unit cell, b_1 and b_2 are the reciprocal lattice vectors of 2D graphite, and Γ, K, and M mark the high symmetry points in the hexagonal Brillouin zone.

(a)   (b)
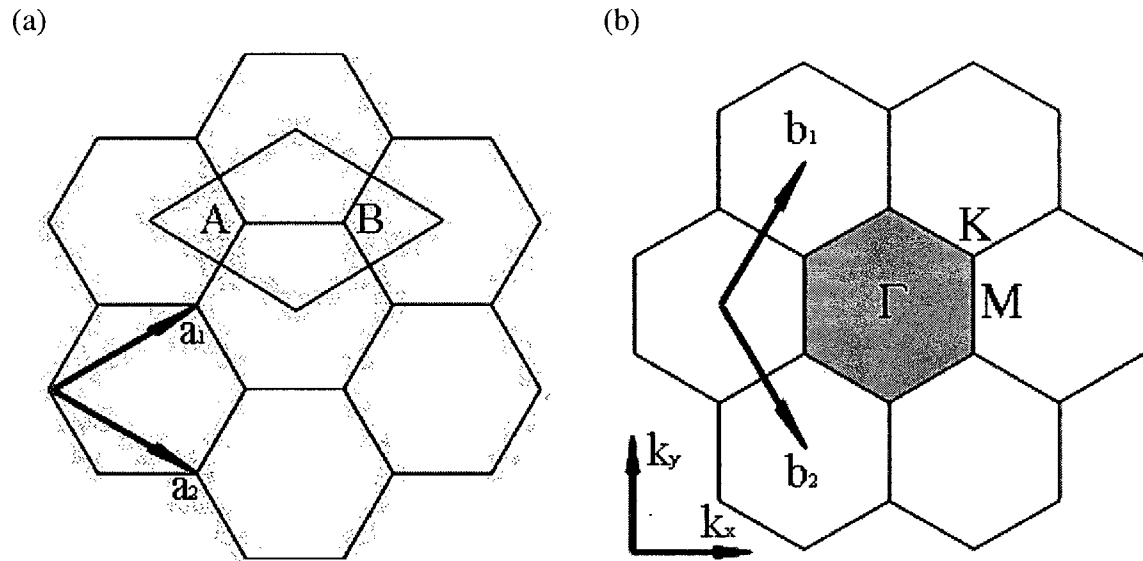
Figure 2.2: (a) The unit cell and (b) the Brillouin zone of 2D graphite shown as a rhombus and shaded hexagon respectively. a_1 and a_2 are the real space unit vectors while b_1 and b_2 are the reciprocal lattice vectors. The high symmetry points in the Brillouin zone are marked as Γ, K, and M.

By applying the tight-binding method [16], the energy dispersion of π energy bands of 2D graphite is given by

$$E_{g2D}(k) = \frac{\varepsilon_{2p} \pm t \times w(k)}{1 \pm s \times w(k)} \qquad (2.8)$$

where ε_{2p} is the orbital energy of the 2p level, t is the carbon-carbon interaction energy (-2.75 eV [18]), s is the overlap integral between the nearest atoms and $-\pi/a < k < \pi/a$. w(k) is given by

$$w(k) = \sqrt{1 + 4\cos\frac{\sqrt{3}k_x a}{2}\cos\frac{k_y a}{2} + 4\cos^2\frac{k_y a}{2}} \qquad (2.9)$$

The directions of k_x and k_y are shown in figure 2.2b.

The energy dispersion relation of 2D graphite calculated by equation 2.8 is shown in figure 2.3. ε_{2p}, t, and s are set to 0, -3.033 eV, and 0.129 to fit both the first principles calculation of the 2D graphite energy bands and experimental data [7]. In figure 2.3, the top half of the energy dispersion curve corresponds to the conduction π^* energy band while the bottom half corresponds to the valence π energy band. The conduction and valence band come into contact at the six corners (high symmetry points K) in the Brillouin zone, implying that 2D graphite is a zero-gap semiconductor. When s = 0, the valence and the conduction band become symmetric about $E = \varepsilon_{2p}$. This is given by

$$E_{g2D}(k_x, k_y) = \pm t\sqrt{1 + 4\cos\frac{\sqrt{3}k_x a}{2}\cos\frac{k_y a}{2} + 4\cos^2\frac{k_y a}{2}} \qquad (2.9)$$

These energy dispersion relations, where s = 0, are commonly used to approximate the electronic structure of a graphene layer, which can be used to approximate the electronic dispersion relations of SWNTs.
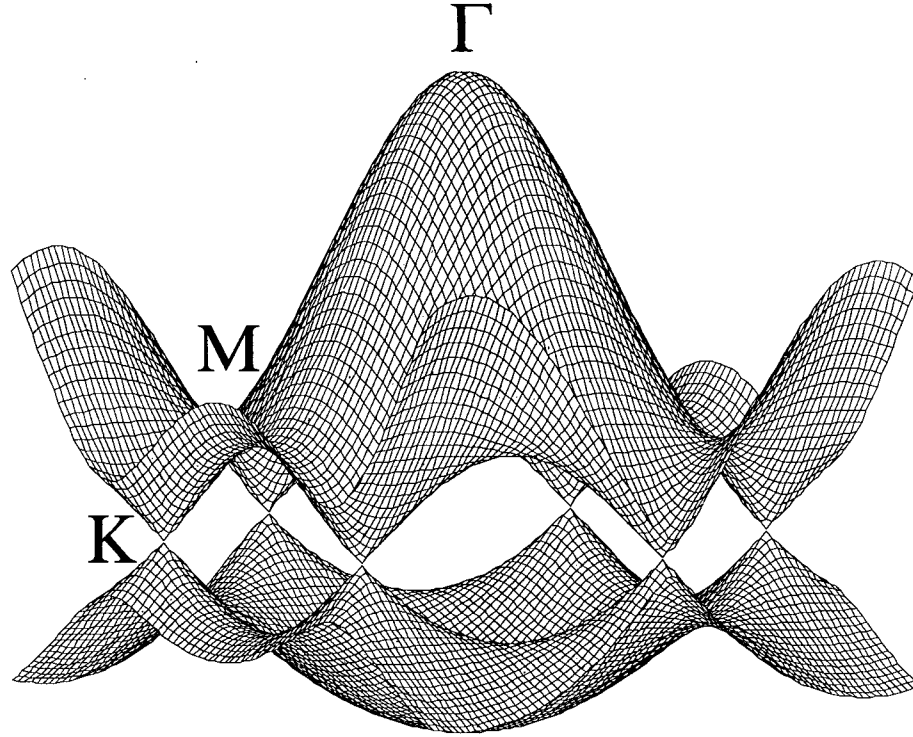
Figure 2.3: Energy dispersion relations of 2D graphite. Γ, K, and M mark the high symmetry points of the energy dispersion relations. The conduction π^* energy band and the valence π energy band touch at the high symmetry K points, indicating that 2D graphite is a zero-gap semiconductor.

As described in section 2.2, a SWNT is a single graphene sheet wrapped into a seamless cylinder. Following the same approach the energy dispersion of a SWNT can be obtained by folding that of graphene. This approach is referred as the zone-folding approximation [7, 16, 17]. The 1D energy dispersion of a SWNT is given by

$$E_\mu(k) = E_{g2D}\left(k\frac{K_2}{|K_2|} + \mu K_1\right) = E_{g2D}(k_x, k_y) \qquad (2.10)$$

with $-\pi/|T| < k < \pi/|T|$, and $\mu = 0, 1, ..., N-1$. K_1 and K_2 are the reciprocal lattice vectors in the circumferential direction and along the tube axis, respectively. The expression for the reciprocal lattice vectors, K_1 and K_2 are given by

10

$$K_1 = \frac{1}{N}(-t_2 b_1 + t_1 b_2), \; K_2 = \frac{1}{N}(-m b_1 + n b_2) \qquad (2.11)$$

where b_1 and b_2 are the reciprocal lattice vectors of 2D graphite given by

$$b_1 = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{a}\right), \; b_1 = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{a}\right) \qquad (2.12)$$

Because of the translational symmetry of T, K_2 is continuous and has a length equal to $2\pi/|T|$. However, K_1, the vector around the SWNT circumference, is quantized due to the periodic boundary condition [17]

$$\mu\lambda = |C_h| = \pi d_t, \; K_1 = \frac{2\pi}{\lambda} = \frac{2\pi}{|C_h|}\mu = \frac{2}{d_t}\mu \qquad (2.13)$$

This boundary condition states that a wave function must have a phase shift of an integer multiple of 2π around the circumference. All other wavelengths will vanish by interference. This gives rise to N discrete values in the circumference direction.

Having described K_1 and K_2, the Brillouin zone of a SWNT then can be viewed as a series of parallel line segments, given by $kK_2/|K_2| + \mu K_1$, that slice the hexagonal Brillouin zone of 2D graphite. The distance between adjacent lines is $2/d_t$. The resulting N pairs of energy dispersion curves given by equation 2.10 correspond to the cross-sections of the energy dispersion surface of 2D graphite (see figure 2.3). Figure 2.4 shows the Brillouin zone of a (4,2) SWNT.
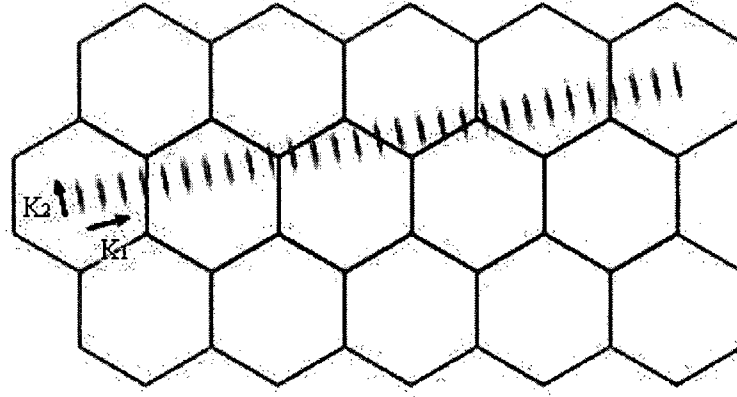
Figure 2.4: The Brillouin zone of a (4,2) SWNT described by the $N = 28$ parallel line segments given by $kK_2/|K_2| + \mu K_1$. K_1 and K_2 are the reciprocal lattice vectors in the circumferential direction and along the nanotube axis respectively. K_1 is quantized by the boundary condition, giving rise to $N = 28$ discrete values in the circumferential direction.

Before applying equation 2.10 (2.9) to calculate the energy dispersion of a SWNT, a conversion from K_1 and K_2 to k_x and k_y (see figure 2.2) is required. The conversion is given by the rotation matrix

$$\begin{bmatrix} kx \\ ky \end{bmatrix} = \begin{bmatrix} \cos\theta_{ra} & -\sin\theta_{ra} \\ \sin\theta_{ra} & \cos\theta_{ra} \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \tag{2.14}$$

where θ_{ra} is the rotation angle given by

$$\theta_{ra} = \left| \tan^{-1}\left(-\frac{n-m}{\sqrt{3}(n+m)} \right) \right| \tag{2.15}$$

An interesting property of SWNTs is that they can either be metallic or semiconducting, depending on their chirality. If one of the lines of $kK_2/|K_2| + \mu K_1$ intersects one of the high symmetry points K in the energy dispersion of 2D graphite, then the SWNT has a zero band gap and is metallic. If the high symmetry points K lie between the lines of $kK_2/|K_2| + \mu K_1$, then the SWNT has a finite band gap and is a semiconductor (see figure 2.4). A general rule for determining the electronic property of

12

SWNTs is that if (2n + m) or equivalently (n − m) is a multiple of 3, then the SWNT is metallic; otherwise the SWNT is semiconducting. The energy dispersion of a semiconducting zigzag (5,0) and a metallic armchair (5,5) SWNT are shown in figure 2.5. Notice that for armchair SWNTs, (2n + n) is always a multiple of 3 thus all armchair SWNTs are metallic. For zigzag SWNTs they are only metallic if n is a multiple of 3.

The density of states (DOS) in units of states/C-atom/eV is given by [7]

$$D(E) = \frac{T}{2\pi N} \sum_{\pm} \sum_{\mu=1}^{N} \int \frac{1}{\left| dE_{\mu}^{\pm}(k)/dk \right|} \delta\left(E_{\mu}^{\pm}(k) - E\right) dE \qquad (2.16)$$

Equation 2.16 shows that when the energy dispersion becomes flat (at band edges), the DOS becomes large and sharp singularities occur. These sharp singularities at band edges are called Van Hove singularities. The DOS of a SWNT can be calculated using the 1D energy dispersion of SWNT given by equation 2.10. The DOS of a semiconducting zigzag (5,0) and a metallic armchair (5,5) SWNT are shown in figure 2.6. Notice that for semiconducting SWNTs, the DOS at the Fermi level is zero while it is non-zero for the metallic SWNTs.
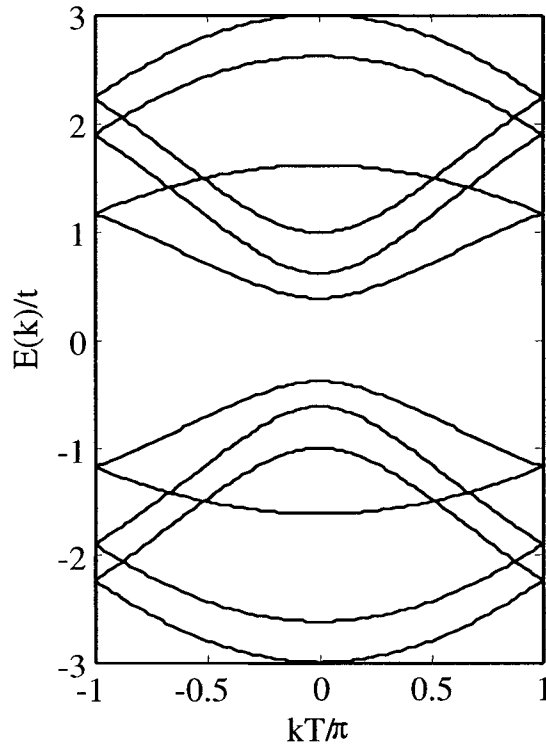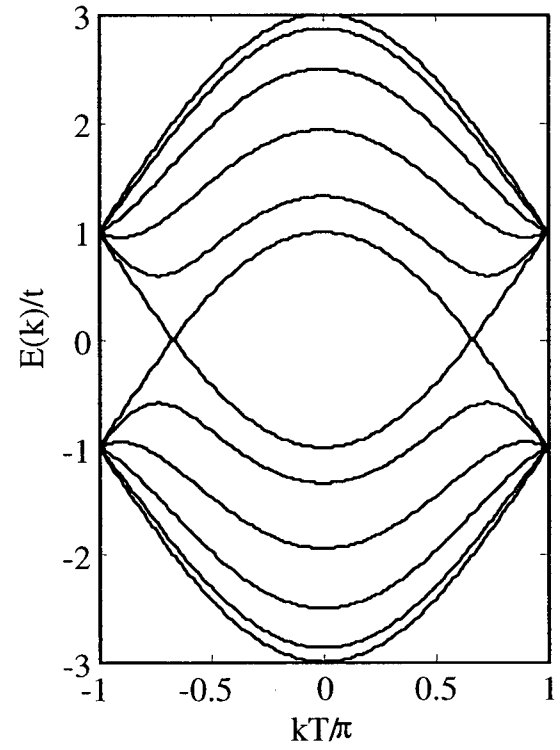
Figure 2.5: The 1D energy dispersion of (a) a semiconducting zigzag (5,0) SWNT and (b) a metallic armchair (5,5) SWNT. |t| = 2.75 eV.

(a)

(b)
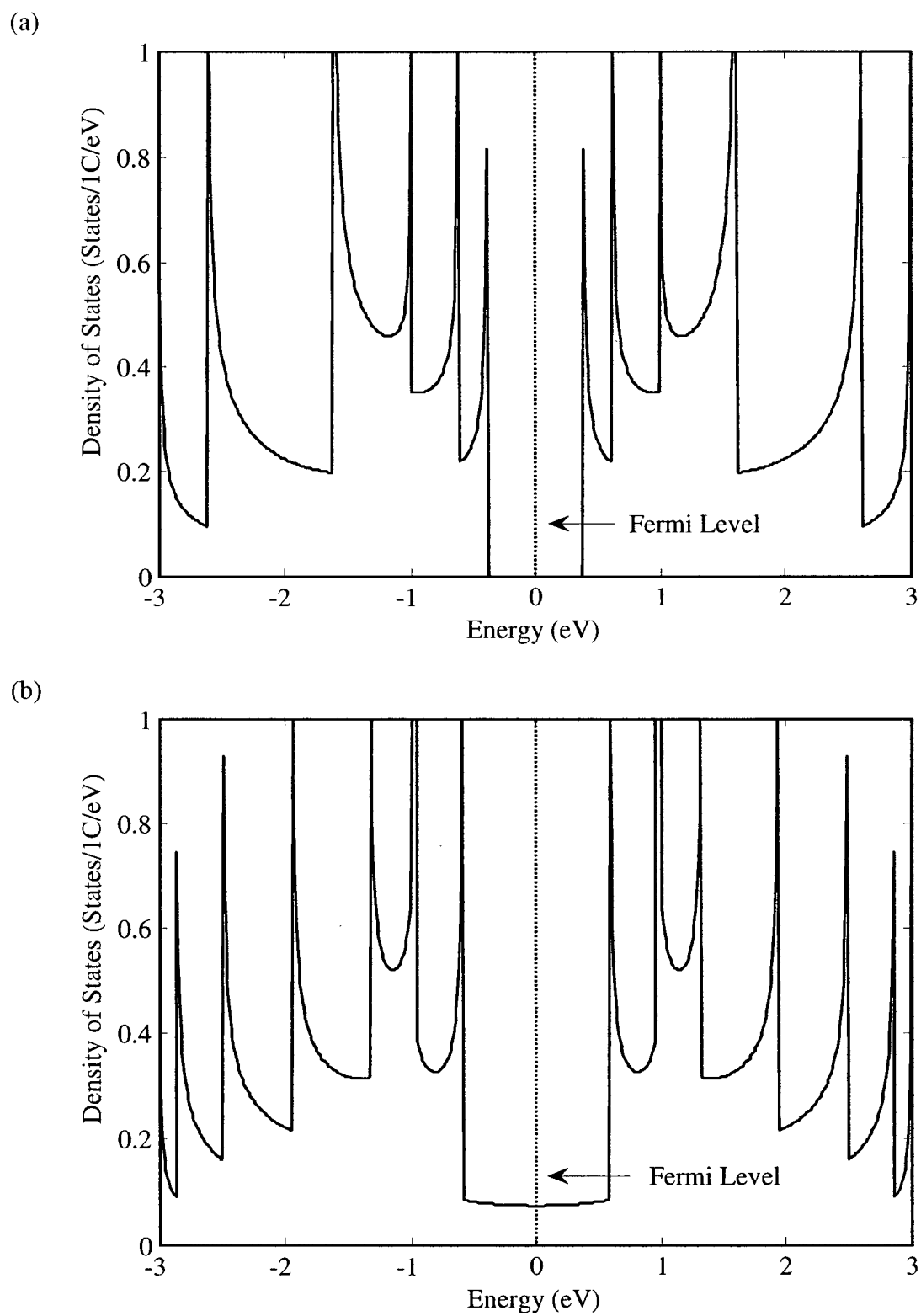
Figure 2.6: The DOS of (a) a semiconducting zigzag (5,0) SWNT and (b) a metallic armchair (5,5) SWNT. |t| = 2.75 eV. The Fermi level is at eV = 0.

15

Assuming light polarized parallel to the tube axis, the selection rules, which will be described in the next section, permit transitions from $-E_\mu$ in the valence π energy band to the E_μ conduction π^* energy band. The energy difference between the two energy bands is the band gap energy. The corresponding band gap energies of $E_{\mu\mu}$ (subband) transitions can be plotted as a function of tube diameters. This plot is shown in figure 2.7 and is called the Kataura plot [18]. The solid circles represent the band gap energies of metallic SWNTs and the open circles represent those of semiconducting SWNTs. In the Kataura plot, the first and second band gaps come from semiconducting SWNTs while the third band gap comes from metallic SWNTs with the same diameter. An important observation is that the band gap energies are dependent upon the tube diameters.

The Kataura plot is very useful for determining the optimal tube diameter for a particular optical absorption peak location. For example, for an absorption peak at 1550 nm (0.8 eV) corresponding to the first band gap, the optimal tube diameter would be ~1 nm. This also implies that the optical properties of SWNTs can be tuned by adjusting the tube diameter, positioning the band gap (absorption peak) at the desired energy level (wavelength).
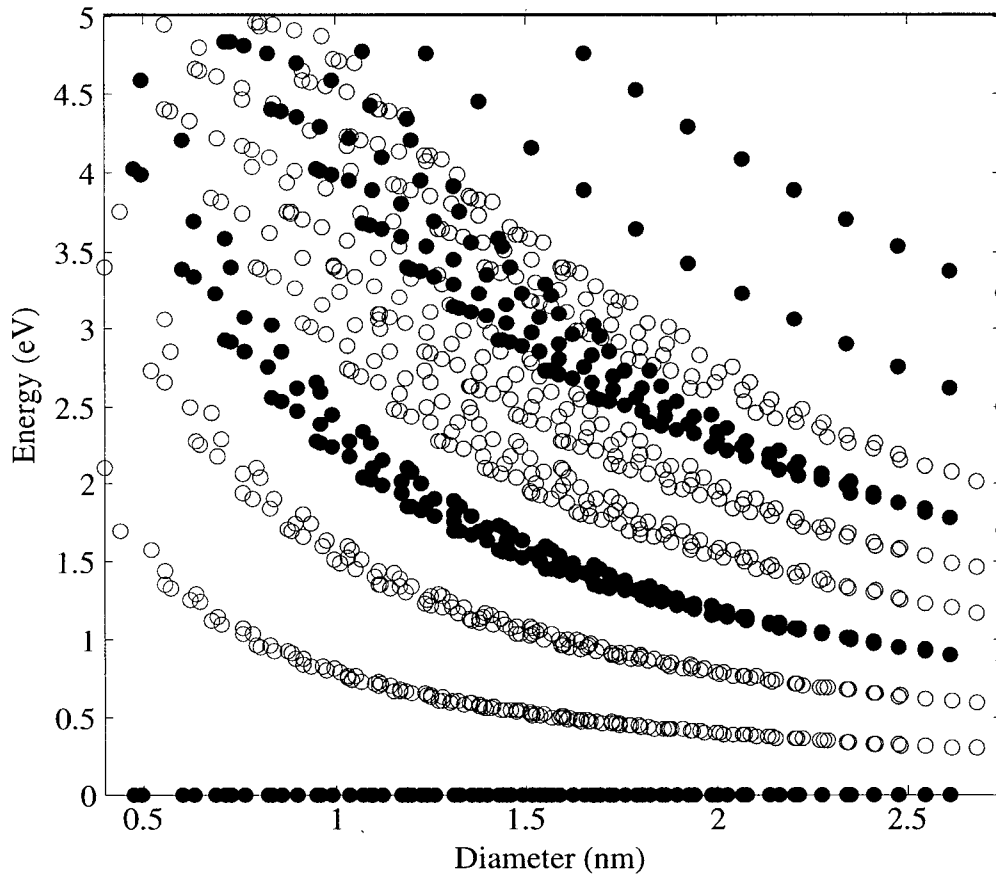
Figure 2.7: The Kataura plot showing the band gap energies for the five subbands of SWNTs with indices (n,m) where $5 \leq n \leq 20$ and $m \leq 20$. $|t| = 2.75$ eV.

2.4 Optical properties of Single-Walled Carbon Nanotubes

Optical transitions in SWNTs follow a set of selection rules. These selection rules depend on the relative polarization of the electric-field to the tube axis. Here only the selection rules for optical transitions induced by z-polarized light (electric-field parallel to the tube axis or E ‖ z) are considered since optical transitions induced by an electric-field perpendicular to the tube axis are suppressed by depolarization. This depolarization effect will be discussed later.

Figure 2.8 shows a few of the energy dispersion curves of a zigzag (5,0) SWNT. The arrows indicate the allowed optical transitions from valence to conduction band for z-polarized light (E ∥ z). For the conservation of total angular momentum when absorbing a photon, the band number μ must remain constant; therefore, only $E_{\mu\mu}$ transitions are allowed.
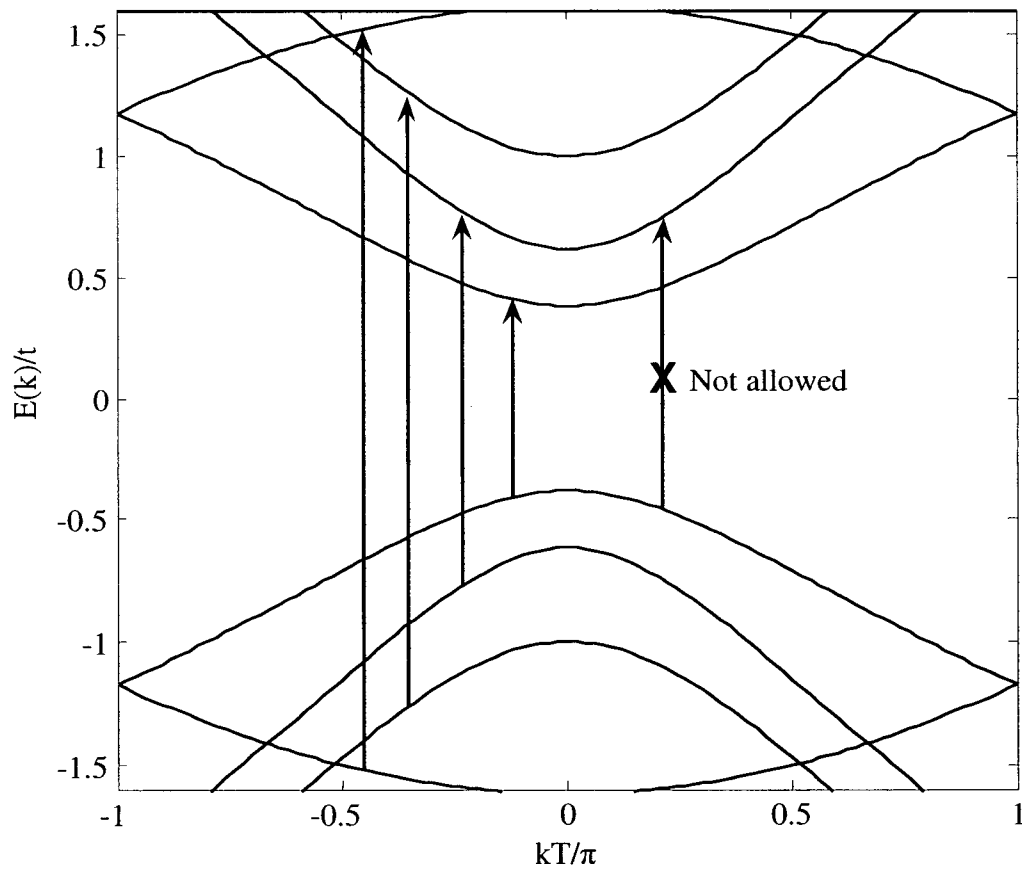
Figure 2.8: The Energy dispersion curves of a zigzag (5,0) SWNT. The arrows show the allowed optical transition for z-polarized light (E ∥ z). Horizontal and intraband transitions are not allowed and only $E_{\mu\mu}$ transitions are allowed due to the conservation of angular momentum.

The suppression of optical transitions induced by an electric-field perpendicular to the tube axis by depolarization is explained in reference [19] as follow: consider an infinitely long and narrow cylinder in an applied electric field. For an external field applied along the cylinder axis, no bound charges are induced on the cylinder wall and the total parallel electric-field is equal to the external field. In the case of an external field applied perpendicular to the cylinder axis, bound charges are induced on the cylinder walls. The resulting induced electric-field opposes the external field and reduces the total perpendicular electric-field. In the case of SWNTs, the locally induced electric-field almost completely cancels the perpendicularly applied electric-field. Based on this explanation, it is expected that the optical response of SWNTs to perpendicular excitation would be much weaker as compare to parallel excitation. A consequence of the depolarization of the external electric field is that absorption-related measurements on randomly oriented SWNTs are completely governed by their properties for z-polarization [17].

SWNTs also exhibit absorptive and refractive nonlinearities. These nonlinear effects are mainly due to interband transitions of electrons across the band gaps (shown in figure 2.8) and the combined motion of electrons associated with intraband-interband transitions [20]. Theoretical calculations performed by Margulis et al. [21] show that the third order nonlinear susceptibility, $\chi^{(3)}$, of SWNTs is proportional to the tube diameter and is greatly enhanced under resonant excitation (i.e., the excitation energy matches the band gap energy of the material), reaching values on the order of 10^{-6} esu. The values of $\chi^{(3)}$ under resonant excitation are three orders of magnitude higher than those under off-resonant excitation (10^{-9} esu [20]). This is because under off-resonant excitation, optical

nonlinearities are created by virtual excitations, where no real carrier populations are generated [20].

2.5 Summary

In this chapter, the basic properties of SWNTs are described. Information about the band gap energies of SWNTs can be obtained from their energy dispersion. The obtained information can be used to estimate the tube diameters of SWNTs in a sample by analyzing the sample's absorption spectrum. Accurate estimates of the band gap energies are also important when designing SWNT based device. As shown in the Kataura plot, the band gap energies depend upon the tube diameters. This means that the enhancement of optical nonlinearities at the resonant frequency of SWNTs is available at all energy levels. This useful property makes SWNTs an ideal material for producing nonlinear optical device.

CHAPTER 3

EXPERIMENT

3.1 Introduction

In this chapter, details of the experimental techniques, the measurement system, and the SWNT sample fabrication process are presented. An overview of the Power-scan (P-scan) method and the pump-probe measurement [17] is given. A detailed description of the construction of the measurement system is provided. Finally, the SWNT sample fabrication process is described and the quality of the fabricated sample evaluated.

3.2 Measurement Techniques

Due to its simplicity in implementation and data analysis, the open aperture Z-scan method [22] is a popular technique for measuring absorptive nonlinearity. However, the open aperture Z-scan method is only suitable for measuring highly uniform samples, which is not the case for the sample used in this work. Instead, a modification of the open aperture Z-scan method, referred to in this work as the Power-scan (P-scan) method, is used. Optical bleaching in SWNTs can be observed by the pump-probe measurement method. This section provides a brief overview of the P-scan and the pump-probe measurement methods. Since the P-scan method is based on the same principles as the open aperture Z-scan method, an overview of the open aperture Z-scan method is also provided.

3.2.1 The Open Aperture Z-scan Method

An illustration of an open aperture Z-scan measurement apparatus is shown in figure 3.1.
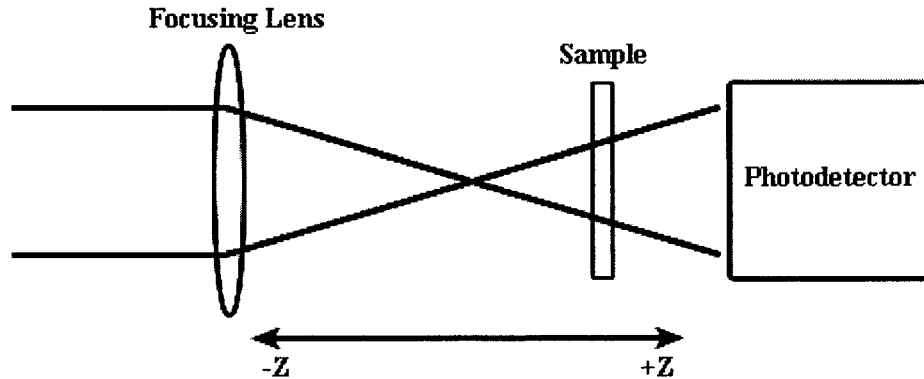
Figure 3.1: The open aperture Z-scan measurement apparatus.

In this method, a thin medium, the sample of interest, is placed in the path of a Gaussian beam focused by a lens. The sample is translated along the axis of propagation (z-axis) of the beam, starting from the pre-focal region (negative z) to the focal point (z = 0), and then to the post-focal region (positive z). At the far-field a large area photodetector measures the power of the beam passing through the sample as a function of the sample's z position. The measured beam power is divided by the weak-field beam power (measured when the sample is far away from the focal point, where no optical nonlinearity is induced) to give the normalized transmittance of the sample. As the sample is moved along the path of the beam, light intensity (measured in watts per cm^2) in the sample changes as the beam diameter changes along the z-axis. The change in light intensity induces a change in the absorption coefficient of the sample. The absorption coefficient of a material as a function of light intensity is given by [22]

$$\alpha(I) = \alpha_0 + \beta I \tag{3.1}$$

where $\alpha(I)$ is the total absorption coefficient, α_0 is the linear absorption coefficient, β is the nonlinear absorption coefficient, and I is the light intensity. Since $\beta \ll \alpha_0$, under low intensity the effect of the term βI is negligible and $\alpha(I) \approx \alpha_0$. However, as the intensity increases, the term βI may no longer be neglected. When the sample is far from the focal point, in the pre-focal region, the light intensity in the sample is low and the resulting βI is negligible; hence the sample's transmittance remains constant. As the sample approaches the focal point, the light intensity in the sample increases and $\alpha(I)$ is changed by βI. Depending on the sign of β, the measured transmittance will either increase (saturable absorption) or decrease (reverse saturable absorption) as the sample approaches the focal point.

Shown in figure 3.2 are examples of the open aperture Z-scan traces of samples exhibiting saturable and reverse saturable absorption plotted in solid and dashed line, respectively. Depending on the nature of the absorption nonlinearity, the open aperture Z-scan trace may either have a peak (saturable absorption) or a valley (reverse saturable absorption) at the focal point. The obtained open aperture Z-scan trace can be analyzed to deduce the value of β.
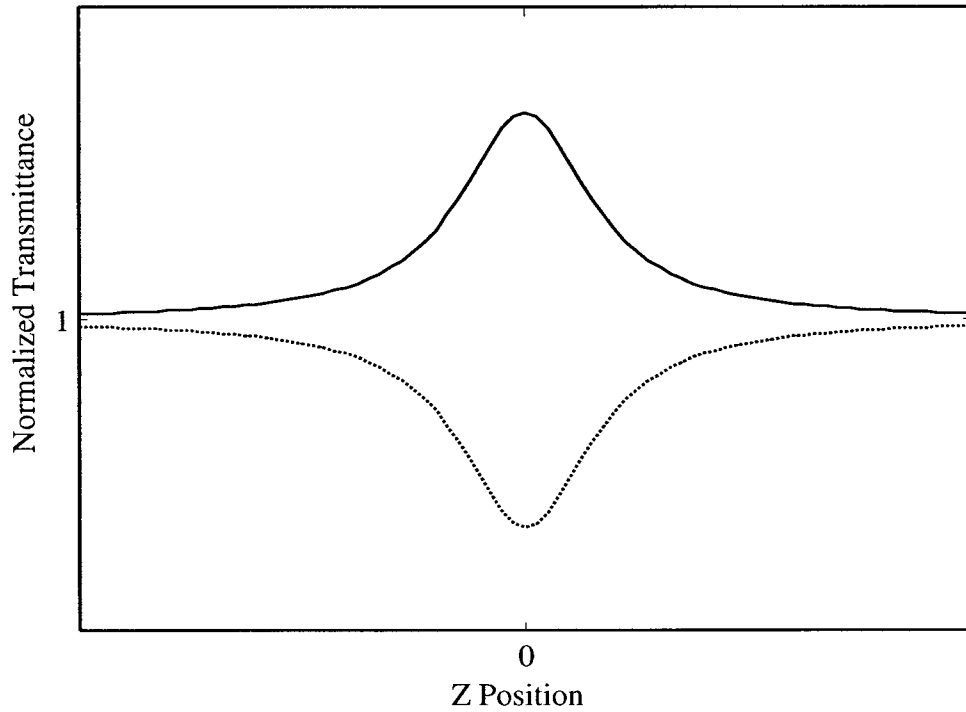
Figure 3.2: Open aperture Z-scan traces of materials exhibiting saturable and reverse saturable absorption plotted in solid and dashed line, respectively. The peak (valley) at the focal point is due to saturable absorption (reverse saturable absorption).

In general, the results obtained by the Z-scan method are only valid if the sample of interest meets the following two requirements: (1) the surface of the sample must be optically uniform down to the smallest spot size (the smallest diameter of the focused beam); (2) the thickness of the sample must be less than $z_R = \pi w_0^2/\lambda$, the Rayleigh range of the focused Gaussian beam [23], such that the sample can be treated as a thin lens. In the equation for z_R, w_0 is the beam waist radius and λ is the laser wavelength.

The requirement to have a highly uniform surface, imposed on a sample by the Z-scan method, is strict. This requirement is especially difficult to meet when a small spot size is used. For a sample that has a non-uniform surface, the Z-scan trace fluctuates in accordance to the changing surface properties of the sample, making it impossible to

extract any useful information from the Z-scan trace. The direct solution to the surface non-uniformity issue is to use a spot size that is large enough for the surface non-uniformity of the sample to average out. This solution, however, cannot be applied in this work since the fabricated SWNT sample requires a large spot size (~160 μm) in order for the sample's surface to appear to be uniform. This large spot size requires a very high power laser to generate sufficient intensity in the sample to induce observable nonlinear optical effects. The surface non-uniformity of the fabricated SWNT sample will be discussed in section 3.4.5.

Another solution is to modify the open aperture Z-scan method so that a sample with a highly uniform surface is not required. One such modification is referred to here as the Power-scan or P-scan method.

3.2.2 The P-Scan Method

The P-scan method works on the same principle as the open aperture Z-scan method. However, instead of varying the applied intensity by changing the beam size in the sample, the power of the laser is changed. In the P-scan method, the sample is placed at the focal point of the focusing lens. The transmitted power of the sample is measured as a function of the applied intensity. After a scan on the sample, another scan is performed on the substrate (without SWNTs) to serve as a reference. Dividing the scan of the sample by that of the substrate yields the transmittance curve of the sample as a function of intensity, which can be analyzed to obtain information about the nonlinear absorption coefficient.

The transmittance curve, T_r, of the sample obtained by the P-scan method at a particular location in the sample can be analyzed by applying Beer's law equation for transmittance given by

$$T_r = \exp(-\alpha(I)L) = \exp(-(\alpha_0 + \beta I)L) \tag{3.2}$$

where L is the sample thickness, and $\alpha(I) = \alpha_0 + \beta I$ (see section 3.2.1). β is assumed to depend hyperbolically on intensity and is given by [24]

$$\beta(I) = \frac{\beta_0}{1 + \dfrac{I}{I_{sat}}} \tag{3.3}$$

where β_0 is the true nonlinear absorption coefficient and I_{sat} the saturation intensity. I_{sat} is defined as the intensity where the change in T_r has reached 50% of its modulation depth. Based on equation 3.2 and 3.3, β_0 can be deduced from the trace of a P-scan if I_{sat}, L and α_0 of the particular location are known.

The advantage of the P-scan method is that surface non-uniformity no longer affects the analyzability of the obtained data. The disadvantage of this method is that the results are localized (i.e., different locations in the sample may yield different responses due to the varying surface properties). This would make interpreting the experimental results difficult, which will be discussed in chapter four.

3.2.3 The Pump-Probe Measurement

Measurements of ultrafast phenomenon, such as the recovery time of a material, are done by the pump-probe measurement method. Shown in figure 3.3 is an illustration of a typical pump-probe measurement setup.
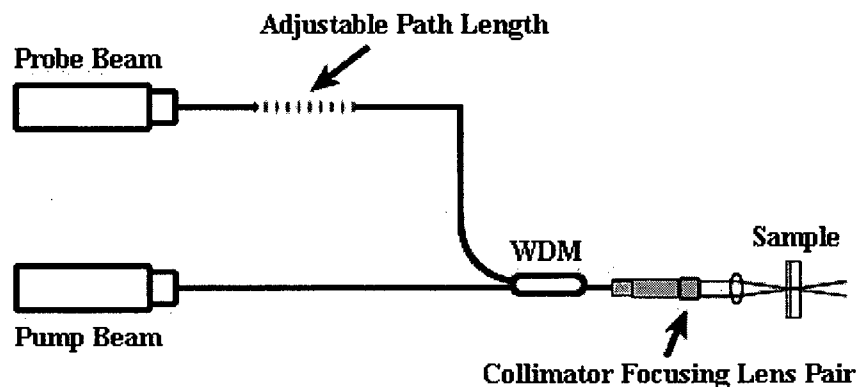
Figure 3.3: A typical experimental setup of the pump-probe measurement. WDM stands for wavelength division multiplexer.

In the pump-probe measurement, two pulses, one with intensity much larger than the other pulse, are projected onto the sample of interest. The larger pulse, the pump beam, arrives first at the sample and generates some degree of excitation in the sample. After an adjustable time delay (t > 0) relative to the arrival time of the leading edge of the pump beam, the weaker pulse, the probe beam, arrives at the sample and its transmittance is measured. The time delay is adjusted by controlling the path length of the probe beam. The measured probe beam transmittance is recorded as a function of delay time and information, such as the recovery time of the sample, can be extracted from the recorded data. The wavelengths of the beams do not need to be identical; pump-probe measurements using two different wavelengths are referred as two-color pump-probe measurements. The time resolution of the pump-probe measurement is controlled by the pulse width: a shorter pulse will yield finer time resolution (ex, a 1 ps pulse width gives a time resolution of 1 ps).

A weak continuous wave laser can also be used as the probe beam, given that the power of the continuous wave laser is sufficiently weak that it will not generate significant excitation in the sample. The advantage of using a continuous wave laser as the probe beam is that a delay circuit (see figure 3.3) is not required.

3.2.4 Summary

The P-scan method offers an excellent means to observe saturable absorption in SWNTs as well as characterizing the absorptive nonlinear behavior in terms of β_0. Since the fabricated SWNT sample does not have a highly uniform surface, the P-scan method is suitable for this work. The pump-probe measurement used in this work is a two-color pump-probe method that uses a continuous wave laser as the probe beam. The design of the measurement system to perform the two described measurement methods is presented in the next section.

3.3 Measurement System

Both the P-scan and the pump-probe measurement methods, described in the previous section, require a measurement system that is capable of delivering high intensity light to the sample of interest. As previously defined, intensity is the power of the laser divided by the beam area; thus, the main concerns in implementing the measurement system are to minimize the spot size and to maximize the laser power. Shown in figure 3.4 is a schematic of the implemented measurement system. This measurement system is capable of performing Z-scan, P-scan, and pump-probe measurements. The measurement system can be divided into five subsystems: (1) the pulse generator, (2) the pulse amplifier, (3) the probe beam source, (4) the lens system, and (5) the demultiplexing circuit. Components (1)-(3) and (5) are part of the laser system. The details of the designs of each subsystem, as well as how they each operate, are presented in the following subsections.
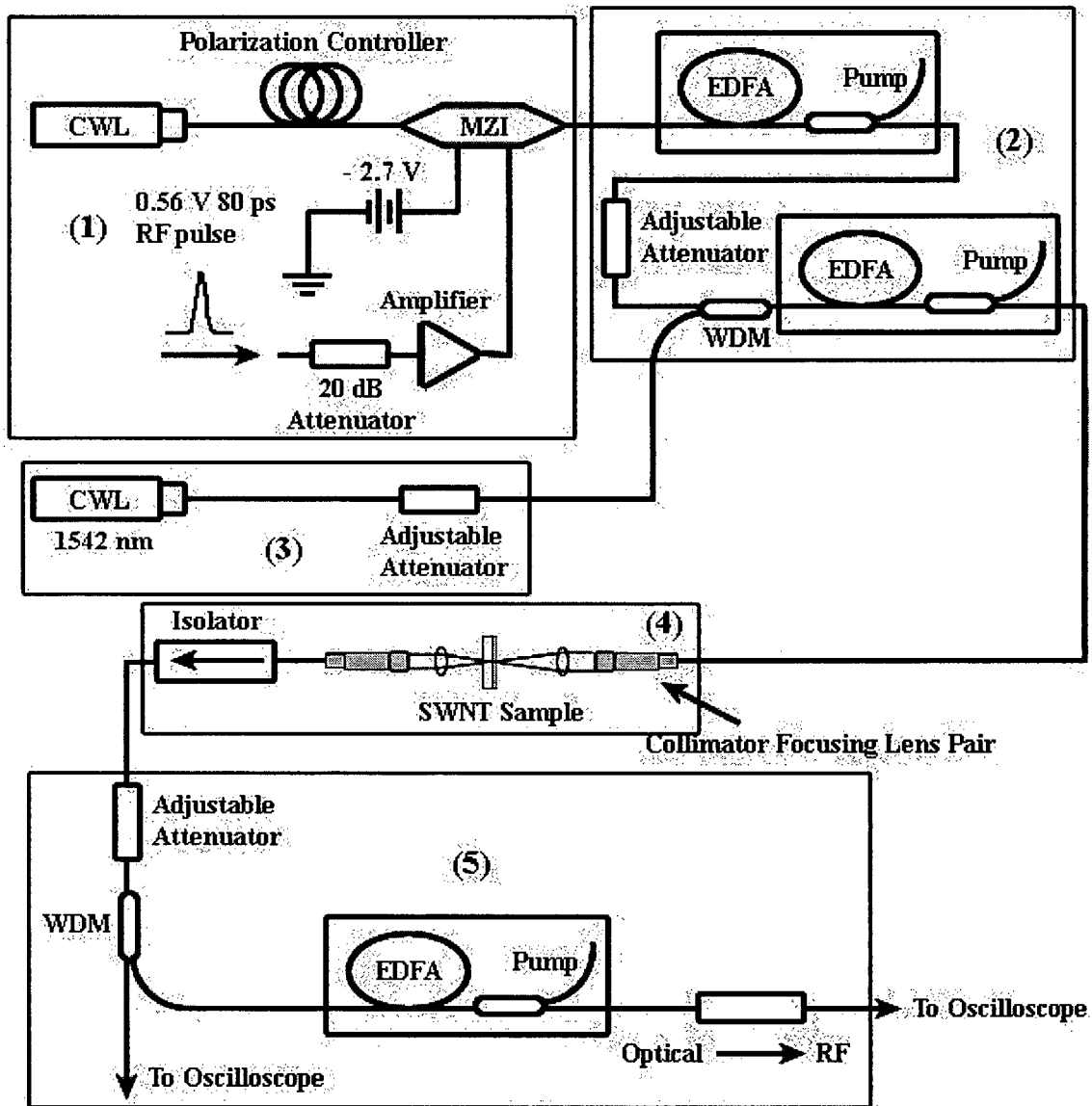
Figure 3.4: A schematic of the implemented measurement system. This measurement system can be used to perform the Z-scan, the P-scan, and the pump-probe measurement.

3.3.1 Measurement System Design Goal

In the work of Set et al. [8], saturable absorption in their SWNT sample at 1550 nm was observed for intensities ranging from 0.036 to 5.8 MW/cm^2. Their results suggest that higher modulation depth can be obtained if higher intensity is applied to their sample; based on this observation, in order to measure saturable absorption in SWNTs, the measurement system should achieve a maximum intensity of no less than 5.8 MW/cm^2 at 1550 nm.

3.3.2 Lens System Design

Shown in figure 3.5 is the implemented lens system ((4) in figure 3.4).
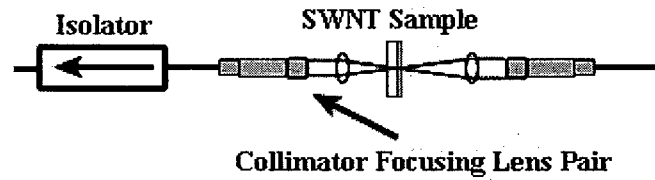
Figure 3.5: The designed lens system. Laser light is focused to a spot size of 6.06 μm using a collimator and focusing lens pair.

In the lens system, laser light from a single mode fiber was collimated using a fiber optic collimator. The collimated beam was then focused to a small spot using a focusing lens. The collimator generated a collimated beam diameter of 2650 μm and the focusing lens has a focal length of 6.24 mm. The spot size produced by the collimator and focusing lens pair was calculated to be 6.06 μm. The calculation of the spot size is discussed later. At the receiving end, an identical collimator and focusing lens pair was used to couple the laser light into another single mode fiber. Aspheric lenses were used

in both the collimators and the focusing lenses to eliminate spherical aberration. The two collimator and focusing lens pairs were each mounted on a translation stage with a displacement resolution of 10 µm. An inline optical isolator was installed at the receiving end to prevent any reflection from entering the lens system. The lenses were also anti-reflection coated to prevent formation of standing waves. To increase the stability of the lens system, the assembly was mounted on an optical bench with vibration control. The assembly was also shielded with black thick cardboard to block out external light from interfering with the system. The stability of the lens system was tested with a logging application installed in an Agilent 8163a Lightwave Multimeter. Test result shows that the lens system maintained its stability with a 1% error in measured power. The sample was secured into position between the two focusing lenses with a plastic clip, which was mounted on another translation stage so that the sample can be positioned at any locations between the two focusing lens.

The spot size, D_{sR}, produced by a focusing lens is given by

$$D_{sR} = \frac{4\lambda f M^2}{\pi D_R}$$ (3.4)

where D_R is the collimated input beam diameter, λ is the laser wavelength, f is the focal length of the focusing lens, and M^2 is the beam quality factor. The derivation of equation 3.4 along with a review of Gaussian beam optics is given in appendix A.

The beam quality factor, M^2, is a measure of how much the laser light's transverse intensity profile deviates from that of a Gaussian beam. The first step in spot size calculation is to measure the M^2 parameter of the laser light emitted from a single mode fiber. M^2 is given by [25]

$$M^2 = \frac{\theta_R \pi w_{0R}}{\lambda}$$ (3.5)

where θ_R and w_{0R} are the half angle of divergence and the beam waist radius of a focused laser beam, respectively. The steps in measuring θ_R and w_{0R} are very simple. Laser light from a single mode fiber is collimated and then focused. w_{0R} and the Rayleigh range, z_{RR}, of the focused laser beam are first measured. Then at far-field ($z > 10z_{RR}$), the radius of the laser beam is measured at two different locations to determine θ_R. The measurement of θ_R must be done at distances far away from z_{RR}, since the measured θ_R near z_{RR} has not approached the limit described by $\theta_R = \lambda M^2/\pi w_{0R}$ [25].

The instrument used for measuring the M^2 parameter and the spot size was the Thorlabs WM100B Omega Meter Beam Profiler. Many measurements were performed using different focusing lenses to give an average M^2 value of 1.31. The following table lists some data obtained from the M^2 measurement.

Table 3.1: Sample M^2 measurement data

f (mm)	w_{0R} (μm)	w_{1R} (μm)	w_{2R} (μm)	Δz (mm)	θ_R (rad)	M^2
11	32	250	282	0.4	0.0400	1.30
8	24	231	272	0.4	0.0524	1.28
18.4	51	363	384	0.4	0.0262	1.36

To verify the accuracy of the measured M^2 value, the spot size of a collimator and focusing lens pair producing a known spot size of 6 μm was measured. Using equation (3.5), a M^2 value of 1.31, and the measured θ_R of 0.21 radian, the spot size was calculated to be ~6.2 μm, which is in close agreement with the actual value. Spot sizes of various collimator and focusing lens pairs were calculated using equation (3.4) with $M^2 = 1.31$, and the resulting spot sizes are in close agreement with the actual spot sizes measured

with the beam profiler. Having determined the value of M^2, the spot size produced by the collimator and focusing lens pair (D_R = 2650 µm and f = 6.24 mm) at the optical communication wavelength was calculated using equation 3.4 to be 6.06 µm.

3.3.3 Laser System Design

In the previous section, the design of the lens system was described and a spot size of 6.06 µm was achieved. This indicates that the laser system must deliver a minimum peak power of 1.67 W to meet the required minimum intensity of 5.8 MW/cm^2 (see section 3.3.1). In this section, the design of the laser system is described. A brief overview of the design concept of the laser system is presented. After that, the designs of each of the subsystem in the laser system (see section 3.3) are described in the following order: the pulse generator, the pulse amplifier, the probe beam source, and the demultiplexing circuit.

3.3.3.1 Design Concept

Laser amplification is achieved by passing the laser light through a gain medium such as an erbium doped fiber amplifier (EDFA). An EDFA is a stretch of erbium doped fiber with energy constantly being pumped into the fiber to excite the erbium ions, creating population inversion. Laser light entering the EDFA stimulates the excited erbium ions to emit radiation at the same wavelength and phase as the input laser light. An EDFA can be viewed as an energy storage device. The stored energy remains in the device until it is emitted either by stimulated emission or by spontaneous emission (noise). If the stored energy can be released all at once in a short time frame, the resulting output power will

be very large. This can be done by sweeping the stored energy with a short pulse with a period equivalent to the charging time of the EDFA. The peak power of the amplified pulse will be inversely proportional to the pulse width. This pulse amplification mechanism is implemented in this work.

3.3.3.2 Laser Pulse Generation

Shown in figure 3.6 is the setup of the laser pulse generator ((1) in figure 3.4).
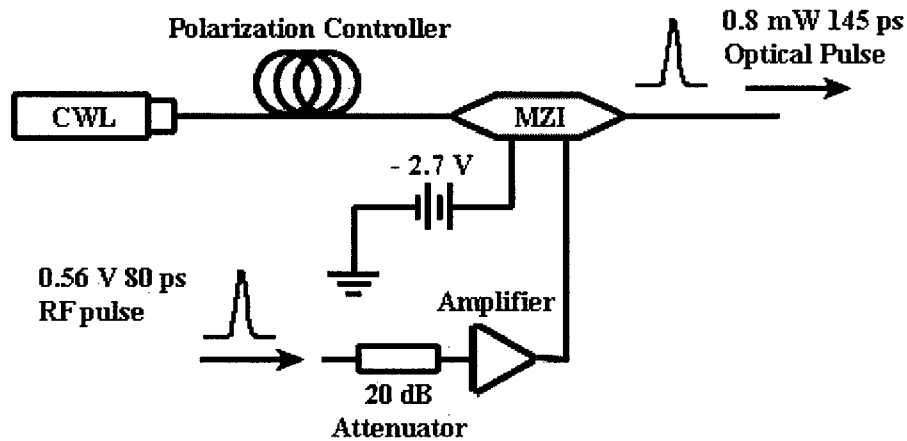
Figure 3.6: The laser pulse generation mechanism. The output from the continuous wave laser is modulated by a MZI. The MZI is biased to 0% output with a DC bias of approximately -2.7 V and the output was generated by a RF pulse.

Laser pulses were generated by modulating a continuous wave laser (Helwet Packard 81689A, 4.5 mW peak power) with an optical modulator. The modulator used was a 10 GHz Mach-Zehnder Interferometer (MZI). See reference [26] for a review on the operation of the MZI. The MZI was biased to 0% output by a DC voltage of approximately -2.7 V and the output was generated by a RF pulse. The MZI was biased to the minimum output level to minimize the output power when the pulse was off. This

minimizes stimulated emission in the EDFA during the charging stage and ensures that the EDFA was fully charged when the pulse arrived. The RF pulse was generated by a 12.5 GHz pulse pattern generator (Anritsu MP1763B). The generated RF pulse has a peak amplitude of 0.56 V and a pulse width of approximately 80 ps. The RF pulse was amplified using an RF amplifier (RF Bay, Inc. LPA-6-26) to produce a final RF pulse with a peak amplitude of ~3 V and a pulse width of ~145 ps. The pulse broadening was due to the lower bandwidth of the RF amplifier (6 GHz). The amplified RF pulse is shown in figure 3.7. The ringing effect shown in figure 3.7 is a typical response of an amplifier when amplifying a short RF pulse. The laser pulse generated from the setup is shown in figure 3.8. The laser pulse has a peak power of ~0.8 mW and a pulse width of 145 ps. The tail of the laser pulse shown in figure 3.8 is the result of the ringing effect shown in figure 3.7 coupled with the output behavior of the MZI. The signal generated has a base (minimum) output of 1.86 µW, which was low enough to minimize stimulated emission in the EDFA when the pulse was nominally off.
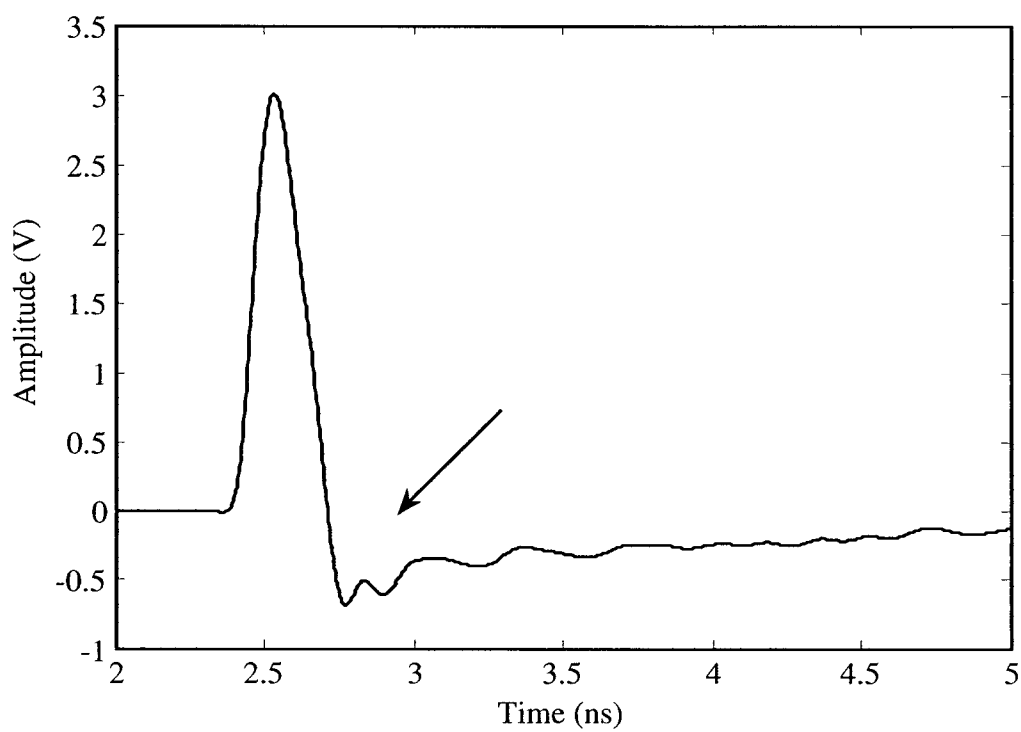
Figure 3.7: The amplified 145 ps 3 V peak amplitude RF pulse. The observed ringing at the tail of the pulse is a typical response of an amplifier when amplifying a short pulse.
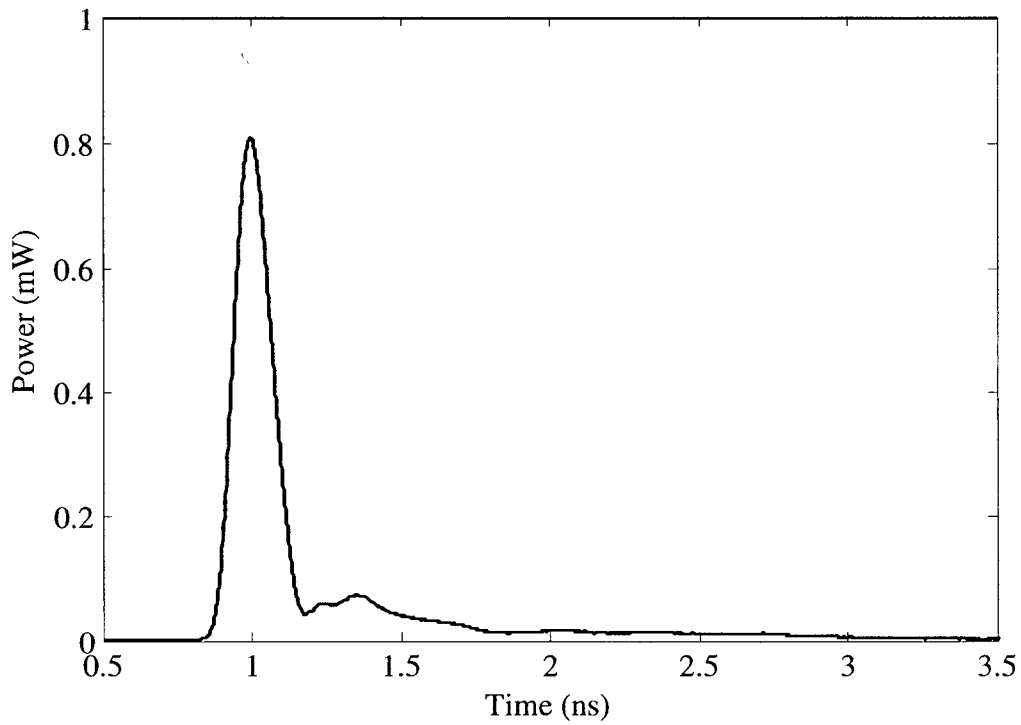
Figure 3.8: The laser pulse generated by pulse generation mechanism employing MZI. The pulse has peak power 0.8 mW and a pulse width of 145 ps.

3.3.3.3 Two Stages Laser Pulse Amplification

The laser pulse generated was amplified with a setup of two EDFAs (JDS Uniphase) arranged in cascade. The schematic of the pulse amplifier is shown in figure 3.9 ((2) in figure 3.4). The period of the pulse was set to the charging time of the EDFA. The EDFA charging time was determined experimentally by varying the pulse repetition rate and monitoring the peak power of the amplified pulse. Figure 3.10 illustrates the evolution of the pulse's peak power with respect to the pulse frequency.
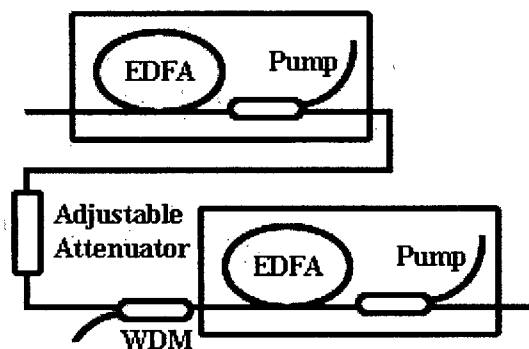
Figure 3.9: The pulse amplifier constructed by arranging two EDFAs in cascade.
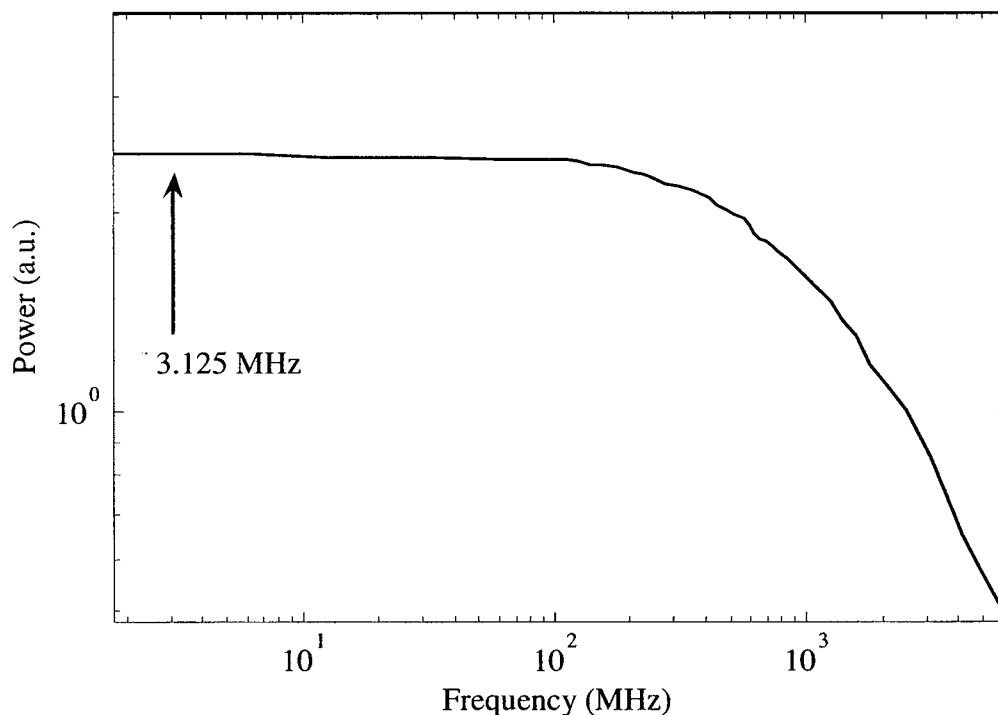
Figure 3.10: Peak Power vs. Pulse Frequency. The gain of the EDFA approaches a constant as the frequency of the pulse decreases.

It was determined that 3.125 MHz was the optimum pulse frequency to obtain the best gain from the EDFA. Each EDFA is capable of generating 20 mW average power before saturation; at a pulse rate of 3.125 MHz and neglecting noise, the energy stored in each EDFA would be 6.4 nJ. The laser pulse entered the first EDFA and then passed

through an adjustable optical attenuator (Agilent 8156A) before entering the second amplification stage. The attenuator provided a means of controlling the output power of the laser system and prevented the pulse from saturating the second EDFA. The saturation of the EDFA was mainly due to the high average/base power of the pulse. An attenuation of 5.5 dB (with the WDM removed) was required to prevent the pulse from saturating the second stage EDFA. Note that a WDM was placed between the optical attenuator and the second stage EDFA. The WDM was used in the pump-probe measurement to multiplex the probe beam into the measurement system. The WDM has an insertion loss of ~3 dB and, therefore, the optical attenuator was set to 2.5 dB. Since the output port of the first stage EDFA has high output isolation (> 35 dB), backward propagating noise from the subsequent EDFA can be neglected. The resulting laser pulse amplified by the two stage EDFA has a peak power of ~16.56 W and a pulse width of ~145 ps. The pulse amplification mechanism provided a gain of ~43 dB.

The energy stored in the pulse was approximately 2.4 nJ, giving an energy efficiency of the laser system of ~20%. This indicated that the implemented laser system should be able to deliver higher peak power. The first contributing factor to the low energy efficiency was the presence of noise, which consumed energy while the EDFAs were charging. The second contributing factor was the low peak power of the pulse generated by the MZI. Prior to saturation, the output power of the EDFA is proportional to the input power. Since the first EDFA was not saturated, a pulse with a much higher peak power (high signal-to-noise ratio) would yield better gain at the first EDFA, increasing the output of the entire laser system. While improvements can be made to further increase the output power of the laser system, the obtained power was already ~10

times higher than the minimum requirement of 1.67 W (see section 3.3.3), thus further improvements were not pursued.

3.3.3.4 The Pump-Probe Measurement Setup

For the pump-probe measurement, the pulse described in the previous section was used as the pump beam while a continuous wave laser at a wavelength of 1542 nm was used as the probe beam (see (3) of figure 3.4). 1542 nm was chosen as the probe beam wavelength because this is the maximum wavelength of the second channel of the WDM used to combine the pump and the probe beam. The probe beam was multiplexed into the laser system prior to the second stage EDFA (see figure 3.9). This was done so that the laser pulse would not breach the damage threshold of the WDM (200 mW). The prevention from damaging the WDM comes with the cost of reduction in the output power of the system. The probe beam, combined with the average power of the pulse from the first stage EDFA, saturated the second stage EDFA. To prevent saturating the second stage EDFA, both the pulse amplified in the first EDFA stage and the probe beam were attenuated by 9 dB and 11 dB, respectively. This gave the pulse peak power of 2.08 W and the average power of the probe beam was estimated to be 0.4 mW measured at the output of the second stage EDFA. To demultiplex the probe beam from the pump beam, another WDM was employed. The demultiplexing circuit ((5) in figure 3.4) is shown in figure 3.11.
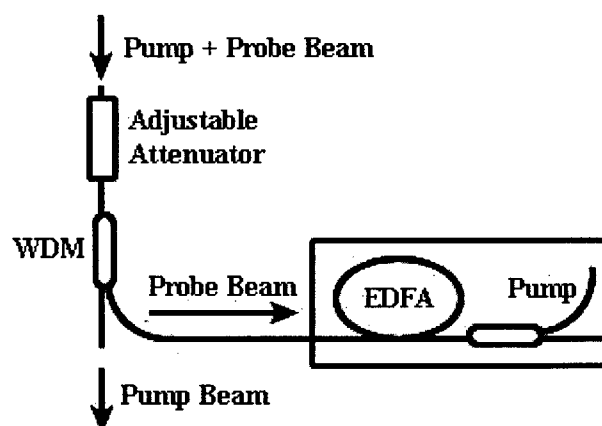
Figure 3.11: Setup for demultiplexing the probe beam from the pump beam.

The demultiplexed probe beam was amplified by an EDFA before entering the digitized oscilloscope. The amplification of the probe beam at the receiving end was necessary because it was expected that the probe beam would be attenuated significantly due to large attenuation of the pump and probe beams prior to demultiplexing to prevent damaging the WDM and the photodetector. Not shown in figure 3.11 is a 45 GHz optical to RF signal converter. The converter was installed at the probe beam output of the demultiplexing circuit to allow the use of high speed digitizing oscilloscope in the pump-probe measurement.

3.3.4 Measurement System Specifications

In this work, a measurement system capable of performing Z-scan, P-scan, and pump-probe measurements was developed and implemented. The measurement system consisted of a laser system and a lens system. The laser system is capable of generating a 145 ps pulse with a maximum peak power of 16.56 W at the optical communication wavelength. The repetition rate of the laser pulse was 3.125 MHz. The lens system consisted of two collimator and focusing lens pairs producing a minimum spot size of

42

6.06 μm. The maximum intensity of the measurement system was ~57 MW/cm^2. In pump-probe mode, the laser system generated a pump beam with a maximum peak power of 2.08 W. The 1542 nm probe beam multiplexed into the measurement system has a peak power of 0.4 mW measured at the output of the second stage EDFA. Shown in figure 3.12 is the complete measurement system with all of the components labeled and listed in table 3.2.
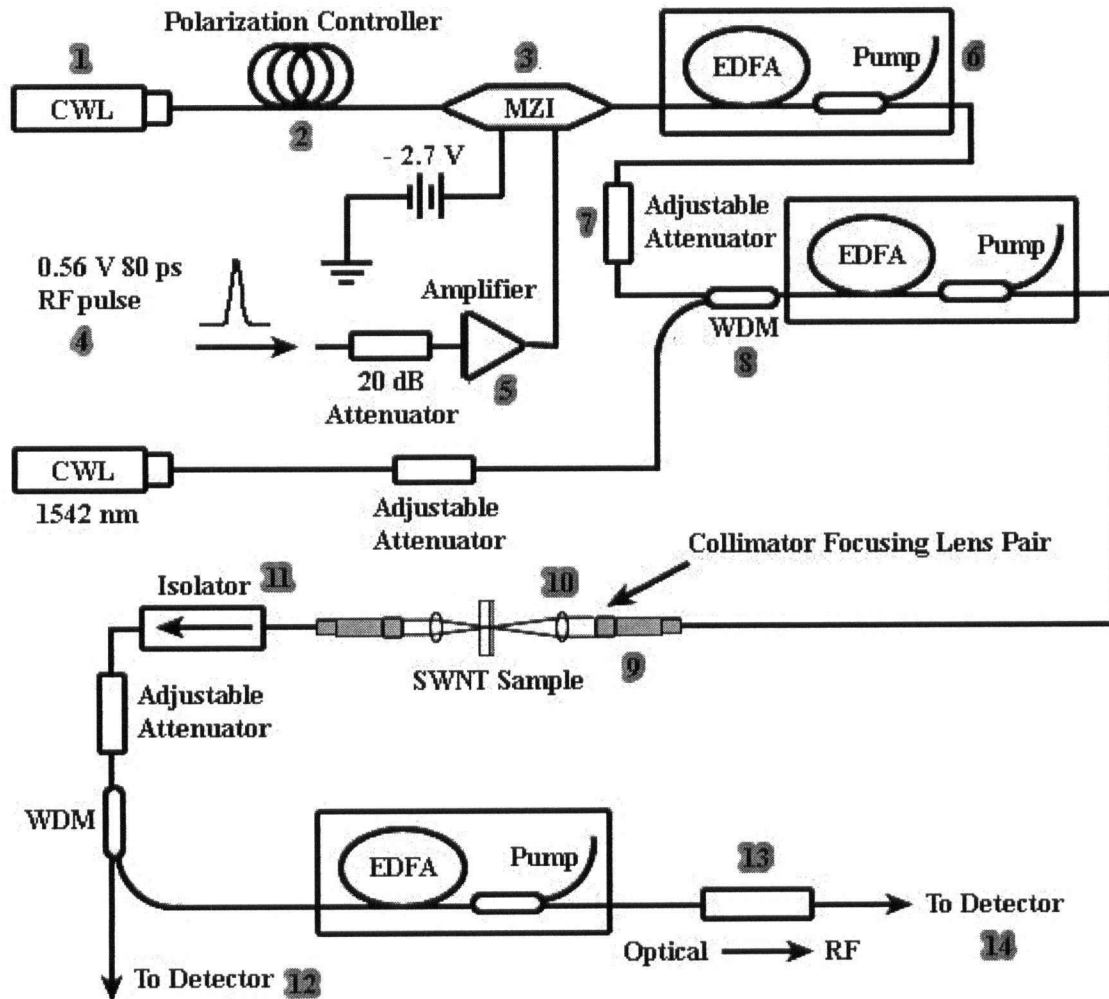
Figure 3.12: The schematic of the complete measurement system.

Table 3.2: List of components used in the measurement system

Label	Component
1	Agilent 8163A Lightwave Multimeter System with HP 81689A Tunable Laser Module
2	Thorlabs Polarization Controller
3	JDS Uniphase 10 Gb/s Intensity Mach-Zehnder Interferometer
4	Anritsu Pulse Pattern Generator MP1763B
5	RF Bay Inc. Low Power Amplifier 6-26
6	JDS Uniphase Erbium Doped Fiber Amplifier Module
7	HP 8156A Optical Attenuator
8	JDS Fitel Wavelength Division Multiplexer 1515RB-NT1
9	Thorlabs Fiber Optic Collimation/Coupling Packages F260FC-1550
10	Thorlabs Glass Aspheric Lens C110TM
11	Princeton Optics Optical Isolator 40-13-SA
12	HP 83480A Digital Communication Analyzer with HP 83485A Optical/Electrical Module 20 GHz Bandwidth
13	Lab Buddy Fiber Input Photodetector DSC10H
14	HP 83480A Digital Communication Analyzer with HP 83484A Dual Channel 50 GHz Electrical Plug-In Module

3.4 Single-Walled Carbon Nanotube Sample Fabrication

In order to measure saturable absorption in SWNTs using the P-scan method and to observe optical bleaching in the pump-probe measurement, the sample must operate in what is known as the "transparent mode". The produced film and the substrate must provide some degree of optical transparency and the sample surface should be highly uniform. In this chapter, the materials and methods used in the fabrication of the SWNT sample are described. The quality of the sample is evaluated to determine if the Z-scan method can be used. Part of section 3.4.1 and all of section 3.4.2 are based on materials provided by Chi Wah Eddie Fok, a graduate student from Dr. John Madden's molecular mechatronics group. The SWNT samples used in this work were also fabricated by Mr. Fok. The SWNTs used in the fabrication of the SWNT sample were provided by Dr. John Madden's research group.

3.4.1   Materials

The SWNTs used in this work were research grade SWNTs purchased from Carbon Nanotechologies Incorporated. These SWNTs were grown by the high pressure CO conversion process (HiPco) [27, 28] and were purified to contain 85 wt.% of SWNTs. The diameter distribution of the SWNTs ranges from 0.7-1.4 nm [28-30] with a mean diameter of 0.9 nm [29-32]. Note that the SWNTs used were not optimized to operate at 1550 nm. However, since SWNTs with a mean diameter of 1 nm (see section 2.3) are not readily available, the commercially available SWNTs were selected. SWNT film formation was realized by the filtration method [33]. The filtration method used a 0.2 micron pore size, 2.5 cm diameter hollow fiber nitrocellulose membranes and a make-up

solution of 1 wt.% Triton X-100 surfactant, pH 11 NaOH, deionized (18 MΩ) filtered (0.2 micron) water. The SWNT sample was fabricated on a sapphire window. Sapphire windows were selected as the substrate due to their high transparency, surface quality, optical flatness, and thermal conductivity.

3.4.2 Method

For SWNT film formation, 0.2 micron mixed cellulose ester (MCE) filter membranes were employed in a vacuum filtration apparatus. See figure 3.13 for an illustration of the vacuum filtration apparatus.
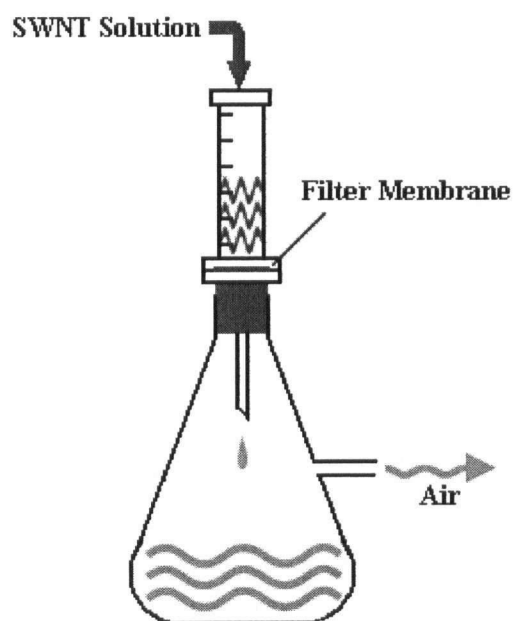
Figure 3.13: The vacuum filtration apparatus.

Prior to film formation, SWNTs were dispersed in 1 wt.% Triton X-100 solution to a concentration of ~2×10^{-3} mg/ml by ultrasonication for ~30 minutes. This was done in order to break and disentangle nanotube bundles. A thin film produced by the

filtration method has a typical density of 0.37g/cm^3 [34]. For the desired film thickness of 1 μm and a film area of $\pi(2.5/2$ cm$)^2$ = ~4.91 cm^2, the required amount of SWNT solution was calculated to be ~90.8 ml. Care was taken when pouring the SWNT solution into the filtering funnel to avoid the formation of bubbles on the solution surface, since bursting bubbles will disrupt the film when it is still wet and fragile. Once the solution was filtered down the vacuum was kept on for ~20 minutes to allow the newly formed SWNT film to dry. Residual surfactant left in the film was subsequently washed away with purified water. Near complete removal of the surfactant is indicated by the initially bubble laden permeate stream going water clear in the vacuum funnel.

Once all residual surfactant was washed away, the SWNT film formed was ready to be transferred to the substrate. Before transferring the SWNT film to the substrate, the area of the SWNT film was reduced to a quarter of its original size so that part of the substrate was left uncovered to serve as a reference surface. The SWNT film, while wetted with water and with the MCE membrane still attached, was placed in contact with the substrate and aligned as desired. The top of the MCE membrane was covered with layers of porous paper and a flat plate while another flat plate was placed at the bottom of the substrate. The two flat plates sandwiched the MCE membrane, SWNT film, and the substrate, applying a compressive load to the assembly to maintain the SWNT film's flatness as it dried. The surface tension of the water as the assembly dries brought the SWNT film and substrate into contact. The assembly was oven dried at 90°C for 45 minutes. When the assembly was removed from the oven, the SWNT film was adhered to the substrate and the SWNT sample was formed with the MCE membrane still attached. The MCE membrane was removed by placing the sample in an acetone bath for

20 minutes. This process was repeated several times to ensure complete removal of the MCE diffused into the sample in the first bath. The SWNT sample was then placed in a methanol bath and then a water bath. Finally, the SWNT sample was oven dried.

The resulting SWNT sample was a SWNT film with an estimated thickness of 1 µm adhered to a 1 mm thick sapphire window. The orientations of the SWNTs were randomly placed on the substrate such that the SWNT sample was polarization insensitive. This is adequate in that the measurement process would be simplified and the SWNT sample can be used as a simple drop-in device. The fabricated SWNT sample is shown in figure 3.14.
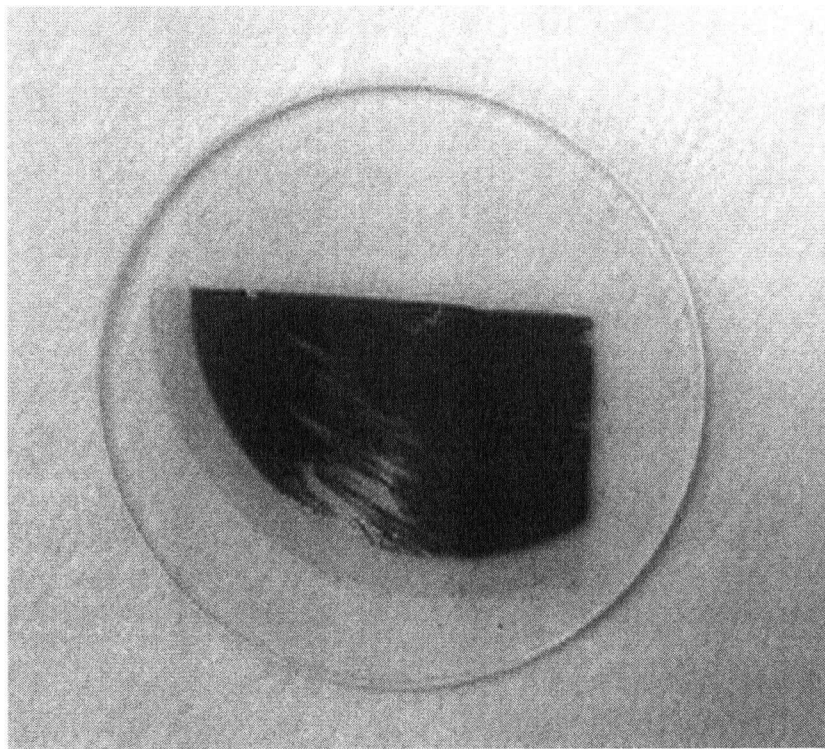
Figure 3.14: The SWNT film fabricated on a sapphire window. This picture is taken after the experiment, where the sample was damaged when transferring the sample into a microscope.

3.4.3   Quality of the Sample

The surface quality of the fabricated SWNT sample was tested by performing several trial Z-scans at low intensity.   The sample was placed in the lens system described in section 3.3.2 and a low power (~450 μW) continuous wave laser was used as the input power of the lens system. The power level of the laser was set at a low level so that the resulting intensity at the focal point (~0.002 MW/cm^2) was not large enough to induce any nonlinear effects in the sample. The Z-scan trace of the sapphire window was a constant transmittance showing no sign of optical nonlinearity, thus the result of a trial Z-scan on a highly uniform SWNT film was expected to be a constant. However, the Z-scan traces of the sample showed otherwise. Two typical results of the trial Z-scan on the sample are shown in figures 3.15a and b. At a far distance away from the focal point, the transmittance (the sample's transmitted power divided by the sapphire window's transmitted power) of the sample was approximately constant. As the sample approached the focal point and reached within a certain distance of the focal point (shown in figures 3.15a and b), the transmittance of the sample changed.   The transmittance either increased by a small amount and then fluctuated randomly or exhibited an effect similar to saturable absorption but with an exceptionally large gain. A simple explanation for the observed gain would be that, for a small spot size, the amount of SWNTs available for optical absorption is less than for a large spot size; therefore, transmittance for a very small spot size is expected to be higher.
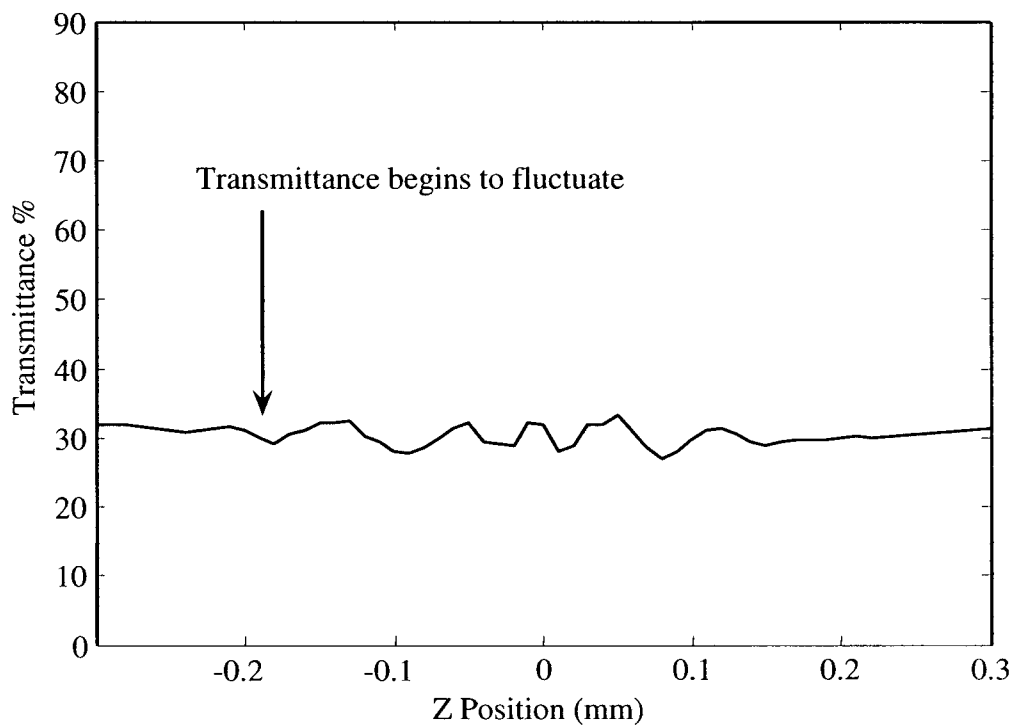
Figure 3.15a: Low power Z-scan trace of SWNT sample (fluctuating transmittance).
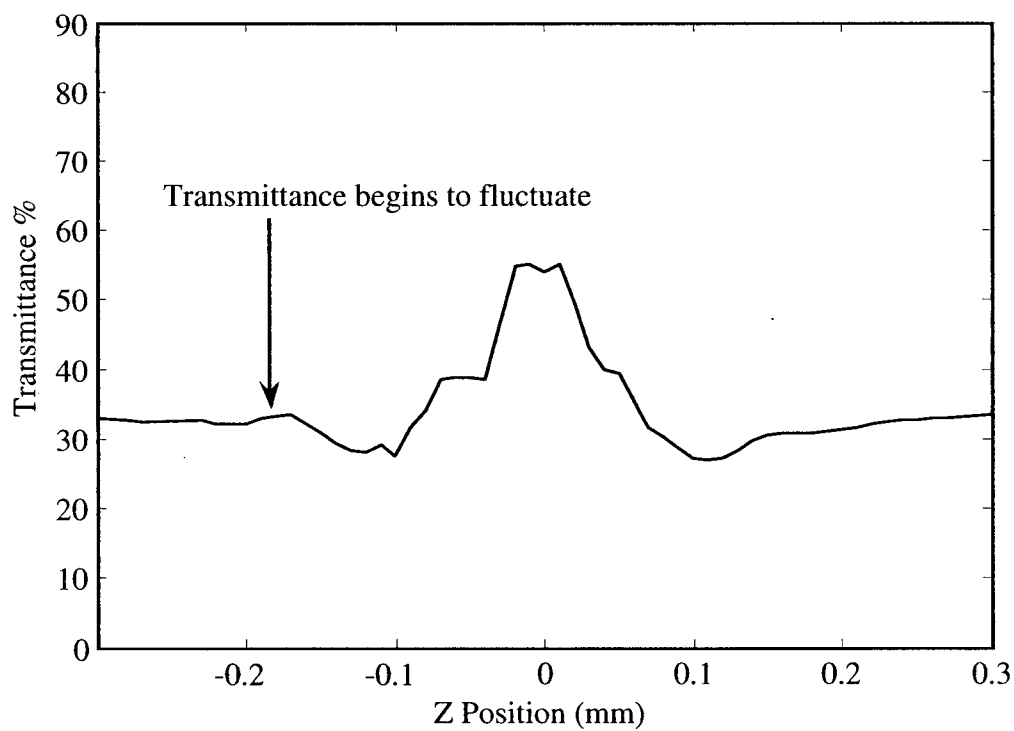
Figure 3.15b: Low power Z-scan trace of SWNT sample (peak at z = 0) .

3.4.4 Uniformity of Single-Walled Carbon Nanotube Sample

The fluctuation in transmittance in the Z-scan trace of the SWNT sample was primary due to the surface non-uniformity of the sample. This is supported by the resemblance of the pre-focal region and the post-focal region of the Z-scan trace shown in figure 3.15a and b. When the beam size in the sample is large, the surface non-uniformity in the sample averages out, and the sample appears to be uniform. When the beam size in the sample becomes small, the transmittance changes as a function of the changing properties per unit area on the sample surface; thus, the transmittance of the sample becomes localized, yielding a different transmittance when the beam waist is placed at different locations on the sample. Since the evolution of the beam size along the Z-scan axis is symmetric about $z = 0$, the transmittance of the sample as a function of z position is expected to show a large degree of symmetry about $z = 0$.

The factors that contribute to the surface non-uniformity can be determined using Beer's law. Beer's law equation for transmittance at low intensity is given by

$$T = \exp(-\alpha_0 L) \tag{3.4}$$

According to equation 3.4, the most obvious factor leading to surface non-uniformity in the sample is the variation in L across the sample surface. The least obvious factor is the variation in α_0. α_0 is dependent upon the surface density and, therefore, if surface density is not constant across the sample surface, α_0 is expected to be different at different locations in the sample.

Simple geometry can be applied to determine the uniformity of the sample. The half far-field divergence angle, θ, of a focused Gaussian beam is approximately constant

(see appendix A). Multiplying the tangent of θ with the distance from the beam waist, where the transmittance begins to fluctuate, gives the beam radius on the sample at the onset of non-uniformity of the sample as it approaches the beam waist. θ, in this case, was calculated to be approximately 24°. The z distance was extracted from the Z-scan traces (see figure 3.15a and 3.15b), giving an average of 180 µm. The beam radius on the sample when the transmittance begins to fluctuate was calculated to be ~80 µm. This indicates that the sample appears to be uniform for an area larger than $\pi(80\mu m)^2 = 2 \times 10^{-8} \, m^2$ or a 160 µm spot size.

3.4.5 Summary

A polarization insensitive SWNT sample was fabricated using the filtration method. The quality of the sample was tested and the results showed that the surface of the sample was rather non-uniform. The surface of the sample would appear to be uniform for a spot size larger than 160 µm, where the non-uniformity of the sample's surface averaged out. Due to the surface non-uniformity of the fabricated SWNT sample, the Z-scan method cannot be used due to the small spot size of the lens system (see section 3.2.1). Instead, the P-scan method described in section 3.2.2 can be used.

CHAPTER 4

RESULTS & DISCUSSIONS

4.1 Introduction

This chapter presents the measurement results obtained form the P-scan and the pump-probe measurements. The transmission spectrum of the SWNT sample is analyzed to estimate the mean tube diameter of the SWNTs in the sample. Data obtained from the P-scans are analyzed to estimate the nonlinear absorption coefficient of SWNTs. Although the implemented measurement system was unable to measure the ultrafast recovery time of SWNTs, optical bleaching was observed in the pump-probe measurement. A short discussion on the carrier dynamics of SWNTs is given. Following the discussion of the experimental results, a figure of merit that compares the performance of saturable absorbers based on their modulation depth, saturation intensity, and recovery time is proposed. The performance of several SWNT samples as saturable absorbers are compared with other multi-quantum well based saturable absorbers.

4.2 Absorption Spectrum

The absorption spectrum (700 to 1700 nm) of the SWNT sample was measured with a UV-VIS-NIR spectrophotometer (Varian Cary 5000). Shown in figure 4.1 is the transmission spectrum of the SWNT sample at low intensity. The inset shows the spectrum prior to background removal. The linear background is due to the presence of metallic SWNTs and the strong higher excitations of SWNTs.
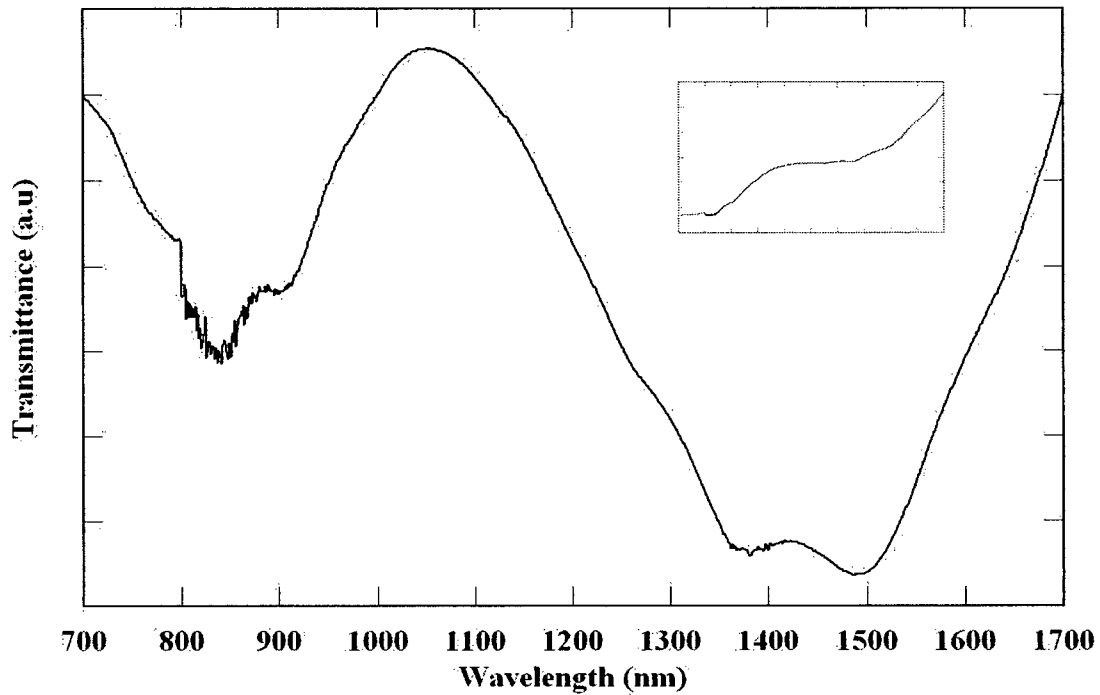
Figure 4.1: The transmission spectrum of the SWNT sample. Shown inset is the spectrum prior to background removal.

Two valleys are observed in the transmission spectrum. The first valley corresponds to the second band gap E_{22} and the second valley corresponds to the first band gap E_{11}, which is the focus of this work. At the second valley, two local minima located at 1381 and 1490 nm (0.9 and 0.83 eV) are identified. Referring to the Kataura plot (figure 2.7), these points correspond to tube diameters of 0.92 and 0.88 nm. The identified tube diameters are consistent with the expected mean diameter of SWNTs produced by the HiPco process (0.9 nm).

Figure 4.1 shows that the strongest absorption band is located at 1490 nm, which is close to 1550 nm. This shows that the SWNT sample is suitable for observing and measuring saturable absorption in SWNTs at 1550 nm.

4.3 P-scan Results

Thirty P-scans were performed at randomly selected locations in the SWNT sample. The transmittance (T_r) curves of the SWNT sample as functions of intensity were calculated from the obtained data and plotted in figures 4.2a, b, c and d. The T_r curves were categorized based on their modulation depth (ΔT) into three groups: $\Delta T > 10\%$, $7\% < \Delta T < 10\%$, $\Delta T < 7\%$.
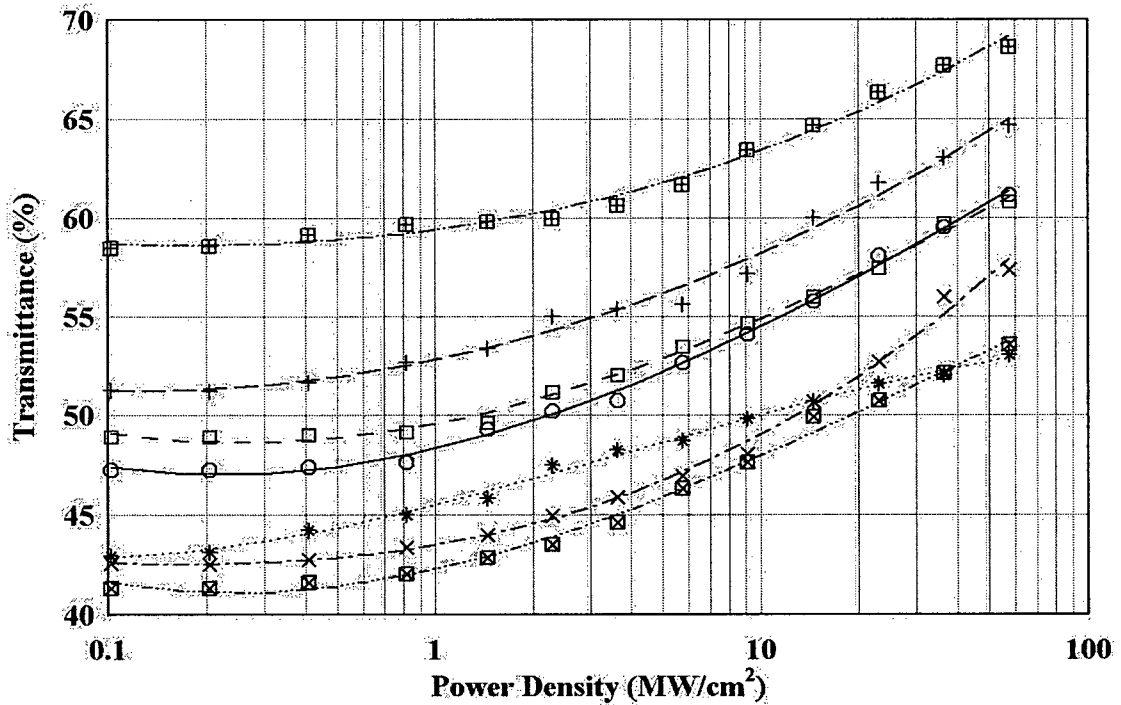
Figure 4.2a: Transmittance curves ($\Delta T > 10\%$).
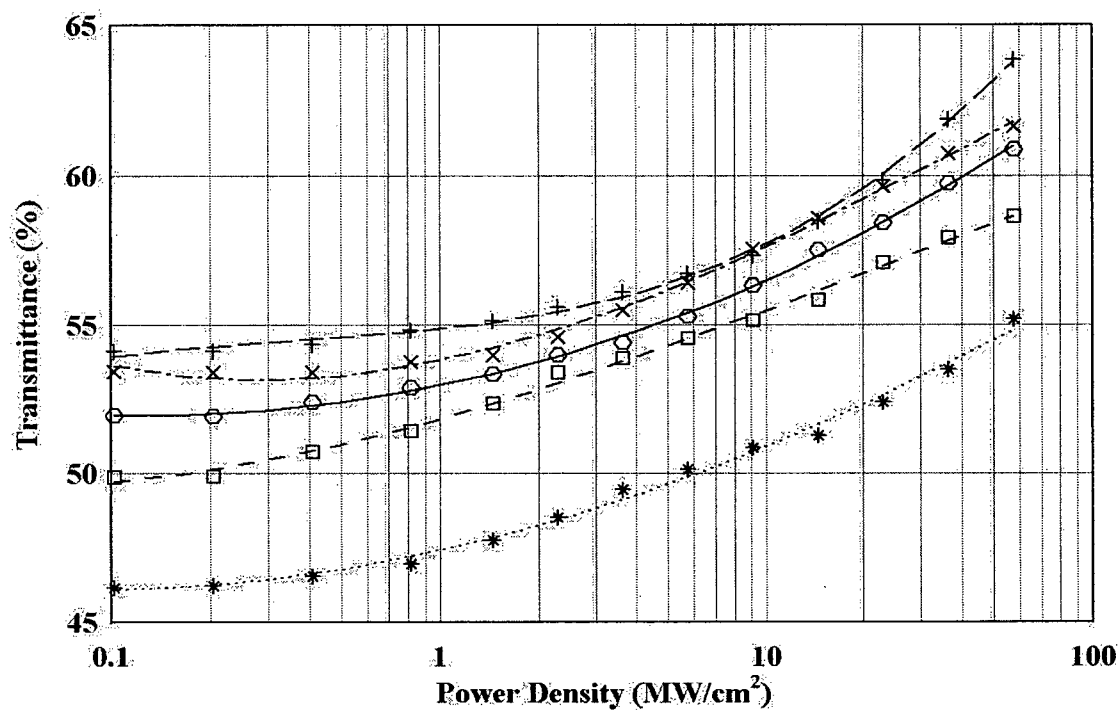
Figure 4.2b: Transmittance curves (7% < ΔT < 10%).
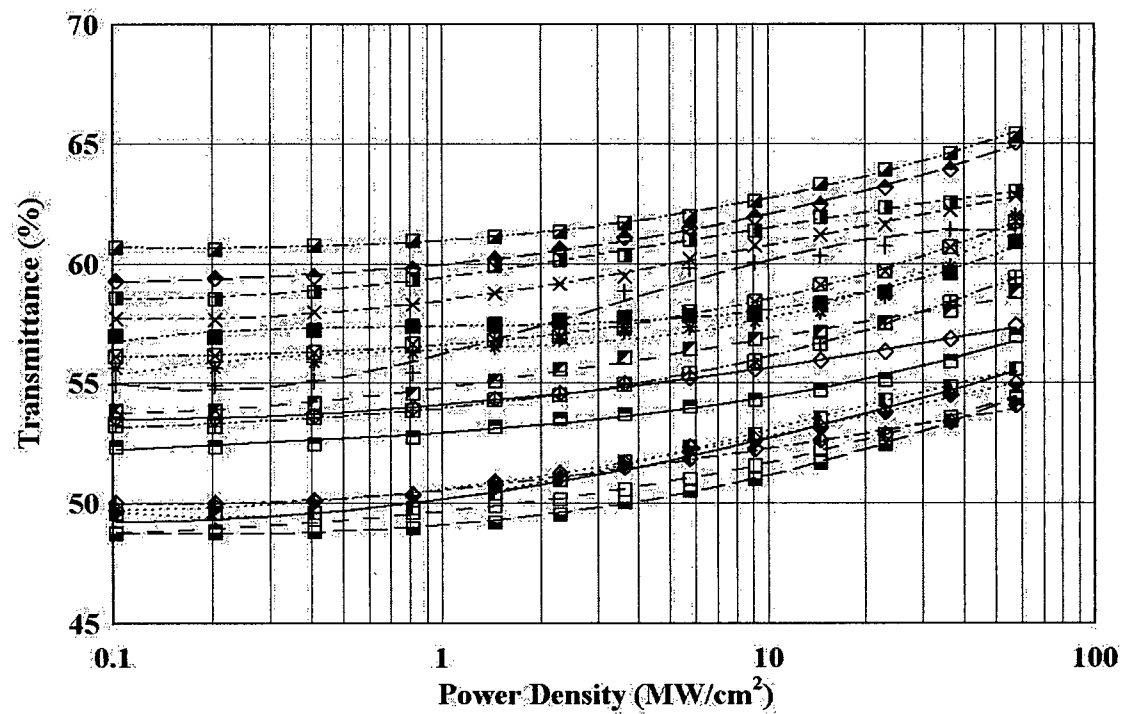
Figure 4.2c: Transmittance curves (ΔT < 7%)

56
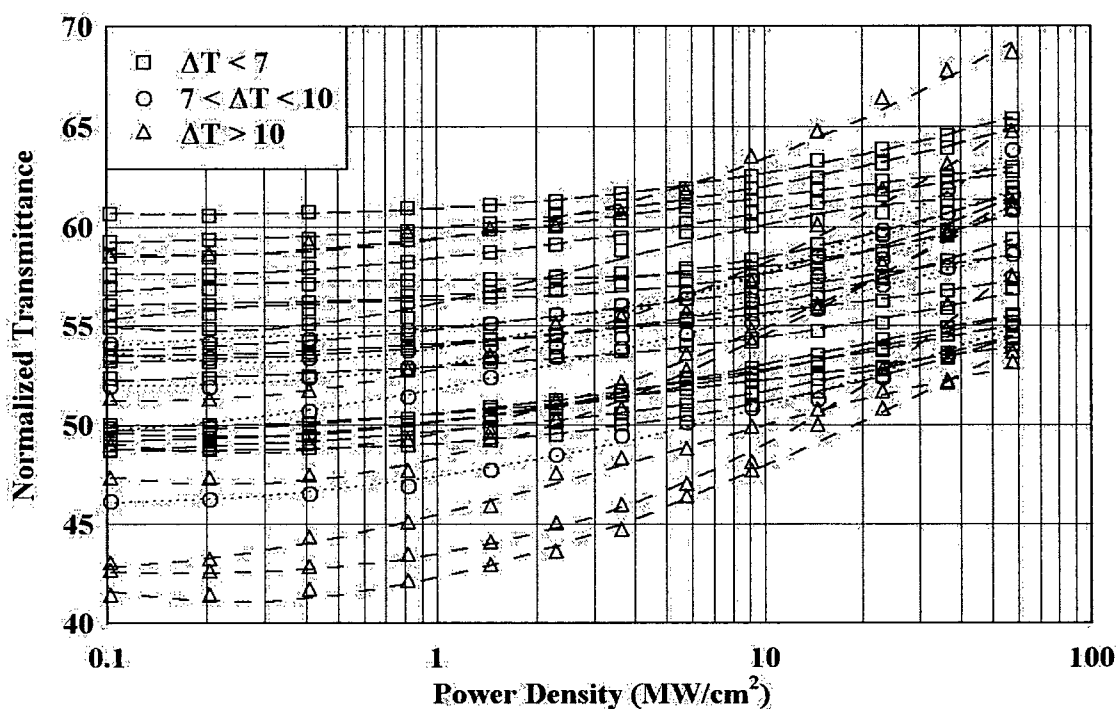
Figure 4.2d: Transmittance curves (all). Curves with high ΔT generally occupy the bottom portion of the plot while curves with low ΔT occupy top portion of the plot.

In these figures, the symbols represent the actual data while the solid lines represent curves fitted to the data. At high intensity, T_r ranges from 53 to 68%. As the intensity decreases, T_r decreases accordingly and reaches a constant value when the intensity is low. This weak field transmittance (WFT) ranges from 40 to 60%. The variation in WFT is due to the surface non-uniformity of the sample and the resulting localization of the response. Figure 4.2d shows that in general, T_r curves with lower WFT tend to have higher ΔT. ΔTs of up to 15% are observed in figure 4.2a. It should be noted that in all P-scans, the SWNT sample was never completely saturated, meaning that higher ΔTs can be achieved at higher intensities. However, since 57 MW/cm^2 was the highest intensity that could be applied to the sample before noticeable optical damage occurred, measurements with higher intensities were not performed. The observed

increase in T_r as the intensity increases is a clear indication of saturable absorption.
Normalizing each curve to their WFT gives the normalized transmittance (T_{norm}) curve
shown in figure 4.3a, b, c, and d.
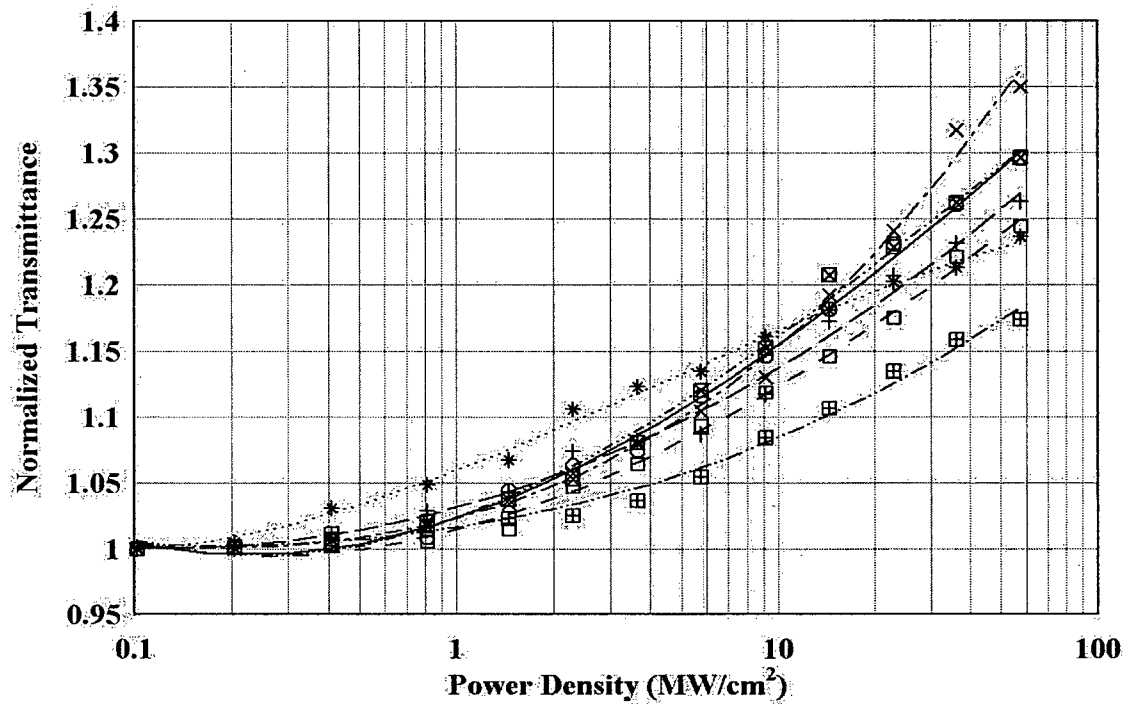
Figure 4.3a: Normalized transmittance curves ($\Delta T > 10\%$).
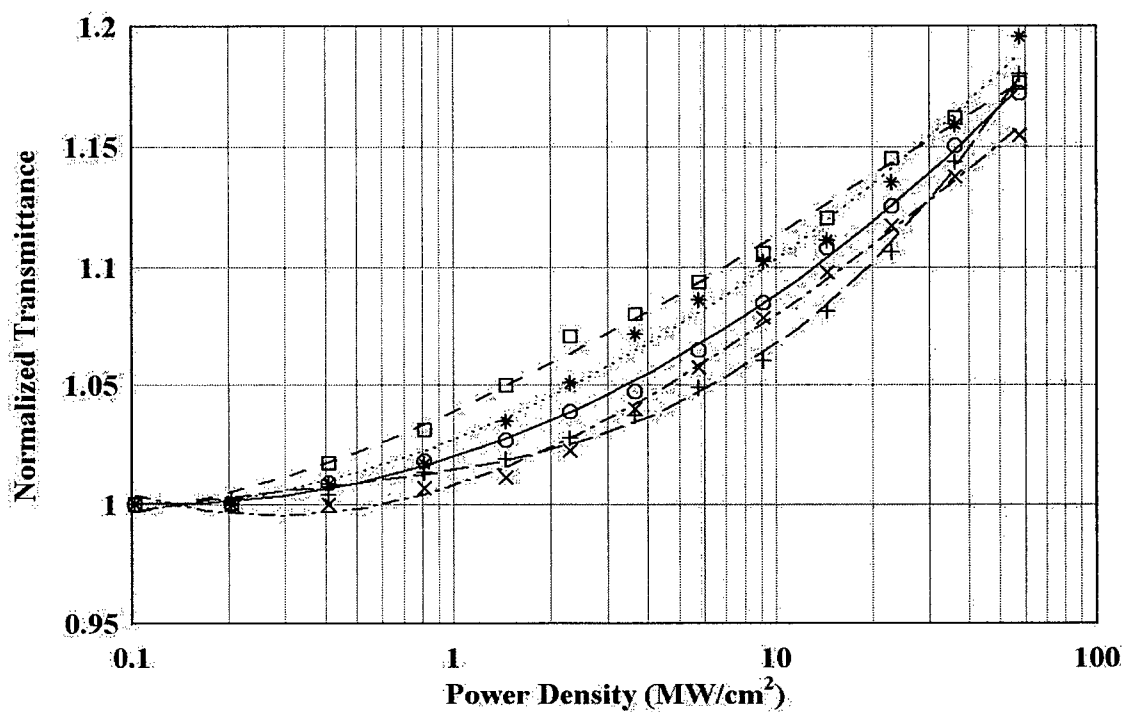
Figure 4.3b: Normalized transmittance curves (7% < ΔT < 10%).
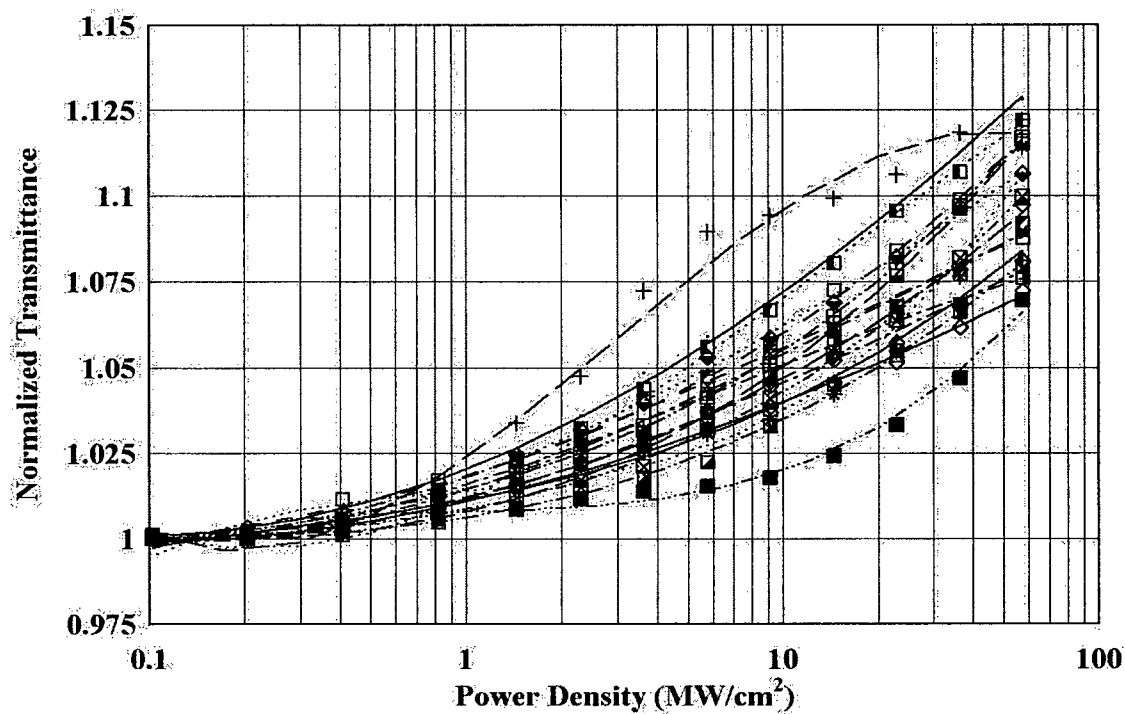
Figure 4.3c: Normalized transmittance curves (ΔT < 7%)
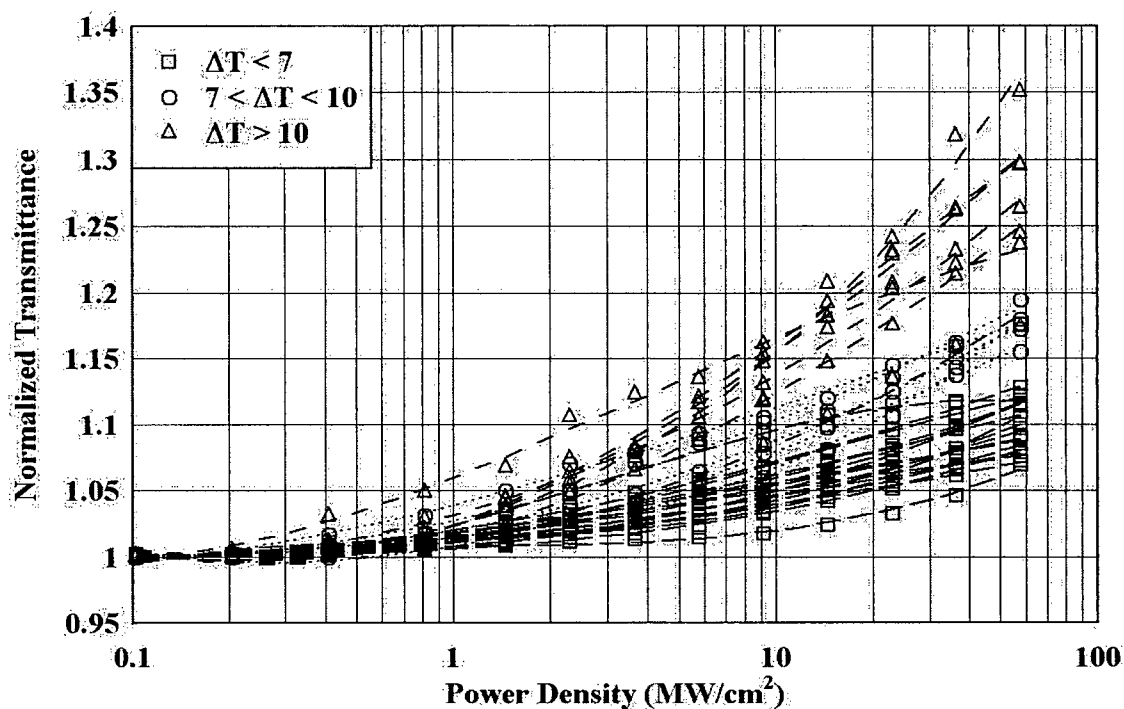
Figure 4.3d: Normalized transmittance curves (all).

4.3.1 Optical Damage

During P-scan measurements, degradation of saturable absorption and increase in WFT in the SWNT sample were observed as more P-scans were performed. The SWNT sample's surface was examined after the measurements, and visible damage to the samples was observed. Figure 4.4 gives a comparison between the damaged area and an area that was not involved in the P-scan. The vertical lines seen in the damaged area were caused by translating the beam waist about the sample surface at high intensity. Sparks observed were indications of damage to the sample. The high intensity burned away layers of SWNTs and reduced the thickness of the sample. This explained the degradation of ΔT (ΔT_{norm}) and enhancement of WFT.

(a)   (b)
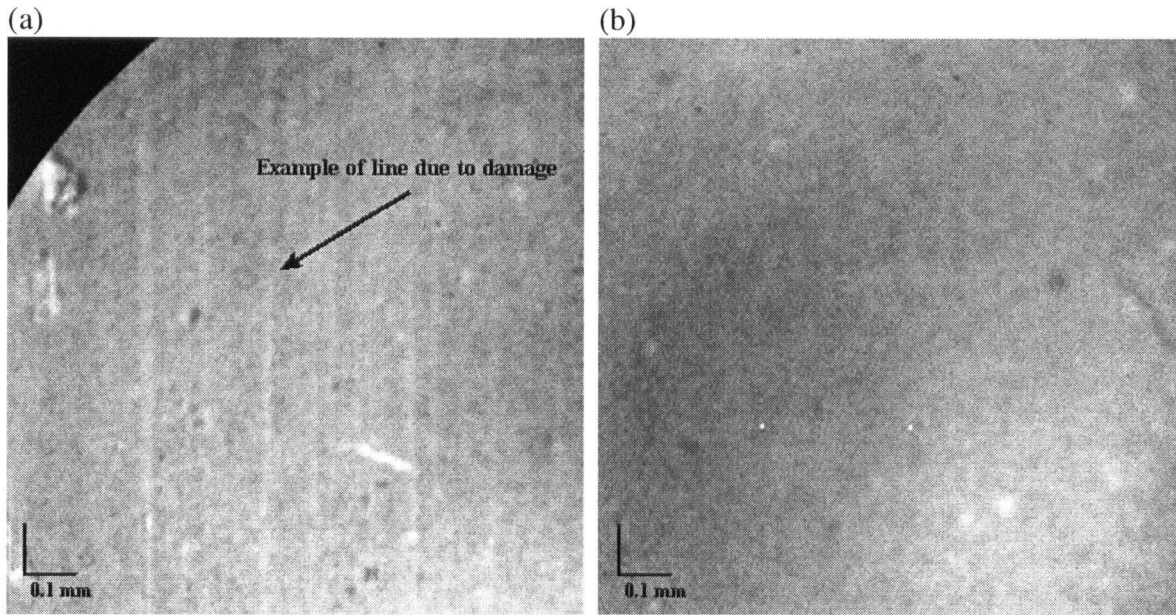
Figure 4.4: (a) Damaged SWNT sample surface compare with (b) SWNT sample surface not involved in P-scan.

Another contributing factor to the degradation of saturable absorption in the SWNT sample was the reaction of SWNTs with nearby reactive oxygen $O_2(^1D_g)$ [35]. Reactive oxygen can be generated from stable oxygen $O_2(^3S_g^-)$ in the surrounding atmosphere by electronic energy transfer. The energy required for this transition is approximately 1 eV/molecule. Since photons at the optical communication wavelength are at about 0.8 eV per photon, it is possible that excitation in the SWNT sample might generate reactive oxygen which would react with the SWNTs and destroy the sample. Tests have been carried out by Schibili et al. [35] and they discovered that in a nitrogen rich atmosphere, saturable absorption in their SWNT sample showed no degradation for prolonged operation and even for extremely high intensity (1 GW/cm^2). A solution to the reactive oxygen problem then is to add a protective coating, such as an anti-reflection coating, on the SWNT sample surface to form a gas barrier. This solution has not been tested but is expected to greatly reduce the probability of damaging the SWNT sample.

4.3.2 Estimation of the Nonlinear Absorption Coefficient

To estimate the nonlinear absorption coefficient of SWNTs, only three T_r curves were considered: the average, the best (maximum ΔT), and the worst (minimum ΔT) curves. The three T_r curves are plotted in figure 4.5. WFT, ΔT, I_{sat} and $\alpha_0 L$ of each curve are listed in table 4.1. It should be noted that since the SWNT sample was never completely saturated, the reported I_{sat} and ΔT in table 4.1 are smaller than the expected values. In order to obtain estimates of β_0, it was assumed that the observed ΔTs in the T_r curves had fully saturated.
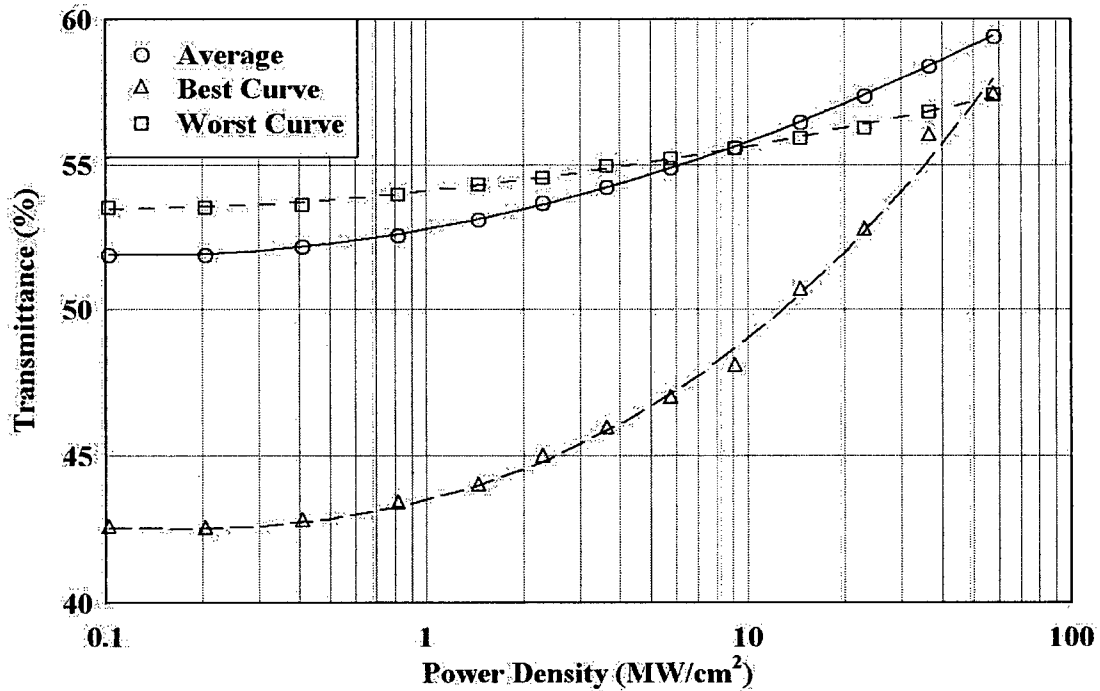
Figure 4.5: Average, best, and worst T_r curves. The symbols represent actual data while the lines represent curves fitted by the graphing tool.

Table 4.1: WFT, ΔT, I_{sat} and $\alpha_0 L$ for average, best, and worst curves

Curve	WFT (%)	ΔT (%)	I_{sat} (MW/cm^2)	$\alpha_0 L$
Average	51.9	7.5	9.1	0.656
Best	42.5	14.9	14	0.856
Worst	53.5	3.9	6	0.625

Combining equations 3.2 and 3.3 (see section 3.2.2) yields

$$T_r = \exp\left(-\left(\alpha_0 + \frac{\beta_0 I}{1 + \frac{I}{I_{sat}}}\right)L\right) \tag{4.1}$$

Equation 4.1 was used along with the data from table 4.1 to estimate the nonlinear absorption coefficient, β_0, associated with each curve. The main difficulty in estimating β_0 was the uncertainties in the values of L and α_0 of each curve due to the sample surface non-uniformity. Only the product terms of L and α_0 associated with each curve were known. In the case of the average curve, L was assumed to be the average L across the sample surface, which was measured by Atomic Force Microscopy in tapping mode to be 600 nm. Using this value for L, α_0 for the average curve was calculated to be 1.09 x 10^6 m^{-1}. For the best and the worst curves, either L or α_0 had to be set to a constant in order to calculate β_0. It is unlikely that L is constant since ΔT is proportional to L and that the two curves clearly have different ΔT; therefore, it was assumed that the α_0 of the two curves are equal to that of the average curve. Based on this assumption, Ls for the best and the worst curves were calculated to be 783 and 572 nm, respectively. Figure 4.6 shows the fitting curves calculated by equation 4.1. The estimated values for β_0 are listed in table 4.2.
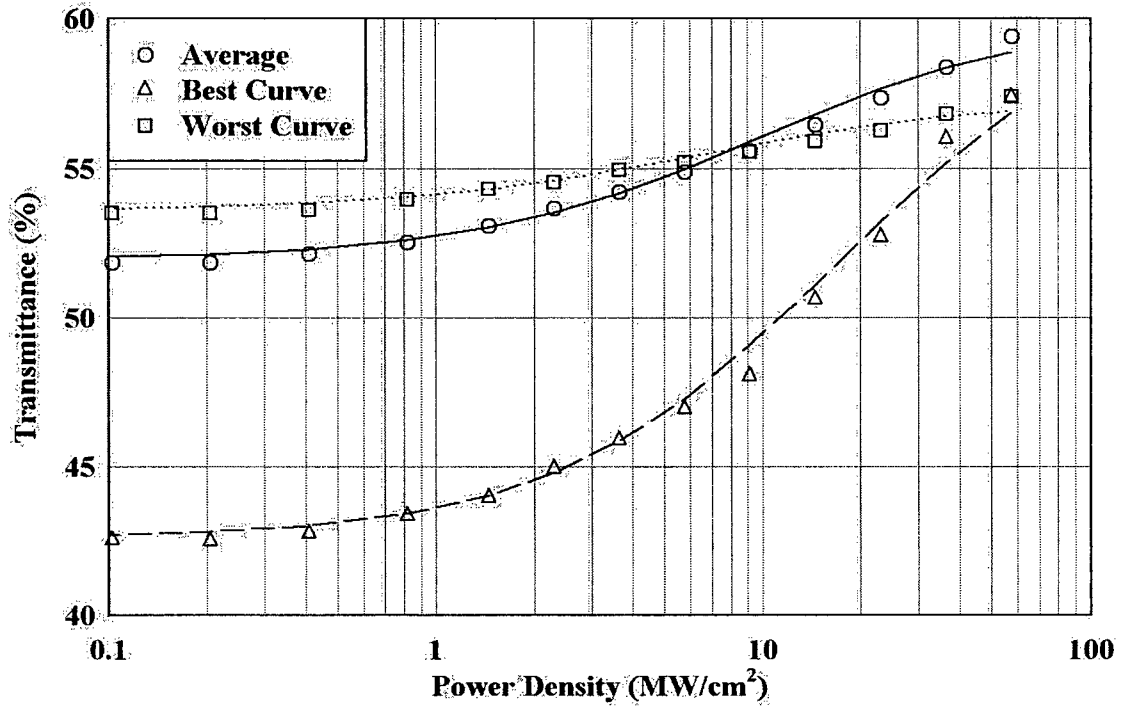
Figure 4.6: Average, best, and worst T_r curves. The symbols represent actual data while the lines represent curves fitted by using equation 4.1 along with data listed in table 4.1.

Table 4.2: β_0 for the average, the best, and the worst curves

Curve	β_0 (m/W)	$\left|\chi_{img}^{(3)}\right|$ (10^{-6} esu)
Average	-2.92×10^{-6}	5.16
Best	-3.32×10^{-6}	5.87
Worst	-2.08×10^{-6}	3.68

In table 4.2, the values of β_0 are also expressed in terms of $\chi_{img}^{(3)}$, which is the imaginary part of the third order nonlinear susceptibility (see section 2.4). For SWNTs, the conversion from β_0 to $\chi_{img}^{(3)}$ is given by [36]

$$\chi_{img}^{(3)} = \frac{\beta_0 n_0 c\lambda}{64\pi^2} \tag{4.2}$$

where n_0, the refractive index of SWNTs, is approximately 2.4 [36]. While the obtained values for β_0 listed in table 4.2 are rough approximations, they provide estimates for the upper and lower bounds of β_0. The estimated values of β_0 are large compared to those of other nonlinear optical materials (ex, β_0 of 2.4 x 10^{-10} m/W was reported for GaAs at 1550 nm [37]). In terms of $\left|\chi^{(3)}_{img}\right|$, they are either within the same order of magnitude or one order of magnitude higher than measured in a recent study (3.0 x 10^{-6} esu and 1.8 x 10^{-7} esu [38] at 1800 and 1400 nm, respectively).

4.4 Pump-Probe Experimental Results

The recovery time of typical SWNT samples were estimated and/or measured in recent studies to be in the sub-picosecond range [9, 38-41]. It was hoped that the implemented measurement system would be able to measure the recovery time of the SWNT sample. However, the width of the pulse generated by the measurement system was too wide to give time resolution in the sub-picosecond range. Measurement of sub-picosecond phenomena requires pulse width of a few hundreds femtoseconds, which was not available in the lab. In addition, the fastest photodetector (45 GHz or ~8.5 ps) available in the lab was also not fast enough to capture sub-picosecond phenomena. While the measurement system was unable to measure the recovery time of SWNTs, optical bleaching was observed in the pump-probe experiment.

4.4.1 Optical Bleaching

In the pump-probe measurement, the power of the probe beam at three different pump intensities (7.37, 3.69, and 1.47 MW/cm^2) was monitored. Figure 4.7 shows the probe

beam's normalized transmittance (T_{p-norm}) at different pump intensities. Prior to the arrival of the pump beam (pulse), T_{p-norm} is a constant. As the pump beam arrived, T_{p-norm} rises. At a pump intensity of 7.37 MW/cm^2, T_{p-norm} rises to a peak of 1.015. The shape of the probe beam is similar to the shape of the transmitted pump beam, which indicates that the recovery time of the SWNT sample is too fast for the photodetector to capture. As the pump intensity decreases, the peak of T_{p-norm} also decreases. T_{p-norm} has a peak of 1.009 and 1.004 for pump intensities of 3.69 and 1.47 MW/cm^2, respectively. Matching these values to the T_{norm} curve (see figure 4.3c) indicates that they are associated with one of the lower response curves. It is estimated that if the pump-probe experiments were performed at a location associated with the high response curve, then at 57 MW/cm^2, T_{p-norm} could reach up to 1.35. This behavior demonstrated by the pump-probe experiment is an indication of optical bleaching in the SWNT sample.

An interesting observation in T_{p-norm} is that at a delay of approximately 100 ps after the probe beam's main pulse, a dip in T_{p-norm} is observed (see figure 4.7). The degree of decrease in T_{p-norm} is proportional to the peak pump intensity. This reduction in transmittance is also observed in the pump beam. Figure 4.8 shows the decrease in transmittance of the pump beam after the second peak of the pulse. The behavior observed in figures 4.7 and 4.8 may be attributed to induced absorption by the lowest excited state [40]. However, in reference [41], it was shown that the magnitude of induced absorption is inversely proportional to the peak pump intensity, which is in contrary to the behavior observed in figure 4.7. Further investigation is needed to determine the cause of this behavior. However, such work would require a pulse with a perfectly Gaussian shape, which is not the case for the pulse used in this work.
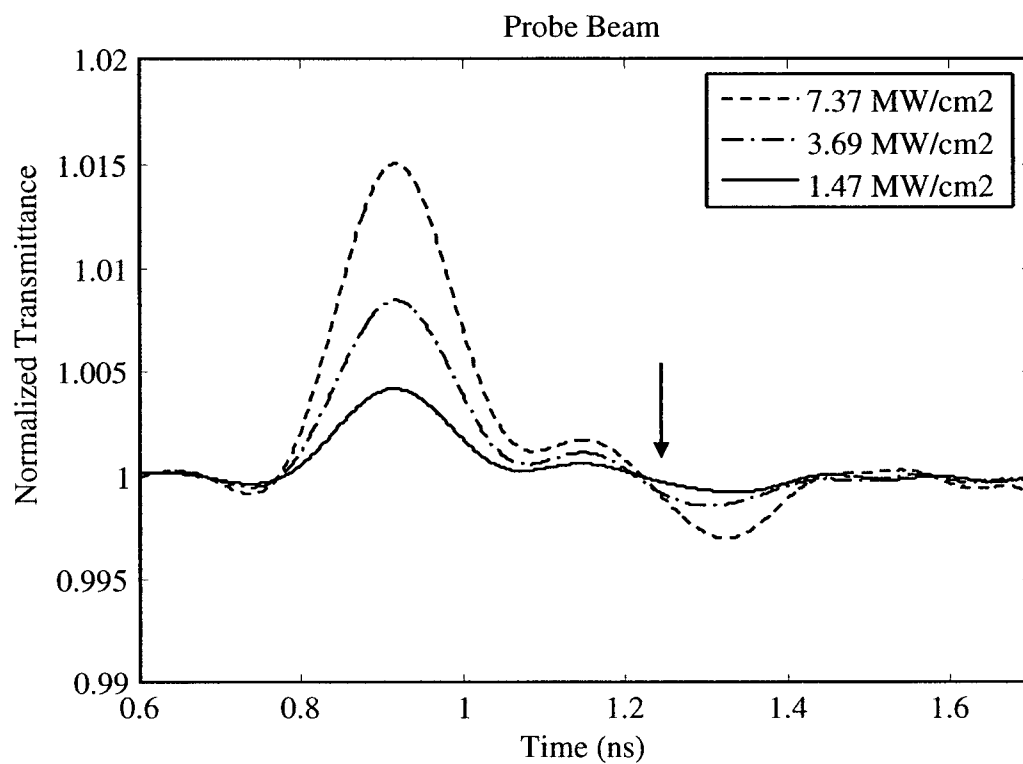
Figure 4.7: Normalized transmittance of probe beam for different pump intensities. Lower pump intensity yields lower peak. A dip in normalized transmittance occurs after the second peaks.
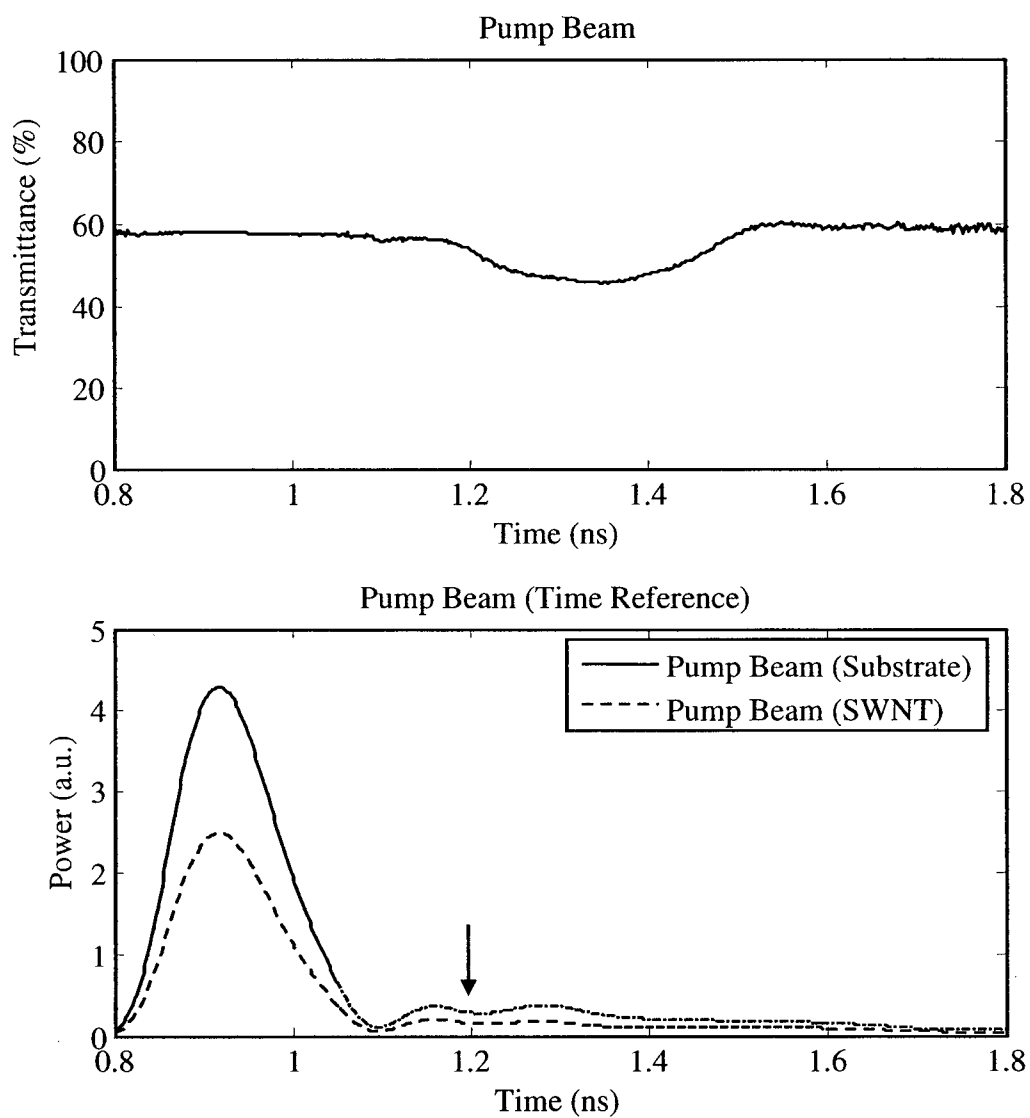
Figure 4.8: Decrease in transmittance after the second peak (as indicated by an arrow). The transmittance is calculated by dividing the power of the pump beam passing through the SWNT sample by that of passing through the substrate (sapphire window).

4.4.2 Carrier Dynamics of Single-Walled Carbon Nanotubes

The observed resemblance of the probe beam to the pump beam is evidence of the ultrafast sub-picosecond recovery time of SWNTs. This ultrafast recovery time is typical in bundled SWNTs, where tube-tube interaction allows electrons excited into the first conduction band to tunnel either into the metallic SWNTs or semiconducting SWNTs with smaller band gaps, which result in the sub-picosecond recovery time.

Recovery times as fast as 100 fs has also been reported in pump-probe measurements involving excitation of the second subband (second band gap) [40]. It is determined that intraband relaxation [17] is responsible for this ultrafast recovery time. Excitation corresponding to the second band gap generates an electron-hole pair in the second conduction and valence band. This second subband excitation is followed by intraband relaxation of the electron-hole pair to the first band gap, resulting in ultrafast recovery of the second subband. This is illustrated in figure 4.9.

To utilize the ultrafast recovery time of intraband relaxation at optical communication wavelengths, the second band gap of SWNTs should be at around 0.8 eV, which correspond to SWNTs having tube diameters of ~ 2 nm (refer to the Kataura plot shown in figure 2.7).
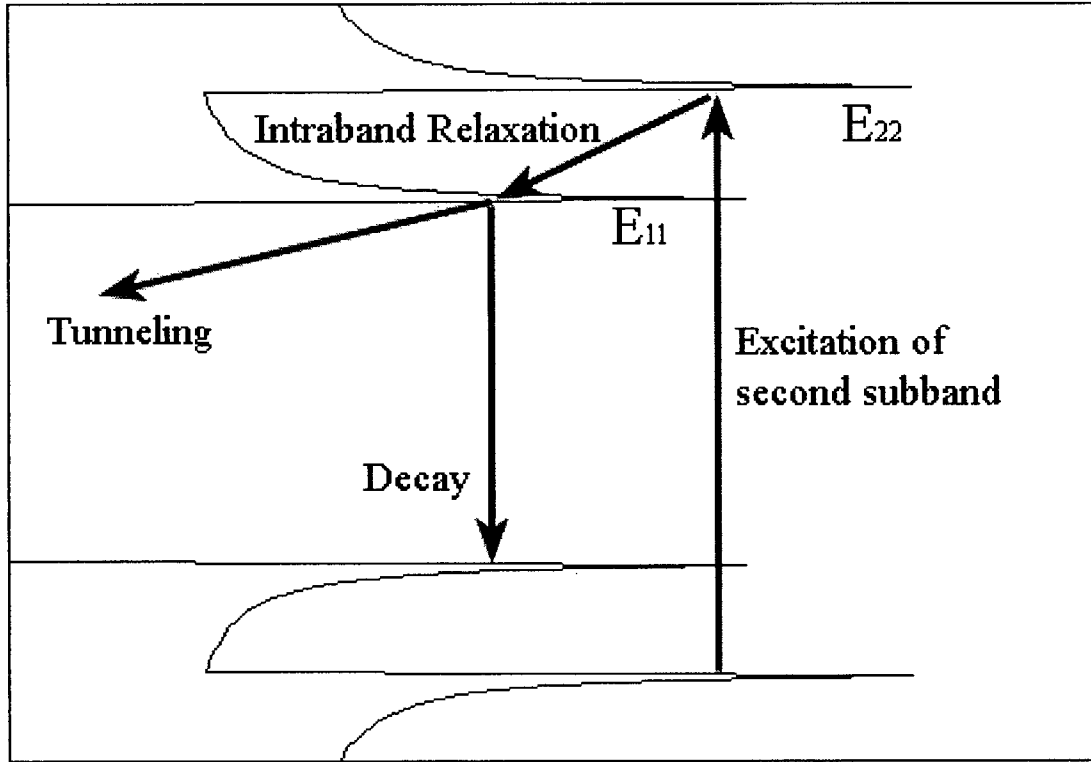
Figure 4.9: The arrows indicate the carrier dynamics in SWNT. Electron excited into the second band gap rapidly relaxes to the first band gap. The result is the ultrafast recovery of the second subband. From the first band gap the excited electron either tunnel into other SWNT or decay back to the first valence band. In the case of isolated SWNT the electron will not tunnel to other SWNT since they are not bundled.

4.5 Performance Evaluation: Figure of Merit

In section 4.3.2, the upper and lower bounds of β_0 of SWNTs were estimated. The magnitude of the estimated values are large compared to other materials exhibiting absorptive nonlinearity [37, 38, 42-44]. In this section, the performance of the SWNT sample as a saturable absorber is evaluated. A figure of merit is proposed to compare the performance of saturable absorbers based on their modulation depth, saturation intensity, and recovery time.

One approach to evaluate the performance of saturable absorbers is to evaluate them based on the maximum induced change in absorption coefficient ($\Delta\alpha_{max}/\alpha_0$); for a perfect saturable absorber, $|\Delta\alpha_{max}/\alpha_0|$ should be equal to one (i.e., the saturable absorber is transparent when saturated). $|\Delta\alpha_{max}/\alpha_0|$ is of particular interest from a device's point of view, since a higher $|\Delta\alpha_{max}/\alpha_0|$ means a larger ΔT and a larger signal-to-noise ratio. Based on this strategy, a figure of merit for evaluating the performance of a saturable absorber is introduced as

$$FOM_{SA} = \frac{\Delta T}{WFT \times I_{sat} \times \tau}$$

(4.3)

Division of ΔT by the weak field transmittance (WFT) is similar to $|\Delta\alpha_{max}/\alpha_0|$ and incorporating I_{sat} into the figure of merit gives insight into the sensitivity of absorptive nonlinearity in the materials. Lower I_{sat} also permits low power operation in applications such as the generation of mode-locked laser pulses [45]. Faster recovery times enable higher bit rates, thus the recovery time τ is incorporated into the FOM_{SA} to give a higher value for shorter τ. For a good saturable absorber, FOM_{SA} should be as large as possible. Table 4.4 lists the calculated $FOM_{SA}s$ of several MQW based saturable absorbers and several SWNT samples. In table 4.4, L-SWNTs and H-SWNTs represent SWNTs grown by laser ablation and the HiPco process, respectively. Note that the values of ΔT listed in table 4.4 for the average, the best, and the worst curves are obtained by interpolation using equation 4.1.

Table 4.4: FOM$_{SA}$ of various saturable absorbers

Material	λ (nm)	ΔT (%)	WFT (%)	I_{sat} (Wcm^{-2})	τ (ps)	FOM (cm^{2}W^{-1}s^{-1})
InGaAs-AlAs-AlAsSb [13]	1620	23	77	22.7 M	0.6	21.93 k
InGaAsP/InP [15]	1570	4.7	11	5.7 M	5	14.99 k
InGaAs-AlAsSb [14]	1880	54	46	48.5 M	2	12.10 k
L-SWNTs [8]	1550	3.5	~65	1.45 M	~1	37.14 k
H-SWNTs [12]	1540	2.9	78.5	2.5 M	~1	14.78 k
H-SWNTs [9]	1550	8	30	38 M	0.51	13.76 k
H-SWNTs Best (measured)	1550	18.3	42.5	14 M	~1	30.76 k
H-SWNTs Worst (measured)	1550	3.9	53.5	6 M	~1	12.15 k
H-SWNT Average (measured)	1550	8	51.9	9.1 M	~1	16.94 k

Table 4.4 shows that in general, the performance of SWNT samples as saturable absorbers are comparable to those of MQW based saturable absorbers at optical communication wavelengths (OCW). The strength of SWNT based saturable absorbers comes from their ultrafast recovery time. The fastest reported recovery time of a MQW based saturable absorber at 1550 nm is 635 fs [46], which is slower than the 510 fs recovery time of SWNTs reported in reference [9] (a recovery time of 0.35 ps measured at 1400 nm was reported in reference [38]). In addition, τ for SWNTs can be further reduced by applying the ultrafast recovery time of intraband relaxation. It should also be noted that none of the SWNT samples listed in table 4.4 are optimized to operate at OCW (i.e., the excitation is off-resonant). Improvements in performance are expected for SWNT samples optimized to operate at OCW (i.e., SWNT samples containing SWNTs with tube diameter of ~1 nm).

Another important consideration when comparing saturable absorbers is their device fabrication processes. For MQW based saturable absorbers, the fabrication

processes involve the use of systems such as metalorganic chemical vapor deposition (MOCVD) and molecular beam epitaxy (MBE). MOCVD and MBE are clean-room-based systems and are very costly. In comparison, excluding the processes in growing SWNTs, most SWNT based device fabrication processes [8-11, 38], such as the filtration method used in this work, are much simpler and less expensive than MOCVD and MBE.

The optical properties and ultrafast recovery time of SWNTs, combined with the low fabrication cost, make SWNTs the ideal material for producing nonlinear optical device.

4.6 Summary

The upper and lower bounds of the nonlinear absorption coefficient of SWNTs has been estimated. The estimated values are large compared to those of other nonlinear optical materials and are within in the same order of magnitude as those reported in a recent study. Optical bleaching in the SWNT sample was observed in the pump-probe measurement. Although the measurement system was unable to measure the recovery time of the SWNT sample, the resemblance of the probe beam signal to the transmitted pump beam pulse indicates that the SWNT sample possesses a recovery time that is much faster than the speed of the photodetector (~8 ps). Using the proposed figure of merit, it has been shown that, at optical communication wavelengths, the performance of the SWNT sample and several SWNT based saturable absorbers reported in the literature are comparable to those of MQW based saturable absorbers reported in the literatures.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

5.1 Conclusions

In this work, saturable absorption in SWNTs at the optical communication wavelength was studied. A polarization insensitive SWNT sample was fabricated using the filtration method. A measurement system capable of performing Z-scan, P-scan, and pump-probe measurement was developed and implemented. The P-scan and the pump-probe measurements were performed to observe and to study saturable absorption and optical bleaching in the SWNT sample fabricated. Based on a newly proposed figure of merit, the performance of SWNT based saturable absorbers has been evaluated and compared with other multi-quantum well (MQW) based saturable absorbers. The obtained results will be summarized in this section.

A modification of the open aperture Z-scan method, referred in this work as the P-scan method, was used to observe saturable absorption in SWNTs. Data obtained from P-scan measurements were analyzed to estimate the nonlinear absorption coefficient of SWNTs. The estimated values are several orders of magnitude larger than those of other nonlinear optical materials at the optical communication wavelength.

The results obtained from the pump-probe measurements show that the fabricated SWNT sample possesses a recovery time that is much faster than the speed of the photodetector used in the measurement system. It is estimated that the recovery time of

the SWNT sample is similar to the sub-picosecond recovery times reported in recent studies.

Using the proposed figure of merit, the performance of a SWNT sample fabricated for this work and SWNT based saturable absorbers reported in the literature [8, 9, 12] are compared with those of MQW based saturable absorbers reported in the literature [13-15]. The results show that the performance of SWNT based saturable absorbers are comparable to those of multi-quantum well based saturable absorbers. The advantages of SWNT based devices are their ultrafast recovery times, high tunability, and low cost of fabrication.

In conclusion, in this work it has been shown that SWNTs possess a large degree of absorptive nonlinearity. This useful optical property, combined with the low fabrication cost, illustrate the potential of SWNTs to find a large number of high speed optical applications. It is expected that in the future, as SWNTs of selectable tube diameters become readily available, SWNTs will greatly impact the design of high speed optical systems, especially at optical communication wavelengths.

5.2 Recommendations for Future Work

In future work, the device fabrication technique should be improved to ensure the fabrication of highly uniform (or high density) SWNT samples. The filtration method is a manual process, hence good quality control is difficult to achieve. In this work, the fabricated sample's surface non-uniformity and the difference between the actual (~600 nm) and the designed (~1 μm) sample thickness are evidence of a need for improved quality control.

The incorporation of SWNT based saturable absorbers in generating mode-locked laser pulses has been demonstrated by several groups [10, 11, 35]. The SWNT sample fabricated in this work can also be used to generate laser pulses in a fiber ring laser setup. An experimental setup of a mode-locked fiber ring laser incorporating SWNT based saturable absorber is shown in figure 5.1.
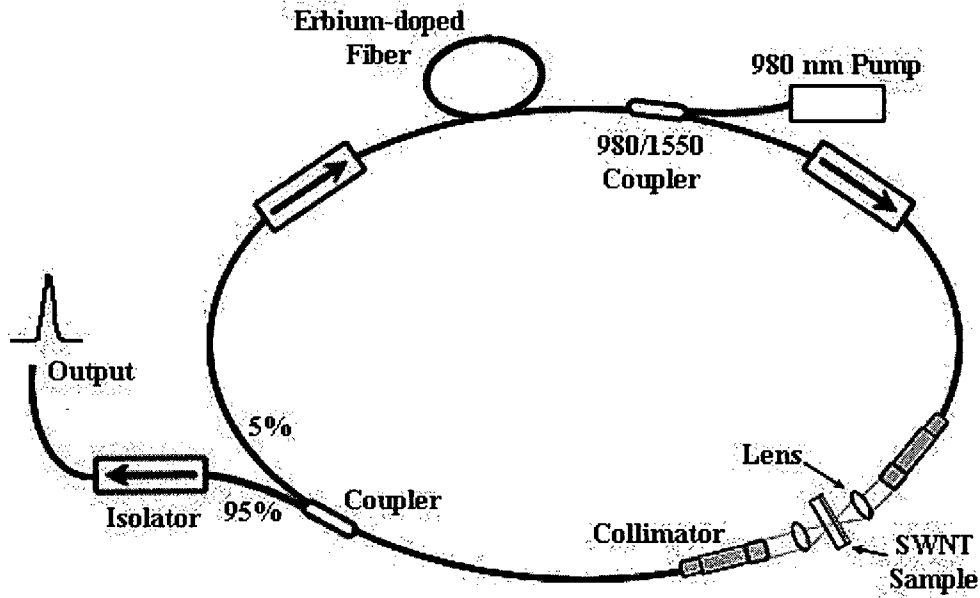
Figure 5.1: A mode-locked fiber ring laser incorporating SWNT based saturable absorber.

The gain medium of the fiber ring laser is an erbium doped fiber backward pumped by a 980 nm adjustable pump source. Erbium doped fiber is chosen for its stimulated emission at optical communication wavelengths. Isolators are installed at both ends of the gain medium to ensure unidirectional operation. Two collimator focusing lens pairs are used to focus laser light onto the SWNT saturable absorber and to collect the transmitted light back into the laser cavity. The laser light is then separated into two

paths: one is used as a feedback for the laser cavity while the other is used as an output of the fiber ring laser. If the pump source could provide sufficient power so that the fundamental mode of the gain medium can breach the saturable absorber, then the fiber ring laser will be able to generate mode-locked laser pulses.

Another potential optical application of interest for SWNT based devices is high speed optical modulator. Change in the absorption spectra of SWNT thin films by both p-type (I_2, Br_2) and n-type (K, Cs) doping has been demonstrated by Kazaoui et al. [47]. Their results, combined with the operation of SWNTs as field effect transistors [48, 49], indicate the possibility in modulating the optical absorption peaks in SWNTs by field gating. This concept has been demonstrated by Wu et al. [50] at a wavelength of 1676 nm. Their SWNT based optical modulator has a modulation depth of up to 48% between gate voltages of ±1.8 V. The same behavior is expected at the optical communication wavelength. While the dynamic behavior of optical modulation in SWNTs has not yet been studied, it is expected that SWNTs would find applications such as optical modulators and sensors due to their useful optical and electrical properties.

REFERENCES

[1]   G. Steinmeyer, "A review of ultrafast optics and optoelectronics," *Journal of Opitcs A: Pure and Applied Optics*, vol. 5, pp. R1-R15, 2003.

[2]   S. Suomalainen, A. Vainionpaa, O.Tengvall, T. Hakulinen, S. Karirinne, T. G. Euser, and W. L. Vos, "Long-wavelength fast semiconductor saturable absorber mirros using metamorphic growth on GaAs substrates," *Applied Physics Letters*, vol. 87, pp. 121106, 2005.

[3]   F. Seguineau, D. Rouvillain, H. Choumane, G. Aubin, J. L. Oudar, P. Brindel, B. Lavigne, and O. Leclerc, "Regeneration capabilities of passive saturable absorber-based optical 2R in 20 Gbit/s RZ DWDM long-haul transmissions," *Electronics Letters*, vol. 39, pp. 857-859, 2003.

[4]   M. Pantouvaki, M. J. Fice, R. Feced, E. P. Burr, R. Gwilliam, A. B. Krysa, J. S. Roberts, and A. J. Seeds, "10-Gb/s All-Optical 2R Regeneration Using an MQW Fabry-Perot Saturable Absorber and a Nonlinear Fiber," *IEEE Photonics Technology Letters*, vol. 16, pp. 617-619, 2004.

[5]   O. Leclerc, B. Lavigne, E. Balmefrezol, P. Brindel, L. Pierre, D. Rouvillain, and F. Seguineau, "Optical Regeneration at 40 Gb/s and Beyond," *Journal of Lightwave Technology*, vol. 21 pp. 2779-2790, 2003.

[6]   S. Iijima, "Helical microtubules of graphitic carbon," *Letters to Nature*, vol. 354, pp. 56-58, 1991.

[7]   M. S. Dresselhaus, G. Dresselhaus, and P. Avouris, *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*. Berlin: Springer-Verlag, 2001.

[8]   S. Y. Set, H. Yaguchi, M. Jablonski, Y. Tanaka, Y. Sakaibara, A. Rozhin, M. Tokumoto, H. Kataura, Y. Achiba, and K. Kikuchi, "A Noise Suppressing Saturable Absorber at 1550nm Based on Carbon Nanotube Technology," presented at Optical Fiber Communications Conference, 2003.

[9]   Y. C. Chen, N. R. Raravikar, L. S. Schadler, P. M. Ajayan, Y. P. Zhao, T. M. Lu, G. C. Wang, and X. C. Zhang, "Ultrafast optical switching properities of single-

wall carbon nanotube polymer composites at 1.55 ʋm," *Applied Physics Letters*, vol. 81, pp. 975-977, 2002.

[10]   Y. Inoue, S. Yamashita, S. Maruyama, Y. Murakami, H. Yaguchi, T. Kotake, and S. Y. Set, "Mode-locked fiber lasers using vertically aligned carbon nanotubes directly synthesized onto substrates," presented at Optical Fiber Communications Conference, 2003.

[11]   S. Yamashita, Y. Inoue, S. Maruyama, Y. Murakami, H. Yaguchi, M. Jablonski, and S. Y. Set, "Satruable absorber incorporating carbon nanotubes directly synthesized onto substrates and fibers and their application to mode-locked fiber lasers," *Optics Letters*, vol. 29, pp. 1581-1583, 2004.

[12]   N. N. Il'ichev, E. D. Obraztsova, S. V. Garnov, and S. E. Mosaleva, "Nonlinear transmission of single-wall carbon nanotubes in heavy water at a wavelength of 1.54 um and self-mode locking in a Er^{3+}: glass laser obtained using a passive nanotube switch," *Quantum Electronics*, vol. 34, pp. 572-574, 2004.

[13]   A. V. Gopal, T. Simoyama, H. Yoshida, J. I. Kasai, T. Mozume, and H. Ishikawa, "Intersubband Absorption Saturation in InGaAs-AlAs-AlAsSb Coupled Quantum Wells," *IEEE Journal of Quantum Electronics*, vol. 39, pp. 1356-1361, 2003.

[14]   A. V. Gopal, H. Yoshida, A. Neogi, N. Georgiev, T. Mozume, T. Simoyama, O. Wada, and H. Ishikawa, "Intersubband Absorption Saturation in InGaAs-AlAsSb Quantum Wells," *IEEE Journal of Quantum Electronics*, vol. 38, pp. 1151-1520, 2002.

[15]   M. Pantouvaki, R. Gwilliam, E. P. Burr, A. B. krysa, J. S. Robers, and A. J. Seeds, "Ultrafast Recovery Times and Increased Absorption Nonlinearity in InGaAsP MQW Saturable Absorbers Implanted at 200°C," in *International Conference on Indium Phosphide and Related Materials*, 2003.

[16]   R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes*. United Kingdom: Imperial College Press, 1998.

[17]   S. Reich, C. Thomsen, and J. Maultzsch, *Carbon Nanotubes Basic Concepts and Physical Properties*. Federal Republic of Germany: Wiley-VCH, 2004.

79

[18]   H. Kataura, Y. Kumazawa, Y. Maniwa, I. Umezu, S. Suzuki, Y. Ohtasuka, and Y. Achiba, "Optical Properties of Single-Wall Carbon Nanotubes," *Synthetic Metals*, vol. 103, pp. 2555-2558, 1999.

[19]   L. X. Benedict, S. G. Louie, and M. L. Cohen, "Static polarizabilities of single-wall carbon nanotubes," *Physical Review B*, vol. 52, pp. 8541-8549, 1995.

[20]   V. A. Margulis, "Theoretical estimations of third-order optical nonlinearities for semiconductor carbon nanotubes," *Journal of Physics: Condensed Matter*, vol. 11, pp. 3065-3074, 1999.

[21]   V. A. Margulis, E. A. Gaiduk, and E. N. Zhidkin, "Electric-field-induced optical second-harmonic generation and nonlinear optical rectification in semiconducting carbon nanotubes," *Optics Communications*, vol. 183, pp. 317-326, 2000.

[22]   M. Sheik-Bahae, A. A. Said, T. H. Wei, D. J. Hagan, and E. W. V. Stryland, "Sensitive Measurement of Optical Nonlinearities Using a Single Beam," *Journal of Quantum Electronics*, vol. 26, pp. 760-769, 1990.

[23]   W. T. Silfvast, *Laser Fundamentals*. United States of America: Cambridge University Press, 1996.

[24]   J. F. Lami and C. Hirlimann, "Two-photon excited room-temperature luminescence of CdS in the femtosecond regime," *Physical Review B*, vol. 60, pp. 4763-4770, 1999.

[25]   H. Sun, "Thin lens equation for a real laser beam with weak lens aperture truncation," *Optical Engineering*, vol. 37, pp. 2906-2913, 1998.

[26]   New Focus, "Practical Uses and Applications of Electro-Optic Modulator," in *Application Note*, 2001.

[27]   M. J. Bronikowski, P. A. Willis, D. T. Colbert, K. A. Smith, and R. E. Smalley, "Gas-phase production of carbon single-walled nanotubes from carbon monoxide via the HiPco process: A parametric study," *Journal of Vacuum Science and Technology A*, vol. 19, pp. 1800-1805, 2001.

[28]   P. Nikolaev, M. J. Bronikowski, R. K. Bradley, F. Rohmund, D. T. Colbert, K. A. Smith, and R. E. Smalley, "Gas-phase catalytic growth of single-walled carbon

nanotubes from carbon monoxide," *Chemical Physics Letters*, vol. 313, pp. 91-97, 1999.

[29] E. Gregan, S. M. Keogh, T. G. Hedderman, B. McCarthy, G. Farrell, G. Chambers, and H. J. Byrne, "Stokes/anti-Stokes Raman Spectroscopy of HiPco Single-Wall Carbon Nanotubes," *AIP Conference Proceedings*, vol. 663, pp. 294-297, 2002.

[30] A. Hagen, G. Moos, V. Talalaev, and T. Hertel, "Electronic structure and dynamics of optically excited single-wall carbon nanotubes," *Applied Physics A: Materials Science & Processing*, vol. 78, pp. 1137-1145, 2004.

[31] S. M. Bachilo, M. S. Strano, C. Kittrell, R. H. hauge, R. E. Smalley, and R. B. Weisman, "Structure-Assigned Optical Spectra of Single-Walled Carbon Nanotubes," *Science*, vol. 298, pp. 2361-2366, 2002.

[32] A. Kukovecz, C. Kramberger, V. Georgakilas, M. Prato, and H. Kuzmany, "A detailed Raman study on thin single-wall carbon nanotubes prepared by the HiPco process," *The European Physical Journal B*, vol. 28, pp. 223-230, 2002.

[33] A. G. Rinzler, J. Liu, H. Dai, P. Nikolaev, C. B. Huffman, F. J. Rodriguez-Macias, P. J. Boul, A. H. Lu, D. Heymann, D. T. Colbert, R. S. Lee, J. E. Fischer, A. M. Rao, P. C. Eklund, and R. E. Smalley, "Large-scale purification of single-wall carbon nanotubes: process, product, and characterization," *Applied Physics A: Materials Science & Processing*, vol. 67, pp. 29-37, 1998.

[34] J. Li, A. Cassell, L. Delzeit, J. Han, and M. Meyyappan, "Novel Three-Dimensional Electrodes: Electrochemical Properties of Carbon Nanotube Ensembles," *Journal of Physics Chemistry B*, vol. 106, pp. 9299-9305, 2002.

[35] T. R. Schibli, K. Minoshima, H. Kataura, E. Itoga, N. Minami, S. Kazaoui, K. Miyashita, M. Tokumoto, and Y. Sakakibara, "Ultrashort pulse-generation by saturable absorber mirrors based on polymer-embedded carbon nanoubes," *Optics Express*, vol. 13, pp. 8025-8031, 2005.

[36] V. A. Margulis and T. Sizikova, "Theoretical Study of Third-Order Nonlinear Optical Response of Semiconductor Carbon Nanotubes," *Physica B*, vol. 245, pp. 173-189, 1998.

[37] Z. Zheng, A. M. Weiner, J. H. Marsh, and M. M. Karkhanehchi, "Ultrafast Optical Thresholding based on Two-Photon Absorption GaAs Waveguide Photodectors," *IEEE Photonics Technology Letters*, vol. 9, pp. 493-495, 1997.

[38] A. Maeda, S. Matsumoto, H. Kishida, T. Takenobu, Y. Iwassa, M. Shiraishi, M. Ata, and H. Okamoto, "Nonlinear and ultrafast optical response in single-walled carbon nanotubes," *Journal of Physics: Conference Series*, vol. 21, pp. 177-182, 2005.

[39] L. Huang, H. N. Pedrosa, and T. D. Krauss, "Ultrafast Ground-State Recovery of Single-Walled Carbon Nanotubes," *Physical Review Letters*, vol. 93, pp. 17403:1-4, 2004.

[40] H. Hippler, A. N. Unterreiner, J. P. yang, S. Lebedkin, and M. M. Kappes, "Evidence of ultrafast optical switching behaviour in individual single-walled carbon nanotubes," *Physical Chemistry Chemical Physics*, vol. 6, pp. 2387-2390, 2004.

[41] Y. Z. Ma, J. Stenger, J. Zimmermann, S. M. Bachilo, R. E. Smalley, R. B. Weisman, and G. R. Fleming, "Ultrafast carrier dynamics in single-walled carbon nanotubes probed by femtosecond spectroscopy," *Journal of Chemcial Physics*, vol. 120, pp. 3368-3373, 2004.

[42] M. N. Islam, C. E. Soccolich, R. E. Slusher, A. F. J. Levi, W. S. Hobson, and M. G. Young, "Nonlinear spectroscopy near half-gap in bulk and quantum well GaAs/AlGaAs waveguides," *Journal of Applied Physics*, vol. 71, pp. 1927-1935, 1992.

[43] F. Kadlec, H. Nemec, and P. Kuzel, "Optical two-photon absorption in GaAs measured by optical-pump terahertz-probe spectroscopy," *Physical Review B*, vol. 70, pp. 125205:1-6, 2004.

[44] A. Vaidyanathan and A. H. Guenther, "Band-structure calculations of the two-photon absorption coefficients fo GaAs, InP, CdTe, and ZnSe," *Physical Review B*, vol. 24, pp. 2259-2262, 1981.

[45] R. Paschotta, L. Krainer, S. Lecomte, G. J. Spuhler, S. C. Zeller, A. Aschwanden, D. Lorenser, H. J. unold, K. J. Weingarten, and U. Keller, "Piscosecond pulse sources with multi-GHz repetition rates and high output power," *New Journal of Physics*, vol. 6, pp. 174:1-13, 2004.

[46] T. Akiyama, N. Georgiev, T. Mozume, H. yoshida, A. V. Gopal, and O. Wada, "Nonlinearity and recovery time of 1.55 um intersubband absorption in InGaAs/AlAs/AlasSb couppled quantum wells," *Electronics Letters*, vol. 37, pp. 129-130, 2001.

[47] S. Kazaoui, N. Minami, R. Jacquemin, H. Kataura, and Y. Achiba, "Amphoteric doping of single-wall carbon-nanotube thin films as probed by optical absorption spectrosocpy," *Physical Review B*, vol. 60, pp. 13339-13342, 1999.

[48] S. J. Tans, A. R. M. Verschueren, and C. Dekker, "Room-temperature transistor based on a single carbon nanotube," *Letters to Nature* vol. 393, pp. 49-52, 1998.

[49] K. Keren, R. S. Berman, E. buchstab, U. Sivan, and E. Braun, "DNA-Templated Carbon Nanotube Field-Effect Transistor," *Science*, vol. 302, pp. 1380-1382, 2003.

[50] Z. Wu, Z. Chen, X. Du, J. M. Logan, J. Sippel, M. Nikolou, K. Kamaras, J. R. Reynolds, D. B. Tanner, A. F. Hebard, and A. G. Rinzler, "Transparent, Conductive Carbon Nanotube Films," *Science*, vol. 305, pp. 1273-1276, 2004.

[51] J. Alda, "Laser and Gaussian Beam Propagation and Transformation," in *Encyclopedia of Optical Engineering*: Marcel Dekker, Inc., 2003.

[52] S. A. Self, "Focusing of spherical Gaussian beams," *Applied Optics*, vol. 22, pp. 658-661, 1983.

[53] C. C. Davis, *Laser and Electro-Optics Fundamentals and Engineering*. United Kingdom: Cambridge University Press, 1996.

[54] A. E. Siegman, "New developments in laser resonators," *SPIE*, vol. 1124, pp. 1-14, 1990.

Appendix A

Spot Size Estimation

The selection of the appropriate lenses to achieve the desired spot size requires good understanding of the behavior of a laser beam traveling through a system of lenses. Here, an overview of the propagation of a Gaussian beam in a lens system is provided. Derivation of the spot size equation incorporating the beam quality factor is shown.

The propagation of a laser beam can be approximated by a Gaussian beam [51], which corresponds to the TEM_{00} mode. A Gaussian beam has a Gaussian transverse intensity profile shown in figure A.1.
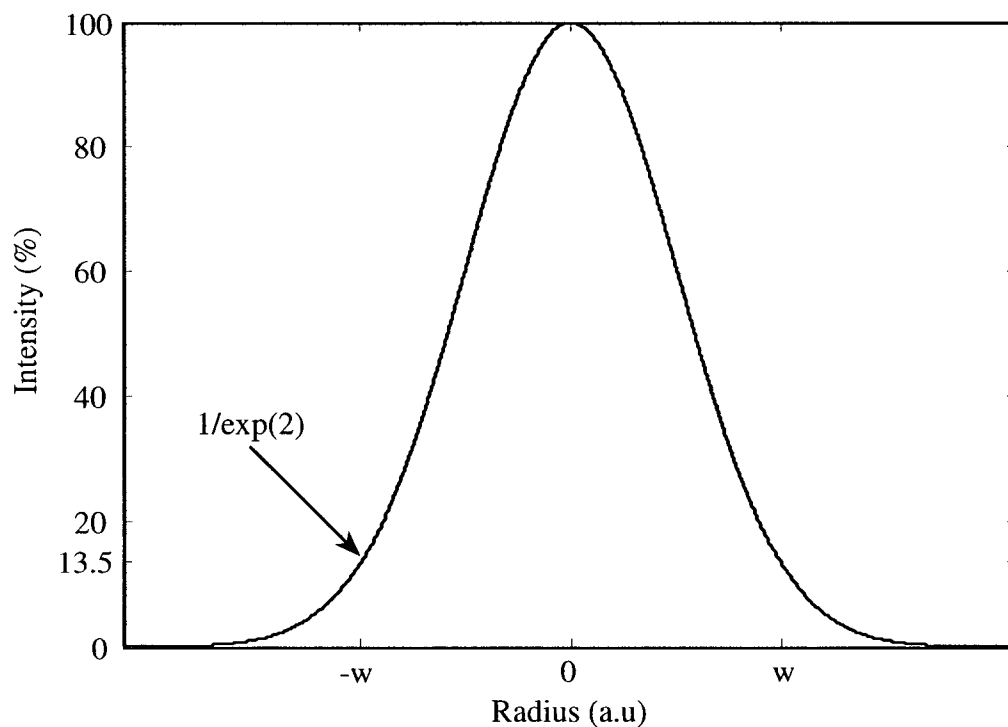

Figure A.1: The intensity profile of a Gaussian beam. The point where the intensity profile drops below 1/exp(2) (13.5%) is referred as the radius of the Gaussian beam.

When a Gaussian beam is emitted from a source, it propagates at the beginning as a perfectly plane wave. As the beam travels away from the source, it gains curvature and spreads out. The spreading of the Gaussian beam is characterized by [51]

$$R(z) = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right] \qquad (A.1)$$

and

$$w(z) = w_0 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]^{\frac{1}{2}} \qquad (A.2)$$

where z is the distance from where the beam originates, λ the wavelength of the beam, w_0 is the radius of the beam at $z = 0$ or the beam waist radius, and $w(z)$ and $R(z)$ are, respectively, the radius and the radius of curvature of the beam after traveling a distance z away from the source. The radius of the Gaussian beam is defined as the radius at which the intensity profile of the beam has fallen to $1/e^2$ (refer to figure A.1). The evolution of a Gaussian beam as it propagates away from the source is illustrated in figure A.2. The location where the beam radius is measured is called the beam waist. At the beam waist, the wave front is flat and $R(z) = \infty$. As the beam propagates in the z direction away from the source, $R(z)$ becomes finite, reaches a minimum, and then approaches ∞ again as z goes to ∞. As z increases, $w(z)$ increases and eventually can be approximated by

$$w(z) = \frac{\lambda z}{\pi w_0} \qquad (A.3)$$

The evolution of the spreading of the beam radius, or diameter, can be approximated by a cone with its vertex located on the z axis at the beam waist. At far-field the spreading of the beam radius makes an angle $\theta = \lambda/\pi w_0$ to the z-axis. This is the far-field angular

85

radius or half angle of divergence of the Gaussian beam. This half angle of divergence θ

along with the nonlinear behavior of R(z) indicate that the beam does not diverge linearly.

At the near-field the beam begins with a small divergence angle which slowly approaches

the far-field angle. The Rayleigh range [51], $z_R = \pi w_0^2/\lambda$, is the point that separates the

far-field from the near-field and is the location on the z-axis where the beam radius has

increased by a factor of square root of 2 compared to w_0. It is also the location where R(z)

has reached a minimum. The separation of the near-field from the far-field is extremely

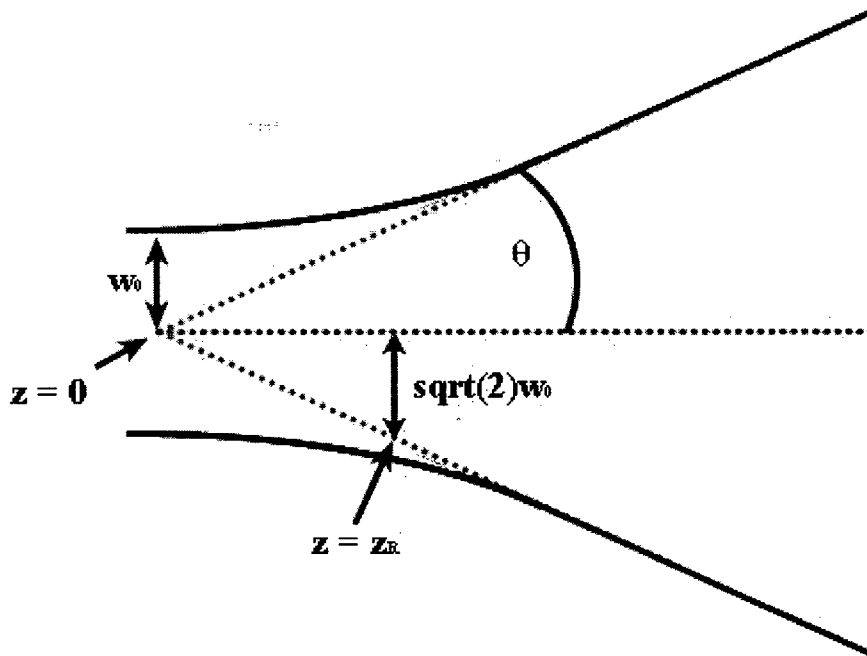important for the determination of the beam quality factor M^2, which will be discussed

later.

Figure A.2: The evolution of Gaussian beam along the z-axis. The beam waist is the location where the spot size is the smallest. The beam spreads out like a cone with its vertex at the beam waist. z_R is the Rayleigh range of the beam.

Self et al. [52] developed a method to model the propagation of Gaussian beam through simple optics using a simple lens equation. The standard lens equation is

$$\frac{1}{s/f} + \frac{1}{s''/f} = 1 \qquad (A.4)$$

where s is the object distance, s" the image distance, and f the focal length of the lens. In Self's model the object represents the waist of the input beam and the image represents the waist of the output beam. See figure A.3 for an illustration. Incorporating the Rayleigh range of the input beam into the model yields the Self's equation [52]

$$\frac{1}{s + z_R^2/(s-f)} + \frac{1}{s''} = \frac{1}{f} \qquad (A.5)$$

For the case where the beam waist of the input beam is at the principle surface of the lens (focusing of a collimated beam), s is equal to zero and rearranging equation A.5 yield s" as

$$s'' = \frac{f}{1 + (\lambda f / \pi w_{0in}^2)^2} \qquad (A.6)$$

where w_{0in} is the input beam's beam waist radius. Subsituiting s" and w_{0in} into z and w(z) respectively in equation A.2 and assuming that $w_{0in} \gg w_{0out}$ (i.e., the collimated beam diameter is large compared to the spot size) and $\lambda f/\pi w_{0in}^2 \ll 1$ (always the case) yield

$$w_{0out} = \frac{\lambda f}{\pi w_{0in}} \qquad (A.7)$$

Transforming equation A.7 to give the beam size parameter in terms of beam diameters yields

$$D_s = \frac{4\lambda f}{\pi D} \qquad (A.8)$$

where D_s is the spot size and D the input (collimated) beam diameter. While the general approach in reducing the spot size is to increase the diameter of the collimated beam and to decrease the focal length of the lens, the spot size cannot be reduced indefinitely. There is a limit on the minimum spot size that can be focused by a lens and it is $D_{smin} \approx \lambda/\pi$ [53].
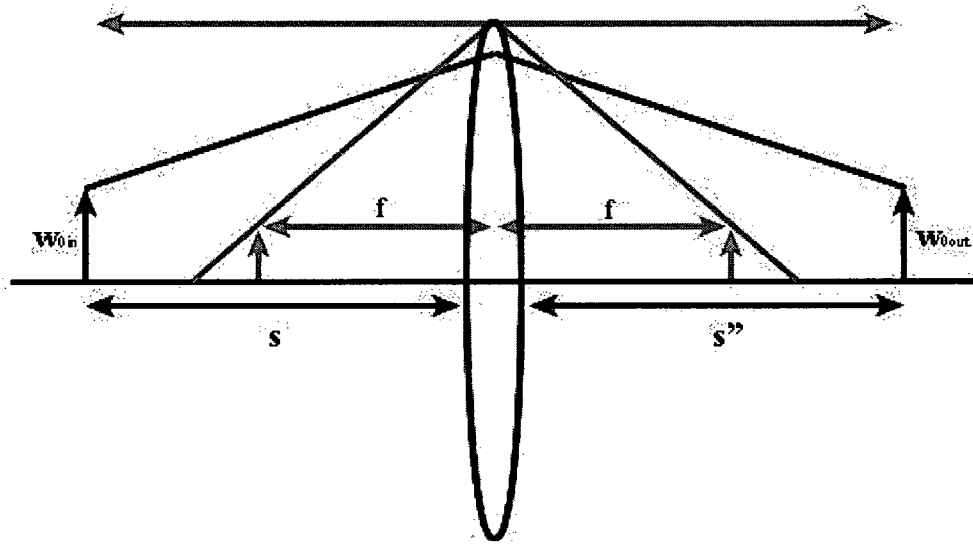

Figure A.3: The standard lens. s is the object distance and s" the image distance.

While equation A.7 gives a good approximation for the spot size of a focused Gaussian beam, a laser with a perfectly Gaussian transverse intensity profile is extremely rare. For a real laser beam, the transverse intensity profile is generally non-Gaussian [54]. The deviation of a real laser beam from a Gaussian beam is characterized by the beam quality factor M^2. Recalling that the half angle of divergence of a Gaussian beam is given by $\theta = \lambda/\pi w_0$. The product of w_0 and θ is λ/π, a constant for any given Gaussian

laser beam. With this product term, a Gaussian beam and a real laser beam can be related by the ratio [25]

$$M^2 = \frac{\theta_R w_{0R}}{\theta w_0} \text{ or } \theta_R w_{0R} = M^2 \frac{\lambda}{\pi} \tag{A.9}$$

where θ_R and w_{0R} are the half angle of divergence and beam waist of a real beam, respectively. Equation A.8 shows that for a real laser beam that is truly Gaussian $M^2 = 1$, while for non-Gaussian beam $M^2 > 1$. Incorporating M^2 into equation A.1 and A.2 gives [25]

$$R_R(z) = z\left[1 + \left(\frac{\pi w_{0R}^2}{\lambda z M^2}\right)^2\right] \tag{A.10}$$

$$w_R(z) = w_{0R}\left[1 + \left(\frac{\lambda z M^2}{\pi w_{0R}^2}\right)^2\right]^{\frac{1}{2}} \tag{A.11}$$

where $R_R(z)$ and $w_R(z)$ are the real beam radius of curvature and real beam radius, respectively. Note that by setting $M = 1$, equation A.10 and A.11 reduces to equation A.1 and A.2, respectively. The real beam Rayleigh range and the real beam half angle of divergence for real beam are given by [25]

$$z_{RR} = \frac{\pi w_{0R}^2}{\lambda} \tag{A.12}$$

$$\theta_R = \frac{\lambda M^2}{\pi w_{0R}} \tag{A.13}$$

By applying the described Self's model, the spot size of a focused real laser beam is given by

$$D_{sR} = \frac{4\lambda f M^2}{\pi D_R} \qquad (A.14)$$

An important factor that affects the spot size is spherical aberration, which is an effect where rays traveling parallel to but different distance away from the optic axis of the lens fail to converge at the same point, leading to a larger spot size. However, since the development of aspheric lens, the effect of spherical aberration is virtually eliminated.