To the Graduate Council:

I am submitting herewith a dissertation written by LiQiang Li entitled "Single Molecule Detection in Solution". I have examined the final copy of this dissertation for form and content and recommended that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Physics.

Lloyd M Davis

Dr. Lloyd M. Davis, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of The Graduate School

# Single Molecule Detection in Solution

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

**LiQiang Li**

August 1994

## DEDICATION

This dissertation is dedicated to

Juliana

and my home country : China

## ACKNOWLEDGMENTS

I would like to thank my major professor Dr. Lloyd M. Davis for his guidance and instruction. His patient hands on help in the laboratory throughout this project was a major factor for the success of this dissertation. I would also like to thank the other committee members, Dr. James W. L. Lewis, Dr. Arthur A. Mason, Dr. Dennis R. Keefer, Dr. Horace W. Crater and Dr. Christian Parigger, for their comments and assistance on this thesis work.

Over the past three years, I received much support and help from faculty and staff members at the Center for Laser Applications. Among those are Newton Wright, Fred Schwartz, David Plemmons, Jim Hornkohl, Jim Few, Bruce Hill, Mary Helen Brittain, Melissa Weatherford and Dee Robinette.

Also I thank members of my family and friends. Without their understanding and support I could not imagine whether I would have continued this difficult but yet competitive profession — Physics.

Finally, the funding for this project from The Whitaker Foundation is gratefully acknowledged.

# Abstract

This work reports the first experiments demonstrating the high overall detection efficiency of single chromophore molecules in solution. The measurements are accomplished using laser-induced fluorescence with time-gated photon counting to isolate solvent molecule signals from promptly scattered light. Compared to prior experiments, the system sensitivity is improved by over an order of magnitude with more efficient light collection and a custom-modified solid-state single photon detector, which is found to have $> 50\%$ quantum efficiency and subnanosecond time jitter. A custom time-gating circuit is implemented to avoid electronic pile-up and loss of signal due to the high count rate ($> 3 \times 10^5\ s^{-1}$). The increased photon detection rate permits single molecules to be detected using flow velocities that are about 40 times faster than in the prior experiments. A novel sample delivery system causes almost all molecules to pass in a narrow stream through the center of the laser beam, whereas in prior experiments only a very small fraction of the molecules would pass through the beam and be detected. In the new system, sample molecules are injected into a high velocity sheath flow through a fine capillary placed directly above the laser beam to reduce diffusional spreading. For sulforhodamine 101 molecules in water, an overall molecule detection efficiency of about 80 % and a quantitative signal indicated by a well defined peak in the histogram of burst amplitudes are obtained. These results agree well with calculations and simulations performed under conditions used in the experiments.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Prior experiments in ultrasensitive detection

In recent times, researchers have made considerable effort to achieve the detection of individual atoms and molecules. In 1977, single atoms in vacuum were detected by S. Hurst *et al.*[1] at Oak Ridge National Laboratory using laser resonance mass ionization techniques. Since then, there has been a variety of novel experiments conducted in condensed phase, although the particular techniques used differ. The major advantage in the case of single atom detection in vacuum is that the surrounding vacuum presents no background signal, while for single molecule detection in condensed matter Raman and Rayleigh scattering from the solid or liquid itself can seriously obscure the signal. New techniques for probing and imaging individual molecules in the condensed phase include electron tunneling microscopy and more recently near-field optical microscopy[2, 3, 4], where an ultra-small probe volume is used to ensure negligible background contributions from Rayleigh and Raman scattering.

For the detection of single molecules in the condensed phase in the presence of a large number of surrounding background molecules, other techniques must be used to enhance signal to noise. In 1989, W.E. Moerner *et al.*[5] demonstrated the detection of single molecules of pentacene in a p-terphenyl host at cryogenic temperatures. The key advantage provided by using low temperatures in experiments

in solids is that the photon absorption cross-section will become enhanced as the homogeneous linewidth becomes narrow. This improvement occurs when contributions from lattice vibrations of the solute molecules in solid matrices are eliminated. The absorption of the single pentacene impurity center is enhanced to such an extent that even the absorption of a laser of very weak intensity can be observed. By using a narrow linewidth laser and low intensity (typically in the $\mu W\ cm^{-2}$ range), the scattering by the host molecules can be essentially eliminated. If the number of solute molecules inside the illuminated sample probe volume is small enough, the broad inhomogeneous absorption profile is resolved into a set of sharp lines corresponding to each individual molecule. In this absorption method, since the laser intensity change that has to be detected is usually very small, fluctuations of laser intensity pose an important limiting factor on the system sensitivity. Subsequently, a fluorescence excitation method was employed due to its zero-background feature.[6] This method, which detects the fluorescence signal intensity rather than the laser intensity, has enabled a variety of experiments on the spectroscopy of single molecules in solid hosts to be conducted.

Unlike the situation in vacuum or in cryogenic solids, where the background is either non-existent or largely eliminated, in liquids the background is always present. Background sources include Rayleigh and Raman scatter from solvent molecules, specular scatter from cell walls, fluorescence from impurity molecules in solution and stray fluorescence from collection objectives, filters and other optical components. A well-designed detection system can largely eliminate all of these except the Raman scattering, because its spectral range overlaps that of the desired fluorescence signal from the sample molecules.

Ultrasensitive detection in solution using laser-induced fluorescence has been a topic of research for some years. Hirschfeld [7] first reported the detection of a single polyethylenimine polymer containing 80 fluorescein isothiocyanate tags bound to $\gamma$-globulin and supported on a microscope slide. Over the years, progressive im-

provements in detection sensitivity have been achieved. Detection techniques at first used analog photodetectors, continuous wave laser excitation and lock-in amplifiers, while later experiments used photon counting detectors. In order to reduce the background due to Raman scatter from solvent molecules and fluorescence emission from the impurities in solution, smaller and smaller sample probe volumes have been employed. In an attempt to reduce the sample volume from which the Raman scattering arises, researchers used a hydrodynamically-focused flow to confine the molecules within a thin stream. Single molecules of the protein B-phycoerythrin with the equivalent of 25 rhodamine 6G chromophores were detected.[8, 9] Unfortunately, in these experiments, the size of the hydrodynamically-focused sample flow is larger than the laser beam diameter so that the property of hydrodynamic focusing was never fully used.

Later, in 1990, the research group at Los Alamos National Laboratory (LANL) reported the first experiment demonstrating the detection of single rhodamine 6G (R6G) molecules in aqueous solution at a concentration of $10^{-13}M$. In that experiment, high-repetition rate pulsed laser excitation and time-gated detection was used to reject Raman scattered light from the solvent.[10] Unlike previous experiments claiming the detection of single molecules from the non-random feature on the autocorrelation function after many large molecules pass individually through the focal region [9], it was possible to implement an algorithm which indicates the passage of individual single chromophore molecules.

## 1.2 Motivation for ultrasensitive detection

Single molecule and single atom detection, the ultimate limit in chemical analysis, has many potential applications. Furthermore, it opens up new insight for the understanding of fundamental phenomena, making it an appealing research topic today. For instance, the detection of single atoms or single ions in vacuum has revealed several phenomena, such as quantum jumps and photon antibunching that

are unobservable in the environment of many individuals. Single molecule detection (SMD) in a solid was motivated by the possibility of optically addressing local spots in solids, which may lead to the development of memory chips with super-high density. It also provides a tool to study the local environment of a single-absorption center in a host of many surrounding molecules. The later may find applications in nondestructive testing and defects inspection for semiconductor chip manufacturing.

SMD and spectroscopy in aqueous solution serves as a tool to understand microphysics at the single molecule level and provides direct access to the distributions of microscopic parameters unavailable by other methods of averaging a group of individuals. S. Soper *et al.*[11] have already demonstrated that simple spectroscopic measurements can be performed on single molecules in solution. They have successfully distinguished fluorescence signals between two different molecular species and have shown that the fluorescence lifetime of a single molecule can be measured.

Single molecule detection in solution was initially motivated by the increasing demand to improve the detection sensitivity in chemical analysis, since it is the ultimate limit for ultrasensitive detection. The experimental technique and the need for improved sensitivity stems back to flow cytometry, which is a technique for the detection of fluorescently-tagged cells and viruses. Thus SMD in the liquid phase is particularly useful in biology and biomedical studies, because in such applications almost all samples are in liquid solutions. Single molecule detection in solution was further motivated recently by its potential application in high-speed DNA sequencing. [12, 13, 14]

## 1.3 Problems with prior experiments

Although an experiment demonstrating single molecule detection in aqueous solution using laser-induced fluorescence has previously been conducted at LANL, there are two main areas which need further improvement.

Firstly, the main criteria in judging the degree of success in a SMD experiment

is the system sensitivity, which in turn largely depends on the fluorescence detection efficiency of the instrument. This detection efficiency depends on the product of the light collection efficiency of the collection objective, the light transmission throughput of the spatial and spectral filters, and the quantum efficiency of the photodetector. High detection efficiency can only be achieved by optimizing the optical arrangement and using components of high optical performance. In the LANL experiment, the detector used was a micro-channel plate photomultiplier tube, which has a quantum efficiency of only $\sim 5\%$. This inherently results in a weak photon burst signal and a low signal-to-noise (S/N) ratio which causes difficulties in determining whether a burst signal is from a single molecule or from noise. Furthermore, because of the low photon detection efficiency, a reasonably long transit time ($\sim 30\ ms$) was required to collect enough photons to detect each single molecule.

Secondly, due to the sample flow cell construction of the LANL experiment, only a small percentage of the R6G molecules was detected. The flow cell used had inside dimensions of $1\ cm \times 0.4\ cm$. The laser beam was focused near one side wall of the cell, because of the small working distance of the collection objective. Since the focused laser beam dimension is much smaller than the cross-section area of the flow cell, most of the R6G dye molecules in the water solution pass through the system without being detected. Also, those molecules which do pass through the probe volume experience illuminations of different intensities, since they pass through different parts of the Gaussian profile of the laser beam.

These two limitations undermine the final goal of rapidly counting each molecule present in the sample solution. Furthermore, the non-uniform illumination of different molecules makes it difficult to quantitatively compare the size of the photon burst from different types of molecules.

## 1.4 Objective and strategy

As discussed above, although the detection of single molecules of R6G was reported in late 1990, there are still areas that need further improvement, namely the low photon burst signal with relatively long molecular residence times and the overall low molecule detection percentage. The objective of this work is to develop a system for the detection of single molecules in solution which addresses the above problems. In this respect the strategies are two fold. Firstly, methods are investigated for increasing the fluorescence detection efficiency, namely by replacing the low quantum efficiency microchannel plate photomultiplier tube (MCP-PMT) by a high quantum efficiency solid state detector, without sacrificing too much on detector timing, and by improving light collection efficiency through the use of a high numerical aperture microscope objective. Secondly, a method is implemented for delivering the sample so that all molecules pass through the tightly focused laser beam with a transit time that is adequate for most molecules to be detected. If these goals can be realized it will be a significant contribution to the area of ultrasensitive detection.

To increase the fluorescence signal, a single photon avalanche photodiode (SPAD) solid state detector was thoroughly tested for its various parameters relating to its application in SMD in solution. These include quantum efficiency, single photon timing-jitter, dead-time and pulse height distribution. The tested results show that this SPAD detector can provide comparable single photon timing but with nearly an order of magnitude higher quantum efficiency than that of the MCP-PMT used in the prior experiments. As a result of this increased detector quantum efficiency and also the implementation of a high numerical aperture collection objective, the count rate due to background goes up significantly, causing the dead time of the time gating electronics to have a significant effect. To counteract the dead time problem, a temporal prefilter is constructed such that most of the counts from Raman scattering photons are eliminated before they reach the slower electronics.

In the second regard, a new flow and sample delivery system is adopted. It

should be pointed out that a hydrodynamically-focused flow system can be used to confine the sample flow into a stream smaller than the focused laser beam as is typically done in flow cytometry. However, hydrodynamic focusing is not suitable for the detection of single chromophore molecules because the flow speed must be very fast in order to achieve a sufficiently small sample stream. This causes the transit time of the molecule through the laser beam to be very short ($\sim 10\ \mu s$). A short transit time means that a molecule can only undergo a small number of cycles of excitation and fluorescence decay, resulting in a photon burst that is too small to detect. The new sample delivery system overcomes these difficulties. It consists of a sheath water flow as in hydrodynamic focusing, but a glass capillary tip that has an opening of only $\sim 0.5\ \mu m$ is used to deliver the ultra-dilute sample solution into the center of the sheath flow. As the sample is carried downstream it spreads due to diffusion, but the tightly focused laser beam is positioned a small distance ($\sim 15\ \mu m$) from the capillary tip so that all molecules pass through the laser before diffusional spreading becomes significant.

A series of measurements, calculations and Monte Carlo simulations were performed to test the new flow system and verify its operation. Then, SMD experiments were conducted using the SPAD and the new flow system using a sample of the multiple-chromophore B-phycoerythrin protein molecule. Due to its large molecular weight ($\sim 250,000$), the diffusion of B-phycoerythrin in water is reduced. Then, sulforhodamine 101 dye molecules were successfully detected in the new flow cell at fairly high flow velocities giving transit times $\sim 1\ ms$. Since all molecules pass through the focused laser beam in similar trajectories, the burst amplitudes are quite uniform resulting in a reasonably well peaked distribution.

The physical principle of SMD by time-gated photon burst detection is the subject of Chapter 2. Appendices A and B give further details on laser-induced fluorescence and competing quenching mechanisms, and on Rayleigh and Raman scattering and their dependence on laser polarization. Chapter 3 gives a detailed discussion of

the SPAD detector. Concise results can be found in Ref.[15] and details of the electrical circuitry are in Appendix C. In Chapter 4, various flow systems are reviewed and the new diffusion-limited flow system as well as the sample stream size are discussed. Details of the experimental setup and the realization of SMD experiments are given in Chapter 5. Chapter 6 addresses issues concerning data acquisition and data processing. Chapter 7 gives calculations of the expected signal size and briefly presents the results of simulations of the experiment. Experimental results, conclusions and suggestions for future work make up the main contents of Chapter 8.

# Chapter 2

# Principle of detection and experimental technique

This chapter explains the method used for the detection of single molecules in solution. §2.1 gives a brief overview of laser-induced fluorescence of organic dye molecules in solution. §2.2 explains the principle of SMD by photon burst detection. §2.3 discusses the major sources of background noise and their reduction. In §2.4 the principle of time-gated detection is elaborated. Time-correlated single photon counting (TCSPC) is discussed in §2.5. Details on the pulse timing are given in §2.6 and the individual major electronic components used in TCSPC are discussed in §2.7. Problems caused by the dead time of the electronics are discussed in §2.8 and a solution to the problem by implementing a temporal prefilter is given in §2.9.

## 2.1 Laser-induced fluorescence

The physical mechanism used in the detection of single dye molecules in solution is *laser-induced fluorescence.*

Fig. 2.1 shows the energy level diagram for a typical organic dye molecule and various mechanisms accompanying the fluorescence process. With light absorption, the molecule passes into a rotationally and vibrationally excited sublevel of the first electronically excited manifold. Thermalization of the rotational and vibrational energy is very fast ($\sim 1ps$). The molecule may then return to the electronic ground

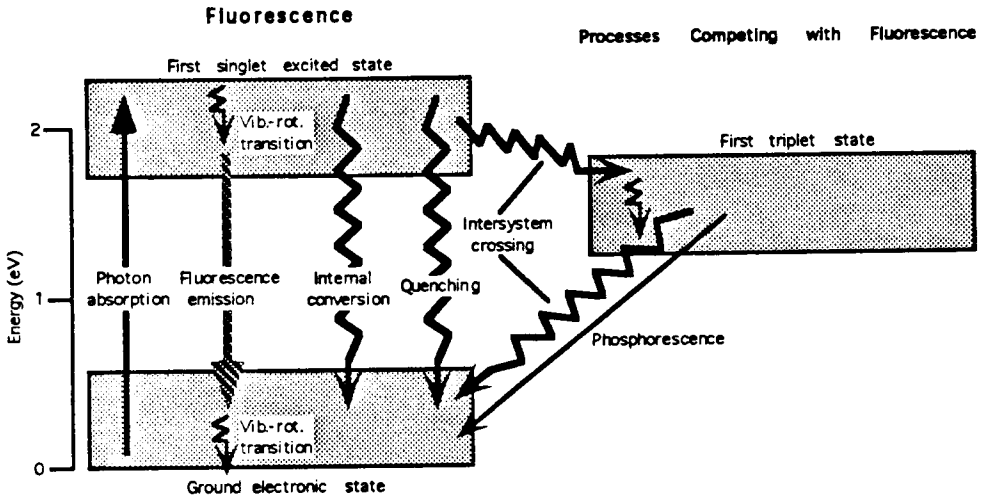



Figure 2.1: Fluorescence and competing processes

state by several competing mechanisms.

The molecule may undergo an electric dipole transition to the ground state with the emission of a fluorescence photon. Besides undergoing fluorescence photon emission, a molecule in the excited singlet state can also decay to the ground state without giving out a photon. The energy is converted either into heat (internal conversion) or into the excitation of another molecule (radiationless energy transfer). Another possibility is that the molecule can pass without emitting a photon to an excited triplet state (intersystem crossing). After a time that is much longer than the fluorescence lifetime, a molecule in the triplet state decays to the ground state, with or without photon emission. If a photon is released, it is called phosphorescence.

No matter what process is involved in the transition, once the molecule returns to the ground state, it should be again available for re-excitation. This is not always the case though. Sometimes absorption causes the molecule to undergo photobleaching, an irreversible change into a different chemical species, after which the molecule can no longer fluoresce. Photodegradation usually occurs from the triplet manifold.[16] The overall probability that an excited molecule will photobleach is

called the photodegradation quantum efficiency. The rate equations and various competing processes as well as polarization properties of fluorescence are discussed in Appendix A.

## 2.2 Photon burst detection

As a single dye molecule passes through the tightly focused laser beam, it undergoes many cycles of laser excitation resulting in the emission of a burst of fluorescence photons. Light is collected from a region in which there are $\sim 6 \times 10^{13}$ solvent molecules around the single dye molecule. The performance of the light collection optics and photon detection apparatus is optimized to efficiently detect the very weak fluorescence photon burst while discriminating most of the strong background, due predominantly to Rayleigh and Raman scattering from solvent molecules. Noise discrimination is accomplished through spectral filtering and time-gated detection, but nevertheless some residual background noise will continue to exist due to unrejected Rayleigh and Raman scattering from the solvent and other sources as discussed in §2.3 below. Fig. 2.2 illustrates a typical sequence of photon counts in a SMD experiment. Photon emission from the illuminated sample volume consists of bursts of fluorescence from single molecules passing through the beam superimposed upon a continuous background of photons.

The number of photons detected from a single molecule should be larger than the size of the photon fluctuations in the background. If this is successfully accomplished, each photon burst above an appropriate threshold will signify the passage of an individual molecule through the laser beam.

## 2.3 Background noise and its reduction

As pointed out earlier, the realization of the detection of single fluorescent molecules in solution depends largely on how successfully one can discriminate the various types of background noise from the weak fluorescence photon burst. Basically, the

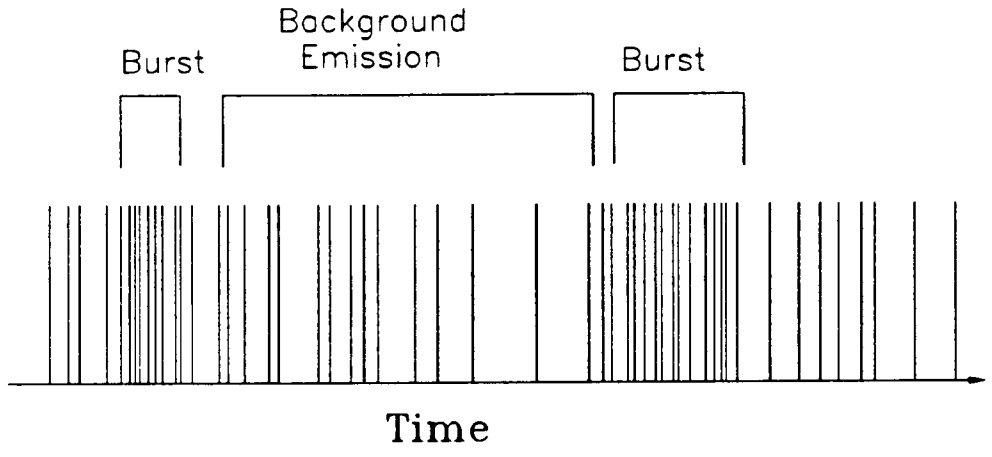



Figure 2.2: Fluorescence photon burst detection

weak fluorescence signal is totally embedded inside the overwhelming noise. Unless the noise can be effectively reduced, one can never achieve sufficient sensitivity to detect the fluorescence signal from a single molecule.

The first source of background comes from specular scattering of light from the cell windows. This can almost completely be eliminated by using a spatial filter at the image plane of the collection objective to limit the region from which light is collected.

The second source of noise comes from Rayleigh scattering from the solvent at the laser excitation wavelength. This noise can be blocked from reaching the photodetector by a bandpass spectral filter whose transmission band is centered near the peak fluorescence wavelength. Also, any stray laser light in the laboratory which leaks into the collection optics will be effectively eliminated by the spectral filter.

The third source of background is Raman scattering from the solvent which is not removed by the spatial or spectral filters. Both Rayleigh and Raman scattering are dependent on direction and polarization (see Appendix B). Thus the first step to

reduce the contribution from these sources of noise is to use horizontally polarized laser excitation and collect fluorescence at right angles to the excitation beam, in the horizontal plane. In an experiment with well-designed optics, one should be able to largely eliminate all scattered light except for the Raman noise, since the spectral range of Raman scattering coincides with the fluorescence emission. Raman noise can only be effectively reduced by the time-gated detection method to be discussed in §2.4 below.

The fourth source of noise, which can not be suppressed by the spatial or spectral filter or by time gating, is induced fluorescence that may possibly come from cell windows, the collection objective and the spectral filters. This can be reduced by using high quality optical components and by placing the spectral filter far from the detector so that possible fluorescence (if any) from the filter losses its intensity considerably before it strikes the detector. Since this source of noise is constant with time it will not lead to bursts of photons with a duration equal to the molecule transit time, but instead will simply add to the background.

Lastly, fluorescence from any impurity molecules in solution simply cannot be eliminated at all. Like the wanted single molecules, impurity molecules can also emit bursts of photons. Thus care must be taken in designing the experiment to make this noise as small as possible. For instance, one has to use very pure solvent and make sure that there are no contamination sources in the rest of the flow system. The design of the flow system should facilitate cleaning with detergent and should avoid dead volumes.

## 2.4 Time-gated detection

Time resolution offers the possibility of reducing the Raman background because Raman scattering is short-lived compared to fluorescence. Researchers doing Raman spectroscopy have been trying to get rid of the interference from fluorescence and doing very well in this respect. This interference is due to spectral overlap as both

Raman and fluorescence are Stokes shifted from the excitation. In this work just the opposite must be done — i.e., Raman interference must be rejected from the "wanted" fluorescence signal. For a homogeneously broadened Raman band, the exponential decay time of the Raman emission, $T_R$, is related to the linewidth of the Raman band, $\Delta\nu$, according to

$$T_R = \pi/\Delta\nu.$$

This equation follows from the Fourier transform relation that applies to homogeneously broadened bands. For organic species in liquids at room temperature, the value of $T_R$ is on the scale of 1 to 10 *ps*. The lifetimes are even shorter for inhomogeneously broadened Raman bands. If a fast photodetector with a time resolution of tens of picoseconds or longer is used, Raman emission appears coincident with the impulse response function.

A typical fluorescence response is illustrated in Fig. 2.3. The nanosecond fluorescence response rises more slowly than the impulse response. In the limit of an infinite fluorescence lifetime it rises as the integral of the impulse response. In general, the fluorescence is delayed compared to the impulse response. This difference in temporal distribution between Raman scattering and fluorescence emission means significant enhancement of fluorescence over Raman can be achieved through time-gating. The reasonably fast time response of the SPAD detector (see §3.3.3) permits us to distinguish photons that are coincident with respect to the laser pulse from photons that are delayed.

Time spectra obtained from pure solvent and ultra-dilute R6G solution are shown in a semilogarithmic plot in Fig. 2.4.

The strong peak at the time of the laser pulse is due primarily to Raman scattered photons. For an ultradilute dye solution, the prompt peak is followed by an exponential decay curve with a lifetime characteristic of the dye. By setting a time window that includes only delayed photons, one can reject the Raman and retain

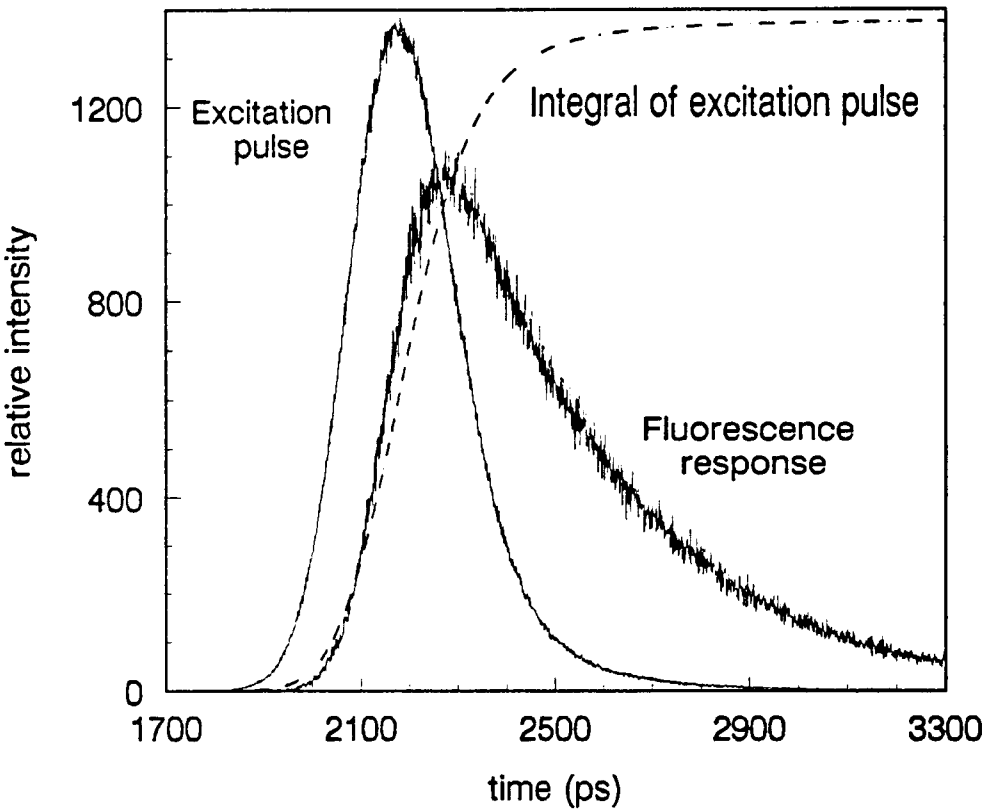



Figure 2.3: Raman and fluorescence emission

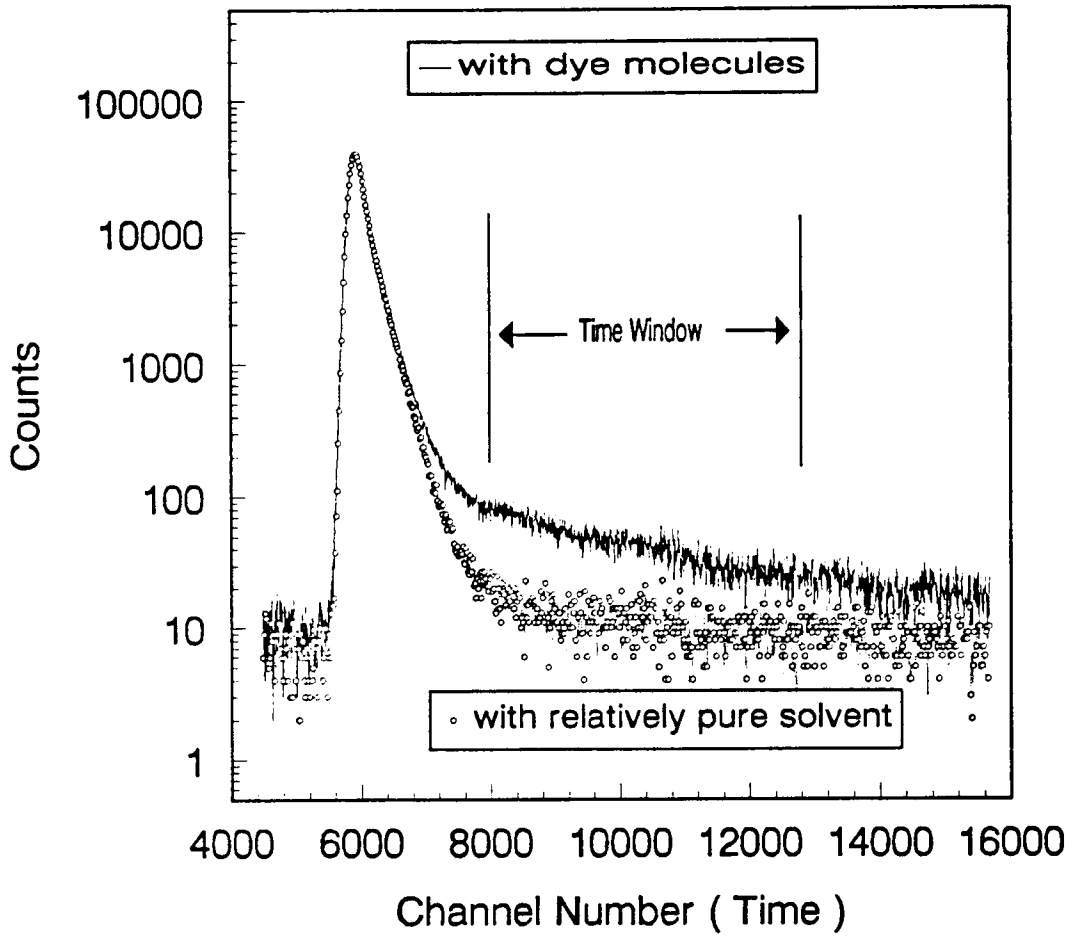



Figure 2.4: Time-gated detection

the majority of the desired fluorescence signal. In this way, the signal-to-noise ratio can be improved substantially. Also, background from sources that are constant with time, such as detector dark count or stray fluorescence that decays with a very long lifetime, can be partially rejected by setting an appropriately short window.

Time-gated detection as described above is achieved by using a technique called time-correlated single photon counting (TCSPC). The principle of TCSPC and its experimental implementation will be discussed in §2.6, but first an explanation of the different time scales involved in using time-gated detection is given.

## 2.5 Laser pulse and photon timing

After promptly scattered photons are rejected by the time-gating electronics, fluorescence photons that fall within the window are used for photon burst detection, as described previously in §2.2. The photon burst events from each single molecule occur on a considerably slower time scale than the nanosecond fluorescence decay.

Fig. 2.5 illustrates the relevant time scales involved in the single molecule detection experiment. In part (a), molecules are assumed to pass through the laser beam at random times, approximately 0.4 s apart on average, as indicated by the vertical lines. In part (b), the average time a molecule spends inside the beam is assumed to be $\sim 1$ *ms*. The molecule emits a burst of photons during this period until it exits from the laser beam. ends

Part (c) indicates laser pulses contained in a time interval during which there is a fluorescence photon. The mode-locked Nd:YAG laser outputs pulses about 0.07 *ns* in duration. The repetition rate is 76 $MHz$ and thus pulses repeat every 13.2 *ns*.

Since the duration of the laser pulses is much shorter than the fluorescence lifetime of the molecule, each pulse can excite the molecule at most once, thus inducing at most one fluorescence photon. Also, the interval between adjacent pulses is several times longer than the fluorescence lifetime. This assures that if a molecule is excited by a pulse it will return to the ground state and be ready for another

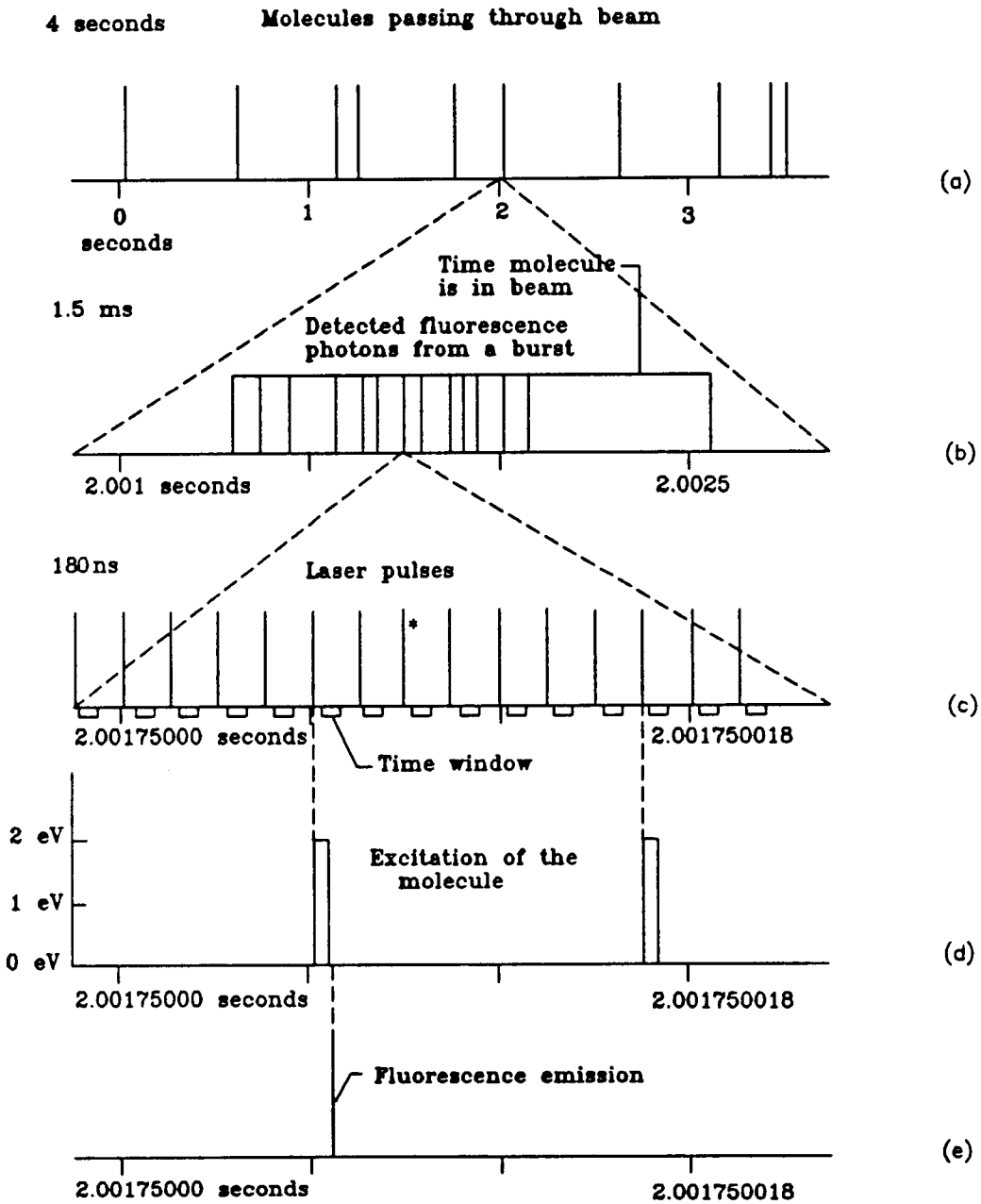



Figure 2.5: Different time scales in single molecule detection

excitation by the time the next pulse arrives. Therefore, in principle, every pulse may cause the molecule to be excited and hence to emit one fluorescence photon.

The time windows during which we count photons are indicated by the small brackets in part (c), which are typically set from 1 to 5 *ns* after each pulse. On average only one laser pulse in every ten excites the molecule, as shown in part (d). For a molecule whose fluorescence quantum yield is 0.5, half the excitations lead to the release of a photon. In most cases a fluorescence photon will fall within the window, as shown in part (e).

## 2.6 Time-correlated single photon counting

Detection of single photons within a preset time window is accomplished by time-correlated single photon counting (TCSPC), as mentioned in §2.4. TCSPC is actually a standard technique used for measuring fluorescence lifetimes. The basic concept relies on the fact that the fluorescence intensity decay profile can be built up by exciting the sample a large number of times to accumulate a histogram of time intervals between the excitation and the emission of a single photon.

A schematic diagram of the instrumentation in a conventional TCSPC set-up is shown in Fig. 2.6. The set-up consists of two channels, the excitation channel, which provides the timing of the excitation laser pulse, and the emission channel, which provides the timing of the subsequently emitted fluorescence photon. Each channel contains a photodetector and possibly a fast timing discriminator to obtain a well-defined timing signal from either the excitation or emission.

The key component is the time-to-amplitude converter (TAC), which converts the time duration between excitation and emission into an output pulse whose amplitude is proportional to the time interval. In a conventional set-up, the TAC START pulse comes from the excitation channel while the STOP pulse comes from the emission channel. However, in the set-up used in this work the excitation channel provides pulses at the laser repetition rate of 76 $MHz$, which is considerably

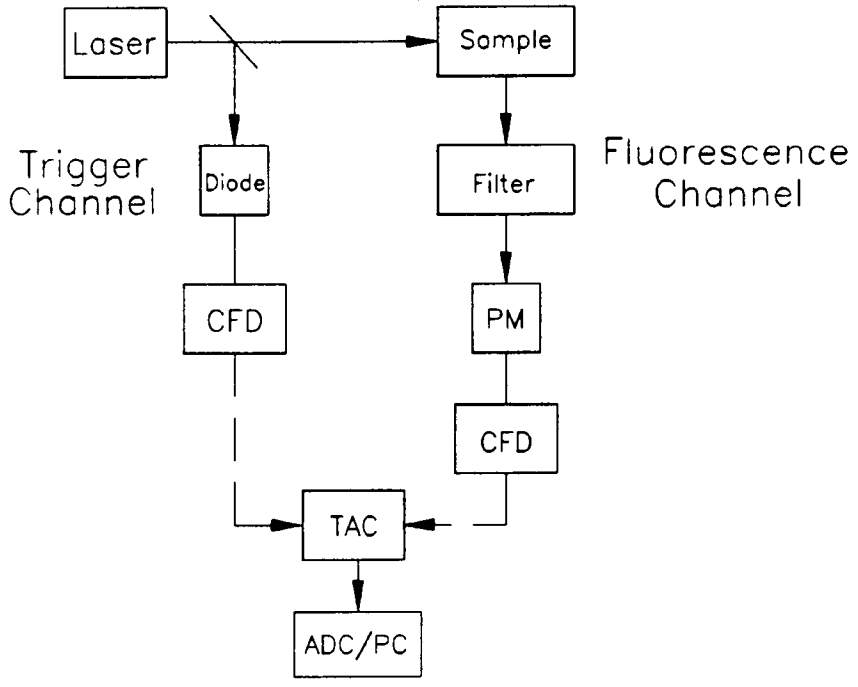



Figure 2.6: Conventional set-up of TCSPC

higher than the maximum rate at which the TAC can accept START pulses. Consequently, in all experiments, the TAC is operated in "reverse mode"[17] by using the signal from the fluorescence photon as the START input for the TAC and the signal from a later excitation pulse as the STOP. In this case, the accumulated histogram of time intervals must be reversed by software to obtain the fluorescence intensity decay profile as a function of increasing time, as shown for example in Fig. 2.4. "Reverse mode" operation has the advantage that every accepted start pulse due to a single photon initiates a successful TAC conversion so that the busy time of the TAC is minimized.

Note furthermore, that we do not use the excitation pulse that immediately follows the fluorescence photon to STOP the TAC. The reason is that the TAC linearity is best if the time interval between START and STOP pulses is near the middle of its conversion range, which is set at its fastest value of 50 $ns$. Instead, the second excitation pulse following the fluorescence is used, as indicated by the

asterisk in Fig. 2.5 (c). This pulse is selected by applying a suitably timed pulse to the STOP GATE. A pulse from the emission channel with width adjusted to be $\sim 13\ ns$ is used for this purpose.

The analog TAC output passes to a multichannel analyzer (MCA) for the construction of the histogram from which the fluorescence decay profile is obtained. Indeed, this is all that is needed for fluorescence lifetime measurement. For time-gated single photon counting, the next step may be performed by passing the analog output of the TAC through a single channel analyzer (SCA) which selects the lower and upper limits of the time window (see Fig. 2.4). Alternatively, as shown in Fig. 2.6, the TAC output can be digitized by an analog-to-digital converter (ADC) and sent to a computer for setting up the time window through software or performing other analysis including the construction of the histogram.

## 2.7 Electronics

This section gives details on the individual electronic components used for TCSPC.

1. NIM Bin:

   Most electronic components used in TCSPC are standard Nuclear Instrument Modules (NIMs). They can be inserted into sockets that are conveniently built on a power supply called a NIM bin. NIMs take their power supply through connections on the rear panel of the NIM Bin crate. Voltages of $\pm 12\ v$ d.c., $\pm 24\ v$ d.c. and $\pm 6\ v$ d.c. are provided by the NIM Bin. The NIM Bins are robust but require occasional calibration and service. In the experiment, two Tennelec TB 3 NIM bins are used.

2. Discriminators:

   Due to statistical variations in the gain, the single photon detector from the fluorescence channel outputs pulses with a distribution of pulse heights. The output from the photodiode in the excitation channel also varies due to laser

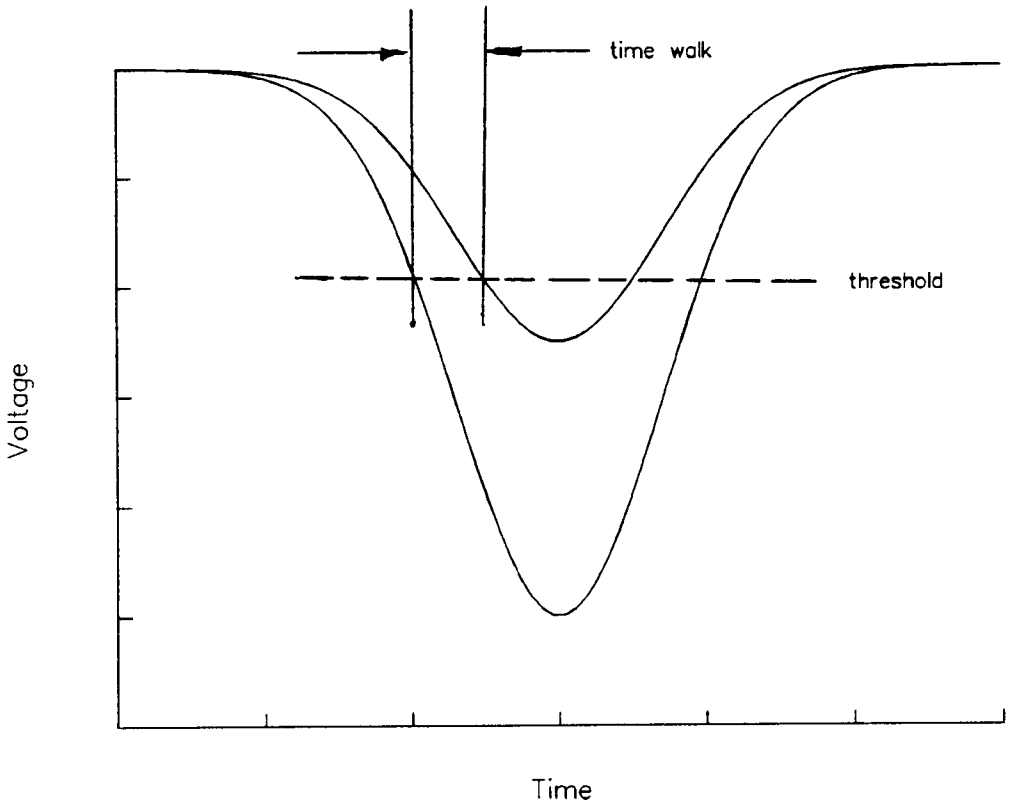



Figure 2.7: Timing-error caused by LED

intensity fluctuations. In order to get good timing from the TAC, it is necessary to furnish the TAC with constant amplitude pulses. Thus detector pulses are conditioned by timing discriminators before being passed to the START and STOP channels of the TAC.

Two types of discriminators are possible: leading-edge discriminators (LED) or constant fraction discriminators (CFD). The LED introduces time walk due to the variation in pulse amplitudes. As illustrated in Fig. 2.7, pulses of different amplitudes, which come from photons emitted the same time after the excitation, cross the discriminator threshold at different times.

To reduce time walk caused at a LED, a CFD can be employed. As the name implies, with this type of discriminator the pulses are timed from a point on the leading edge that is a fixed fraction, $f$, of the pulse amplitude. For pulses of constant or similar shape but varying amplitudes this timing point

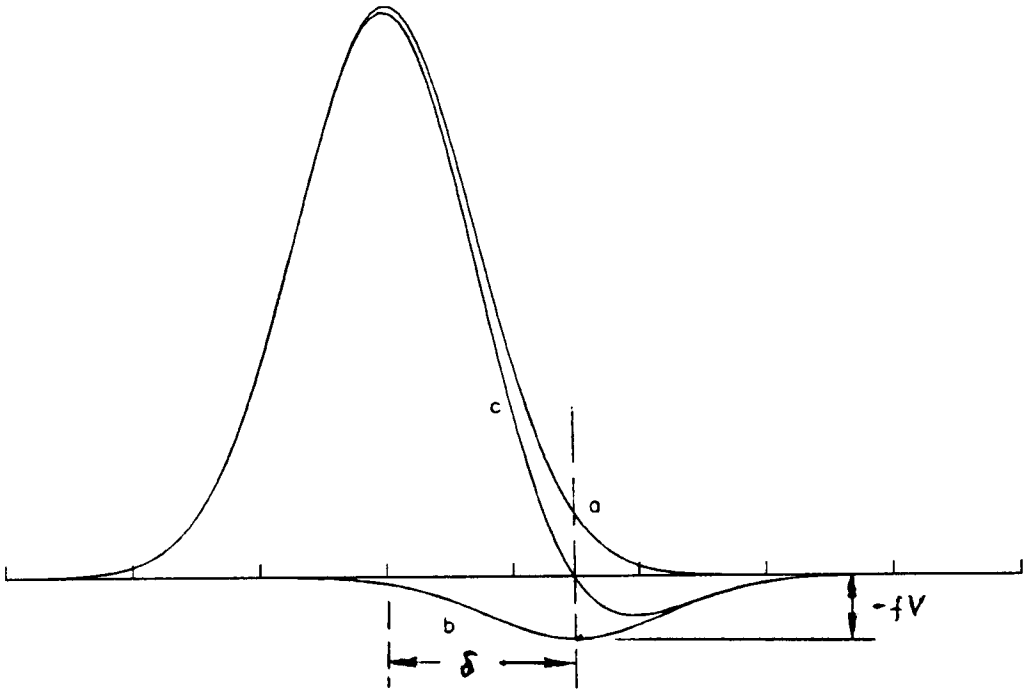



Figure 2.8: Timing using a CFD

is constant.

Fig. 2.8 illustrates the way by which the constant fraction timing is achieved. Suppose the input pulse has amplitude $V$ and is to be timed at an amplitude $fV$ on its leading edge. Inside the discriminator, the input pulse is first split into two path, one of which is delayed by an amount $\delta$ and inverted [a], while the other is attenuated to a amplitude $fV$ [b] but with no delay. Note that time is increasing towards the left. The two signals are then recombined to form the zero crossing signal [c]. In this way, even though there is a change in pulse amplitude, the zero crossing point remains the same.

Two Tennelec quad CFD discriminators (TC 454 and TC455), each with four separate discriminators in one NIM module, are used. Each discriminator can be configured as either a CFD or LED and provides up to three outputs with standard NIM negative signals.

3. Amplifier:

   If possible, the single photon detector should provide enough gain for its output pulse to directly trigger the CFD. If an amplifier is used before the CFD, it may introduce noise which gives time jitter. Also it may act as an antenna picking up radio frequency noise. However, when the detector output pulse has a small amplitude, better performance from the CFD is obtained by using a low noise wide bandwidth amplifier. A fast amplifier from B& H Electronics (Model AC2010H) which has a 150 $ps$ rise time (2.2 $GHz\ BW$) and a 20 dB gain is used.

4. The TAC:

   The TAC works in the following way. Upon receipt of a START signal, a timing capacitor is charged linearly from a constant current source. The charging process is interrupted upon acceptance of a STOP pulse and after a fixed delay an output pulse is generated with an amplitude derived from the final charge in the capacitor. Therefore the output pulse height is proportional to the time difference between the START and STOP pulses. If a STOP pulse is not received after a time called the TAC range, charging is automatically discontinued. Again a fixed time elapses before the capacitor is reset, at which time the TAC is ready for another START pulse. The time which elapses between a START pulse and the time at which the TAC is ready for the next START pulse is called the dead time of the TAC. It limits the maximum photon count rate which the TAC can process.

## 2.8 Dead-time of electronics

It is known from the previous section that it takes time for the TAC to process a pair of START and STOP pulses and to convert this time interval into a pulse amplitude. The TAC will not accept another START pulse during the period of the

TAC dead time, which is approximately 5 $\mu s$.

The burst nature of the fluorescence photons determines that the instantaneous photon count rate is high. Furthermore, the Raman count rate in our experiments is considerably higher than that of previous experiments due to the improved detector quantum efficiency and the higher numerical aperture collection objective used. A count rate of 2–3$\times 10^5$ $s^{-1}$, due largely to Raman photons, is typical in most of the experiments. At such a count rate, the TAC spends most of its time processing the Raman photons while a good portion of the "precious" fluorescence photons are lost.

For a steady state random emission, the probability distribution for the number of photons arriving in a fixed time interval follows Poisson statistics. Hence, the probability density function of time intervals between successive photon arrivals, $p(u)$, is exponential, i.e.,

$$p(u)du = \frac{1}{\overline{u}} \exp(-\frac{u}{\overline{u}})du, \tag{2.1}$$

where $\overline{u}$ is the mean time between photons. If the count rate is $2 \times 10^5$ $s^{-1}$, (i.e., $\overline{u} = 5$ $\mu s$) and the TAC dead time is $t_{dead} = 5$ $\mu s$, then the electronic throughput drops to

$$\int_{t_{dead}}^{\infty} p(u)du = 1 - \int_0^{t_{dead}} \frac{1}{\overline{u}} \exp(-\frac{u}{\overline{u}})du = 0.37, \tag{2.2}$$

or $\sim 37\%$, since only those photons which arrive after the dead time will be accepted by the TAC. Thus we see that the TAC dead time gives rise to a considerable loss in signal.

## 2.9 Temporal prefilter

To avoid the loss of fluorescence photons due to the TAC dead time, a temporal prefilter is implemented before the TAC, reducing the number of Raman pulses which arrive at the START input of the TAC. The temporal prefilter acts overall

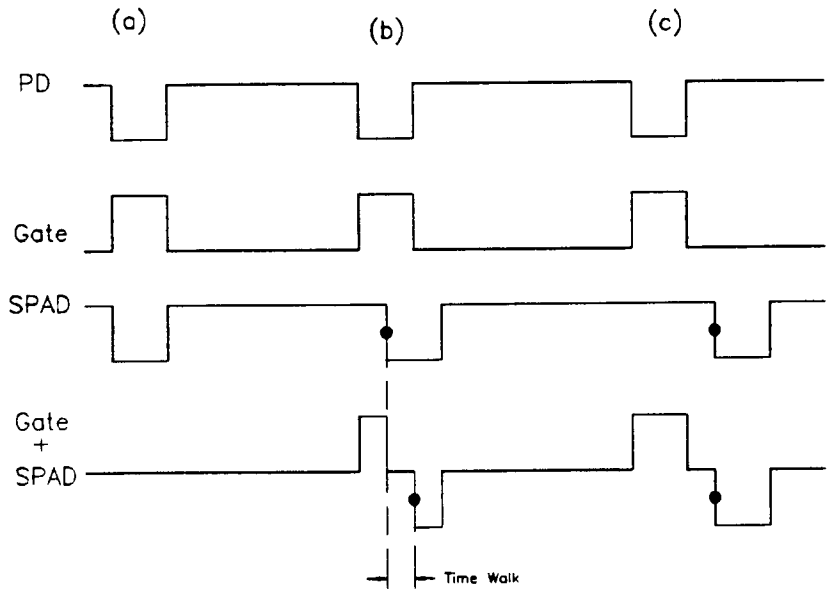



Figure 2.9: Conventional anticoincidence gate

as a subnanosecond anticoincidence logic gate to filter out those pulses that occur within a certain interval about the excitation pulse. In this way, the START count rate will be much smaller than the SPAD total count rate and most of the pulses accepted by the TAC for processing will come from fluorescence.

Although there is a START GATE input on the TAC that may be used as an anticoincidence gate, it does not permit subnanosecond gating. Furthermore, it causes the timing of accepted START pulses near the edge of the GATE to become degraded. The reason for the latter is that this conventional anticoincidence gating logic in essence works by simply adding the START pulse (from the SPAD) with the GATE pulse (from the photodiode (PD)) and passing the *sum* signal to a LED. This works fine if the SPAD pulses are exactly coincident (due to Raman) or completely anticoincident (due to delayed fluorescence) with the GATE pulse. This is illustrated, using square pulses for simplicity, in Fig. 2.9 *a* and *c* respectively. However, due to the time jitter of the SPAD, Raman photons occupy a certain time

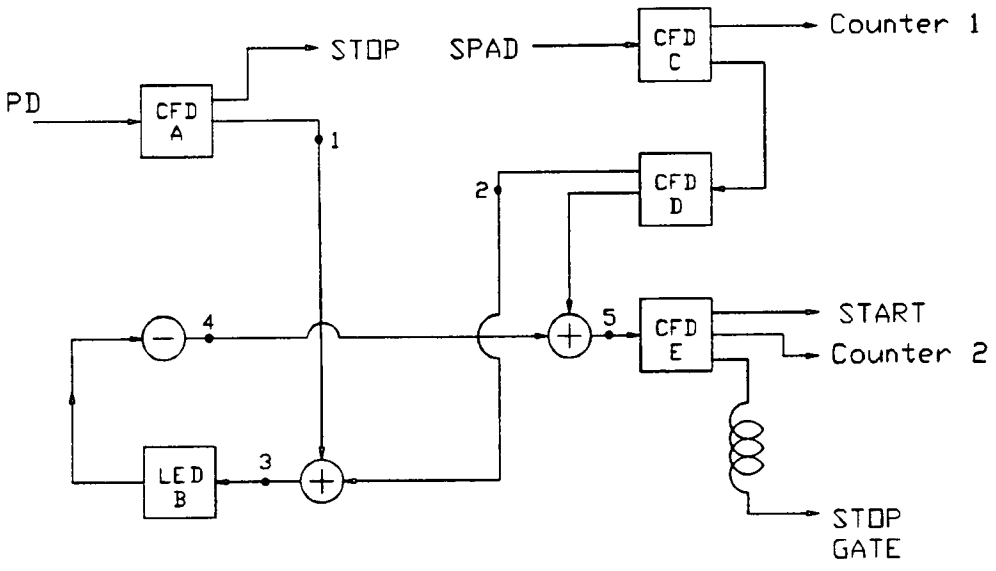



Figure 2.10: Schematic of temporal pre-filter

interval on the detector prompt time spectrum (discussed in §3.3.3). When a SPAD pulse due to Raman is not exactly coincident with the GATE pulse but arrives late as shown in Fig. 2.9 *b*, the sum of the pulses will not be zero. Consequently, the discriminator will trigger and furthermore will not retain the timing of the SPAD pulse. Thus Raman photons which arrive near the falling edge of the GATE will not be eliminated but will become delayed from their original time position and hence appear to be fluorescence photons.

In order to overcome this problem, a two stage coincident-anticoincident sub-nanosecond gating scheme is implemented. Fig. 2.10 shows the logic diagram of this gating circuit. The signals at different positions of Fig. 2.10 are shown in Fig. 2.11.

CFD A provides pulse conditioning for the PD pulses, while CFD C provides pulse conditioning for the SPAD pulses.

In the first stage, SPAD pulses are added with PD pulses (CFD D + CFD A) and the sum signal, shown at point 3 in the figures, passes to LED B, which acts

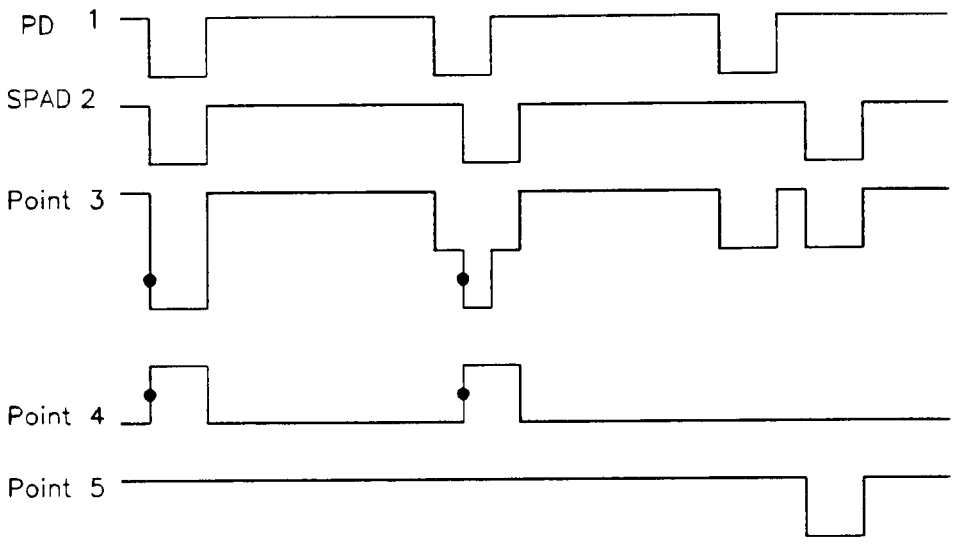



Figure 2.11: Signals at different positions in the prefilter

as a coincidence gate. Thus LED B outputs a signal only when the pulses from CFD D and CFD A overlap to some degree. As in the case of the anticoincidence circuit discussed in Fig. 2.9, the timing of the output pulse of LED B may not exactly follow the SPAD pulse. In reality, the discriminators produce Gaussian–like standard negative NIM pulses, rather than the square pulses used to illustrate the circuit. The threshold of LED B can be adjusted to control the amount of overlap between SPAD and PD pulses, which is also the width of the cut on the prompt time spectrum.

In the second stage, the output of LED B is inverted, as shown at point 4, and added to the SPAD pulse again to give a signal shown at point 5, in which the timing of the SPAD pulse has been recovered. The anticoincidence circuit is completed by passing this sum signal to CFD E, which provides a START signal to the TAC.

Due to the extra delay introduced by each discriminator (either CFD or LED), various signal delays must be carefully adjusted using cables of appropriate length. The timing relationships among the different signals must be checked using a fast

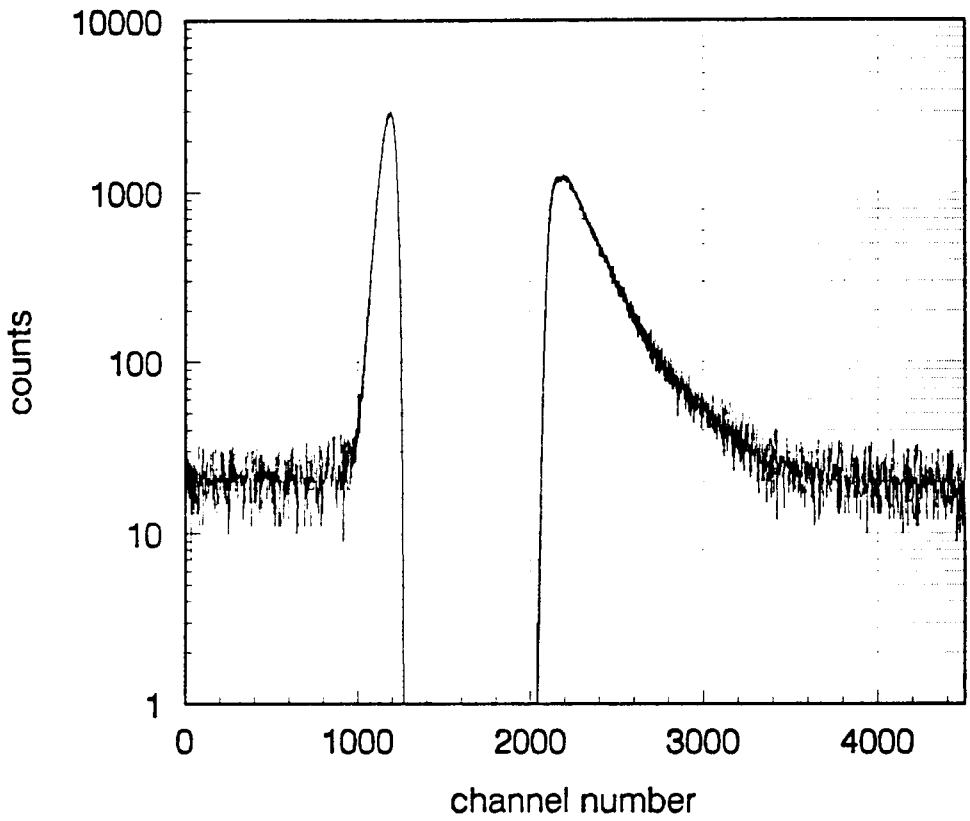



Figure 2.12: Time spectrum after the rejection of Raman photons

oscilloscope.

Fig. 2.12 shows the time spectrum of the SPAD response after employing the pre-filter before the TAC. After the prefilter, the photon counts concentrated near the peak of the response function due to Raman scattering have been removed from the TAC START signal. This alleviates the TAC dead time problem caused by the high count rate so that the system is less likely to loose fluorescence photons.

A related problem concerns the dead-time of the SPAD detector itself. Its relative importance and relation with the TAC dead-time will be discussed in §3.4.

# Chapter 3

# Single photon avalanche diode detector

## 3.1 Principle of photon counting

In this world of digital displays, almost all sensors, whether mechanical, electrical or optical, generate analog rather than digital signals. One of the few exceptions is the photomultiplier tube (PMT). Its digital nature stems from the photo-electric effect, discovered by Hallwachs in 1888 and explained by Einstein in 1905. The theory basically says that the ejection of electrons from illuminated surfaces may be understood if light is considered to be comprised of individual quanta, which were later called photons. The rain of photons striking on the cathode surface of a PMT produces discrete charge pulses at the anode, causing statistical fluctuations known as "shot noise." Although this process entails burdensome noise in analog measurements, it is the basis of a technique called photon counting (PC).

### 3.1.1 Theory of single photon counting

Single photon detection and timing essentially means the detection and timing of a single photoelectron released by light from a photosensitive surface of a photomultiplier tube, solid state avalanche photodiode, or other photodetector. Single photoelectron detection requires sufficiently low noise amplification to overcome the Johnson noise in the first resistor of the eventual electronic circuit, plus additional

amplification to trigger the standard discriminator circuits. Most of the gain usually comes before the electronic circuit, within the photodetector itself. This gain, which is obtained by secondary electron emission processes down a chain of dynodes, or alternatively avalanche processes within the junction of the photodiode, is subject to statistical fluctuations.

For a PMT, if $R_k$ is the incoming photon rate and $q_k$ is the photocathode quantum efficiency, then the rate of photoelectrons leaving the cathode is $q_k R_k$. Only a fraction, $f$, of these photoelectrons generate secondary electrons at the first dynode of the PMT. Thus the rate of anode pulses $R_a$ is

$$R_a = f q_k R_k. \tag{3.1}$$

Each of these pulses is amplified in the dynode chain so that it generates between $10^5$ and $10^7$ electrons at the anode. This number is known as the gain, $G$, of the PMT.

Johnson noise (or thermal noise) is a random variation in the current or voltage at the electrical contacts of a resistor due to the random thermal motion of electrons within the resistor. It is present in all resistive materials at temperatures above absolute zero and can be interpreted as being due to the excitation of electrons by the blackbody radiant energy which is present within any material. From the application of statistical mechanics to an electronic circuit it is known that the average free energy that can be delivered by the resistor to a matched load is $kT$, where $k$ is Boltzmann's constant and $T$ is the absolute temperature. For matched conditions, the power is then $P = kTB$, where $B$ is the noise bandwidth. Relating the power to the current, $i$, using $P = i^2 R_L$, where $R_L$ is the load resistor, the root mean square current fluctuations, $i_{rms}$, due to the Johnson noise is found to be

$$i_{rms} = \sqrt{kTB/R_L}\,.$$

For the detector used in this work, $T \sim 300K$, $R_L = 50\Omega$ and $B \sim 2.2GHz$ (the

bandwidth of the fast amplifier) and thus

$$i_{rms} = 4.3 \times 10^{-7} A \quad \Longrightarrow \quad v_{rms} = 2.1 \times 10^{-5} V \ ,$$

where $v_{rms}$ is the root mean square voltage fluctuations. The amount of internal gain $G$ needed for single photon detection can be calculated as follows. Consider an electron of charge $e$ produced by photoemission at the surface of a photodevice. This electron is amplified to a charge $Ge$, which in turn gives rise to a pulse of voltage $V \sim GeBR_L$ at the load resistor. Thus to produce a voltage pulse greater than the Johnson noise at the fast amplifier the detector must have a gain

$$G \gg \frac{v_{rms}}{eBR_L} \sim 1200. \tag{3.2}$$

However the output from the amplifier, which has a gain of 10, must trigger the discriminator, which gives $R = 50\Omega$, bandwidth $B' \sim 200\ MHz$ and a minimum threshold of $v_t \sim 5mV$. Thus the photodetector gain must be

$$G > \frac{v_t}{10eB'R_L} \sim 3 \times 10^5. \tag{3.3}$$

### 3.1.2 Analog detection

In analog detection the anode current of the PMT is measured directly with a pico-ammeter and the noise associated with DC amplification is high. The most common means of reducing noise is by adding a RC low-pass filter, which serves as an averaging device to smooth out fast fluctuations.

The major sources of detector noise are from dark current, shot noise, Johnson noise, amplifier noise and $1/f$ noise. The type of noise known as 1/f noise has a dependence on signal modulation frequency of the form $1/f^n$, where $f$ is the modulation frequency and $n$ is some number. The origins of this noise are not well understood.

To improve the signal-to-noise ratio, S/N, it is possible to modulate or chop the signal and use a narrow bandwidth amplifier tuned to the modulation frequency to reject noise outside the amplifier bandpass. In this way, the signal-to-noise ratio can be significantly improved, since 1/f noise decreases rapidly with increasing frequency. However, instabilities in the chopping frequency would require an increase in the bandpass, which leads to increased noise. To overcome this, lock-in amplifiers with a very narrow bandpass locked to the chopper frequency may be used. Despite such advances, analog detection methods still have several troublesome characteristics. For example, slow drifts in the PMT high voltage cause the gain, and hence output current, to vary typically by ten times as much. If the drift of the high-voltage is $\sim 0.01\%$, changes of $\sim 0.1\%$ in the output current can be expected. This imposes a limitation on the ability of the amplifier to distinguish between two signals of nearly equal amplitude.

### 3.1.3 Photon counting versus analog detection

The omission of $G$ in Eqn. 3.1 illustrates the stability advantage of PC. Unlike the situation in analog amplification, the anode voltage on a PMT can vary as much as 2.5% around the nominal bias voltage without disturbing the count rate. Even in the dark, thermal emission releases electrons from the dynodes as well as from the cathode. In analog detection, one cannot separate photon induced signal from current due to thermal emission. In PC mode, photon induced pulses can often be distinguished from thermal pulses through the difference in pulse height. This ensures maximization of the light count to dark count ratio. Also counting individual photons makes possible a digital integration technique, promoting further improvement in S/N over lock-in amplification. By fast switching to a gate, pulses are counted and stored for any amount of integration time and with this form of digital integration, 1/f noise is eliminated.

The sensitivity for analog detection is low since many photons are needed to

generate a measurable signal. Although it can be improved through the methods discussed above, namely longer measurement time, RC smoothing and lock-in amplification, still it can not match the high sensitivity provided by single photon counting. Linearity and dynamic range in analog detection are limited due to gain saturation caused by space charge effects. However, for single photon counting, the dynamic range is very large, especially towards the low intensity end, and the linearity is excellent within the dynamic range.

PC is not without its limitations. One of these is the restriction in the maximum acceptable light level. From the moment a discriminator begins to produce an output pulse, any other pulse arriving within the dead time of the detector or electronics will be ignored (see §2.8 and §3.4).

## 3.2 Avalanche photodiodes

### 3.2.1 Principle

Detection of single photons requires high internal gain as discussed in §3.1.1. PMTs have traditionally been the detector of choice for such applications. In principle, an avalanche photodiode (APD) is a similar device to a PMT. An APD is a solid-state photodiode that operates at a reverse bias voltage which causes avalanche photoelectron multiplication near the p-n junction region. APDs have excellent potential for replacing PMTs in many applications, partly due to their simplicity, smaller size and ease of operation. However, until recently it has been very difficult to make APDs with sufficient internal gain for single photon detection. A reverse bias of a few hundred volts across the APD can only generate a gain factor in the range of $10^2$ – $10^3$. For a typical photodiode junction, a reverse bias of tens of volts will lead to an unacceptably large dark current. Moreover, there will be excessive noise associated with "leakage" current across the edges of the junction, unless the edges are well protected. At reverse bias levels in the hundreds of volts, typical photodiode junctions break down completely, acting as open circuits instead of as

diodes. Recent technological breakthroughs in APD manufacturing have overcome most of these difficulties and new APD devices are beginning to be used more and more as alternative devices to PMTs. In an APD, the dark current increases linearly with the size of the active area. Consequently, dark current can be decreased significantly by reducing the size of the active area. Newly designed APD devices that are used at a reverse bias voltage above breakdown to achieve very high internal gain ($> 10^5$) are called single photon avalanche diodes (SPAD).

### 3.2.2 RCA C30902S detector

In ultrasensitive detection and spectroscopy experiments such as single molecule detection, detectors with single photon detection capability and high quantum efficiency are required. Furthermore, in order to implement the time-gated detection scheme for rejecting photons from Raman scattering (see §2.4), the detector must have subnanosecond single photon timing jitter. In previous experiments, the single photon detector used was a microchannel plate photomultiplier tube (MCP-PMT). This detector has very fast single photon timing with full width at half maximum (FWHM) $< 200\ ps$, but its response curve has a long tail which is unwanted for time-gated detection. More importantly, its quantum efficiency is only $\sim 5 - 8$ % which results in a low overall system detection efficiency. To improve the detection sensitivity we set out to find a new single photon detector, focusing our search on SPAD detectors, since they alone offer the prospect of obtaining higher quantum efficiency together with subnanosecond timing. Tests were first conducted on the EG&G avalanche diode, RCA C30902S, which was the most promising commercially available SPAD device on the market at the time this project was started. It has a 500 $\mu m$ diameter active area and an internal gain approaching $10^6$. According to the data sheet, one should be able to achieve a single photon detection quantum efficiency of up to $\sim 50\%$ from this detector. Although the RCA C30902S is selected for low noise, as designated by the S in the part number, the dark count rate was

found to be $\sim$ 7900 $s^{-1}$ at a bias of 10 $V$ above breakdown, much too high for our SMD experiment. For experimental testing purposes, we then constructed a detector housing that allows the SPAD to be cooled down to liquid nitrogen ($LN_2$) temperature while still permitting the light to be focused tightly onto the SPAD using a microscope objective. When the SPAD was fully cooled, the dark count rate did drop significantly to a very low level ($< 10\ s^{-1}$). However, charge trapping, which causes after-pulse formation, becomes a serious problem at this temperature and causes the count rate to become unstable. As for the single photon timing, it is relatively easy to achieve a FWHM of $\sim$ 600$ps$. However, we only achieved a quantum efficiency of a few percent which is considerably less than the maximum value attainable according to the data sheet, although another research group[18] reported a quantum efficiency of $\sim$ 15% and a dark count rate of 500 $s^{-1}$ at $-25°C$. We found it was very hard to adjust the focusing of the light onto the small active area of the detector since it was necessary to mount the microscope objective in close proximity to the SPAD within the cooled housing. Also, the housing did not permit the temperature to be held constant for long, and consequently the breakdown voltage alters, making it difficult to determine the exact bias voltage applied above breakdown. Following correspondence with EG&G regarding the RCA C30902S, EG&G agreed to try to modify the SPAD from their newly available SPCM-200 photon counting module to obtain faster timing. They loaned us such a modified prototype to test for our application in SMD. The modification and test of this SPAD will be discussed in detail below.

### 3.2.3 SPAD detector custom modification

Instead of the usual detector-preamplifier module, the new SPAD detector loaned to us by EG&G Canada is custom modified without the preamplifier from their commercial SPCM-200 module. The SPAD is mounted on a two stage thermoelectric cooler, in series with a 560 $K\Omega$ chip resistor for passive quenching and a 150 $\Omega$ chip

resistor as a load resistor. This is different from the commercial module which has a larger load resistor and greater circuit capacitance. The thermoelectric cooler, the SPAD and the quenching and load resistors are all hermetically sealed inside a TO-8 package. In order to minimize the stray capacitance across the load resistor, a 50 $\Omega$ output cable was soldered directly across the load resistor. We constructed the SPAD bias circuit and the thermoelectric cooler driver, and a detailed electric circuit diagram can be found in Appendix C. The driver with a feedback circuit delivers the appropriate current for the cooler by sensing the resistance of a thermistor mounted alongside the SPAD on the thermoelectric cooler. The thermoelectric cooler keeps the operating temperature of the SPAD $\sim -45 \pm 4^{\circ}\ C$. The breakdown voltage of the SPAD detector depends on the operating temperature and it increases with temperature. At a temperature around $-45^{\circ}\ C$, the breakdown voltage is $-330\ V$.

### 3.2.4 SPAD quenching

Above the breakdown voltage, an optically or thermally induced electron initiates an avalanche chain reaction of ionization. The avalanche current continues to flow until a circuit quenches the avalanche by lowering the bias voltage close to or below the breakdown voltage. After a finite dead time, the working voltage of the device is finally restored in order to make possible the detection and timing of another photon. The detector can be actively or passively quenched. Passive quenching is the simplest, and yet in many cases a perfectly adequate method of quenching a breakdown pulse. This method was employed in our experiments. The electronic arrangement of the passively quenched diode is schematically shown in Fig. 3.1, together with the equivalent circuit of the diode.

The detector is placed in series with a quenching resistor $R_q$, which serves as a current-limiting load resistor. For quenching purposes, we want the current-limiting load resistor, $R_q$, to be large, since large load resistor causes the current through the SPAD quickly drop below the latch current and force the diode into a nonconducting

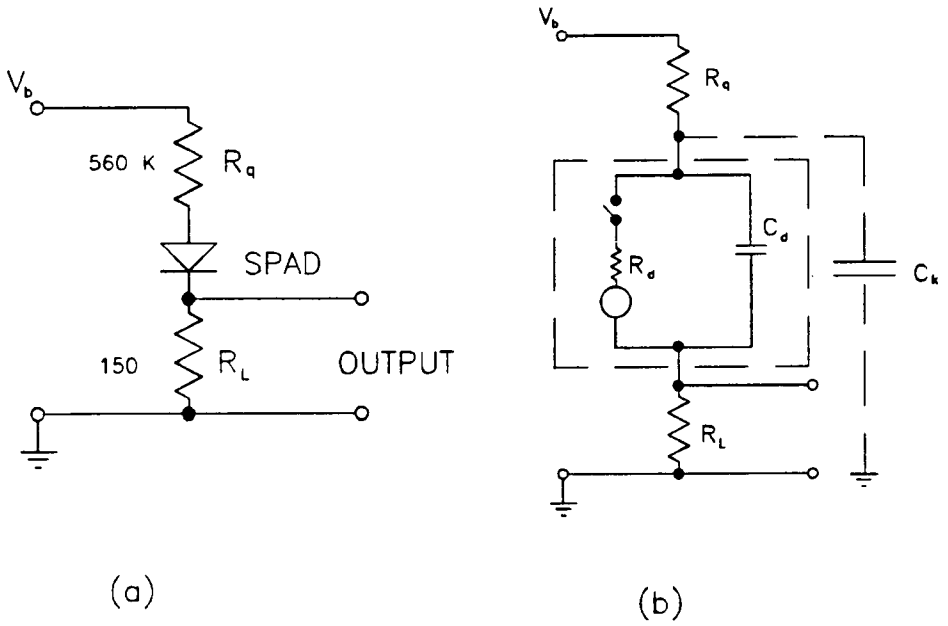



Figure 3.1: Electronic configuration of passive quenching of the SPAD

state. $C_d$ is the capacitance of the device and $C_k$ is the stray capacitance to ground of the diode terminal connected to $R_q$. The avalanche triggering corresponds to closing the switch in the equivalent circuit. The avalanche current discharges the capacitances, the diode voltage drops, and the diode current correspondingly decreases. When the voltage gets close to the breakdown value, the avalanche is quenched, corresponding to opening the switch in the equivalent circuit. The diode voltage then slowly rises towards the bias voltage as the capacitances are recharged by the small current through $R_q$ with a time constant of $R_q(C_d + C_k)$. Passive quenching has the disadvantage of longer dead time and therefore inhibits the use of very high count rate, as discussed in §3.4. The disadvantage of the slow recovery can be avoided by using active quenching. The principle of the method is to force the quenching and resetting transitions to occur in short times by using a controlled low-impedance source for biasing the diode. The rise of the avalanche pulse is sensed by a fast comparator, whose output switches the bias voltage down to the quenching level.

After a short but well-defined dead time, the bias voltage is swiftly restored to the full operating level.

## 3.3 SPAD test results

The tested results on the SPAD have been published (see Reference [15] for details). It was found that tightly focusing the laser beam in the center of the active area of the SPAD is essential for acquiring both high quantum efficiency and good timing. Here we briefly summarize some of the main results.

### 3.3.1 Quantum efficiency and dark count rate

Quantum efficiency and dark count rate are two very important features of a photodetector. At a particular photon energy, these two parameters are functions of the applied bias voltage. Also the dark count rate depends on the temperature.

The dark count rate increases steadily with extra-bias, while the quantum efficiency shows a maximum at 25 $V$ above breakdown. At this bias value, the quantum efficiency of the detector is over 50% and the dark count rate reaches $\sim 80s^{-1}$. It should be noted that it is possible to obtain detectors with lower dark count rate from the manufacturer. Also a small improvement could possibly be made by lowering the temperature further provided that after-pulsing is negligible.

### 3.3.2 Pulse height distribution

Because the avalanche process is statistical in nature, the resultant gain is also statistical. The pulse height distribution (PHD) of the SPAD will affect the rise time, and hence the timing. A CFD is used for triggering the incoming signal. To acquire the PHD, we recorded the decrease in the count rate, averaged over 1 minute, from a stable light source as the discriminator threshold level is incrementally raised. The PHD is then calculated from the derivative. It is found that the relative width of the PHD becomes narrow as the bias voltage increases, particularly for bias voltages

just above breakdown.

### 3.3.3 SPAD timing

For time-gated detection, it is desirable to use detectors of fast timing, since this allows the start of the time-window to be set earlier thereby increasing the throughput of the time filter. Detector timing does not signify how short the rise time and the width of the detector output pulse is. Instead, it is the jitter in the time interval between the laser excitation pulse and the detector output.

The time-jitter, otherwise known as the impulse response or simply "the prompt", is measured with a conventional TCSPC set-up by collecting a histogram of the time intervals, as discussed in §2.6. The pulse width from the dye laser at 585 $nm$ is $< 4$ $ps$ FWHM and so does not contribute significantly to the prompt. To minimize additional time jitter introduced by the electronics, the input voltage threshold and the zero-crossing voltage or "z-walk" of the CFDs are carefully adjusted by a trial and error process. In this manner, by tightly focusing the incoming light onto the active area, a quantum efficiency of over 50% and a single photon timing jitter of 168 $ps$ FWHM are achieved. The FW(1/100)M is 790 $ps$, which is comparable to that from a MCP-PMT. For these measurements, the time calibration is accomplished using an EG&G Ortec 462 time calibrator.

Note that the 168 $ps$ FWHM timing is measured by focusing the laser beam directly onto the SPAD, generating a spot size of $\sim 16$ $\mu m$. However, in the SMD experiment, the size of the spot on the SPAD is determined by the diameter of the spatial filter through which the fluorescence light passes. For a spatial filter with a diameter of $800\mu m$ and a 10 $\times$ microscope objective as the focusing lens, the spot size will be $80\mu m$. The larger spot size degrades the timing somewhat, giving a prompt with a FWHM value of $\sim 250$ $ps$.

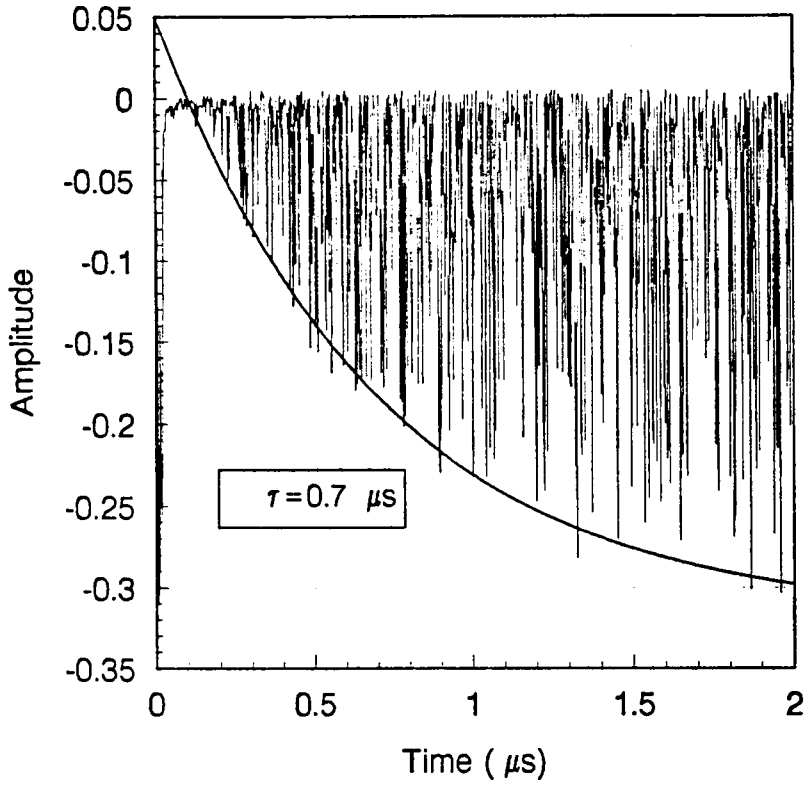



Figure 3.2: SPAD dead-time measurement

## 3.4 Problems related to SPAD dead time

Once an avalanche and subsequent quenching occurs, it takes time to recharge the SPAD, as discussed in §3.2.4, and consequently there is a period for which the SPAD is unresponsive or dead. Fig. 3.2 shows that the SPAD is completely unresponsive for about $\sim$ 200 $ns$ and has reduced response as the bias voltage recovers with a time constant of $\sim$ 700 $ns$, as discussed in section D of reference [15]. Thus a photon arriving before the SPAD bias voltage has fully recovered will, if detected, generate a pulse with a lower amplitude than normal. Furthermore, a longer time is required for the avalanche to develop due to the lower bias. Thus photons which are detected during the recoveries from previous photon events will be timed with additional delays and the time resolution will correspondingly be degraded.

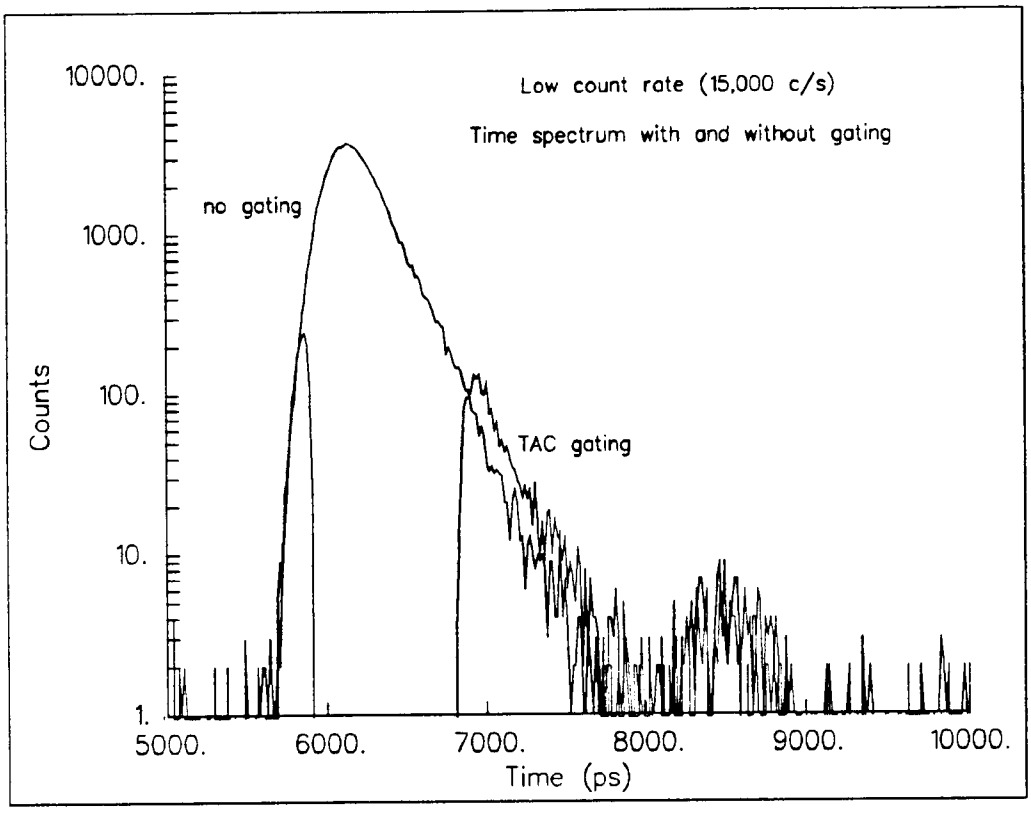



Figure 3.3: Time spectra–low count rate with and without gating

Because the dead time of the SPAD is less than that of the TAC, photons detected during the bias recovery periods are not usually processed. However, as mentioned at the end of §2.9, the SPAD dead time may cause problems when the temporal prefilter is used. Since most of the SPAD pulses originate from prompt Raman scattering, the temporal prefilter effectively eliminates pile-up at the TAC by removing prompt photons thereby allowing the TAC to analyze more delayed photons due to fluorescence.

At low count rates, both the SPAD and the TAC dead-time have negligible effects. Time spectra with and without the prefilter differ only in the existence and disappearance of the prompt peak, as seen in Fig. 3.3. When the two spectra are scaled in order to match the leading edges, the tails are essentially the same.

At high count rates, when the prefilter is applied, apart from the disappearance of the peak of the prompt, the tail of the spectrum increases significantly as seen in

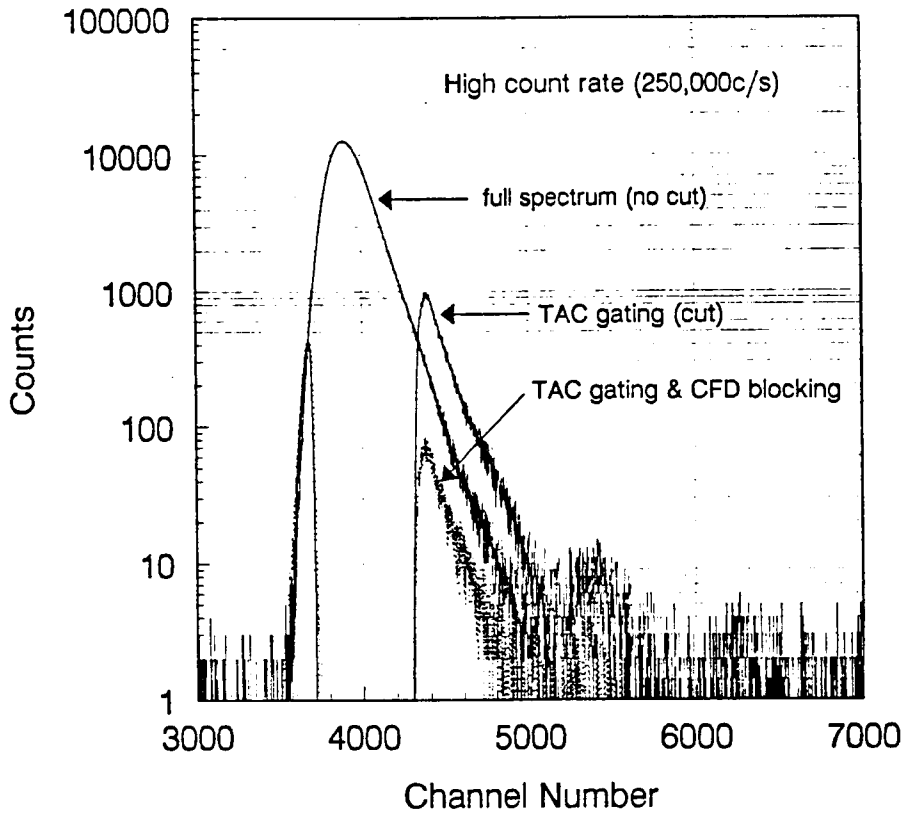



Figure 3.4: Time spectra–high count rate with gating and CFD blocking

Fig. 3.4. It seems that some of the counts eliminated from the peak by the prefilter reappear at later times.

The increase in the tail is thought to be due to the SPAD dead-time. In experiments at high count rates, although most photons are due to promptly scattered Raman, those that occur during the SPAD recovery time will appear delayed, as if they were due to fluorescence photons. When the TAC prefilter is not activated, photons that occur during the recovery time are not processed, even at high count rates, because the TAC dead time is larger than the recovery time.

The delayed Raman photons that give rise to the increase in the tail cause a significant increase in the background count rate. This can be reduced by increasing the lower limit of the time window, but then there is a loss in true fluorescence

counts. Instead, the blocking width of the CFD that accepts the SPAD pulses (CFD C in Fig. 2.10) is increased to $\sim$ 700 $ns$. This causes photons which are detected during the SPAD bias recovery time to be blocked by the CFD. Consequently the tail of the prompt is reduced as shown in Fig. 3.4.

Note that increasing the CFD blocking width does not reduce the SPAD dead time. While most of the photons that occur during the recovery time are due to Raman, those that are due to fluorescence are also lost. This causes a loss in fluorescence signal during photon bursts, when the instantaneous count rate is high. From the results of §8.4, the rate of fluorescence photons during a burst is seen to be typically $\sim$ 20 photons per 1 $ms$, i.e., $4 \times 10^4\ s^{-1}$, while the rate of Raman counts is typically $\sim 3 \times 10^5\ s^{-1}$. Using Eqn. 2.2, for a SPAD dead time of 0.7 $\mu s$, the loss of fluorescence signal is found to be $\sim$ 20%.

Future experiments should use an actively quenched bias circuit for the SPAD to reduce the dead time and hence decrease the loss of fluorescence during photon bursts.

# Chapter 4

# Flow system

## 4.1 Design considerations

In previously published reports on SMD, the cross sectional area of the flow cell was considerably larger than that of the tightly focused laser beam that is used to define the probe volume. Consequently, only a very small fraction of the molecules which pass through the flow cell would also pass through the probe volume and be detected. Also, molecules that would pass near the edges of the laser beam would give smaller amplitude photon bursts so that overall a wide monotonically decreasing distribution of burst amplitudes would be obtained. One of the goals of this work is to design and implement a flow system in which the sample is delivered in such a way that all molecules pass through the center of the laser beam at a flow velocity that will permit SMD. If this can be realized, high detection efficiency and a distribution of burst amplitudes with a well defined peak should be expected.

There are several considerations in designing the flow system in order to achieve the above goal including flow cell size and shape, size of the probe volume, methods for delivering the sample such that it follows a narrow stream and the corresponding flow velocities.

The flow cell size and shape should enable the maximum amount of fluorescence light to be collected. A very important factor that must be considered in choosing a flow cell is the short working distances ($< 0.5\ mm$) for high numerical aperture

microscope collection objectives (see §5.5.3). As for the shape of the flow cell, since a cylindrical capillary cell would introduce serious image aberrations, a square-bore capillary cell must be used. To cause all molecules to pass through the laser beam and be detected, it is not possible to use a flow cell smaller than the laser beam, since this would no longer allow the spatial filter to effectively eliminate specular laser scattering and fluorescence from the cell walls. The flow cell must be large enough for the cell walls to be far from the probe volume, but on the other hand it must be small enough to accommodate the small working distance of the collection objective.

The probe volume is the region in which molecules are efficiently excited and light is efficiently collected. It is determined by the overlap of the tightly focused laser with the field of view of the spatial filter, provided that the collection objective collects light efficiently from throughout this region. However, if the collection objective has a shallow depth of focus, which is usually the case for high numerical aperture objectives, then efficient light collection through the spatial filter will occur only over a shallow depth, as discussed in §5.5.3. In this case, the depth of the probe volume may be determined by the effective depth of collection (DOC) of the objective rather than by the depth of the focused laser beam, as is indicated in Fig. 4.1.

In order to reduce noise due to Rayleigh and Raman scattering from solvent molecules, the probe volume should be as small as possible. Typically the probe volume is $\sim 1.7\ pl$ and hence contains $\sim 6 \times 10^{13}$ water molecules. Thus if there happens to be *one* single dye molecule inside the probe volume at a particular time, there would also be $\sim 10^{14}$ water molecules present. Although Raman scattering has a low cross section compared to absorption and subsequent fluorescence emission, it is additive and thus the huge number of Raman scatterers inside the probe volume produce an enormous background. Similarly, for a fixed laser intensity, Rayleigh scattering also increases linearly with the probe volume.

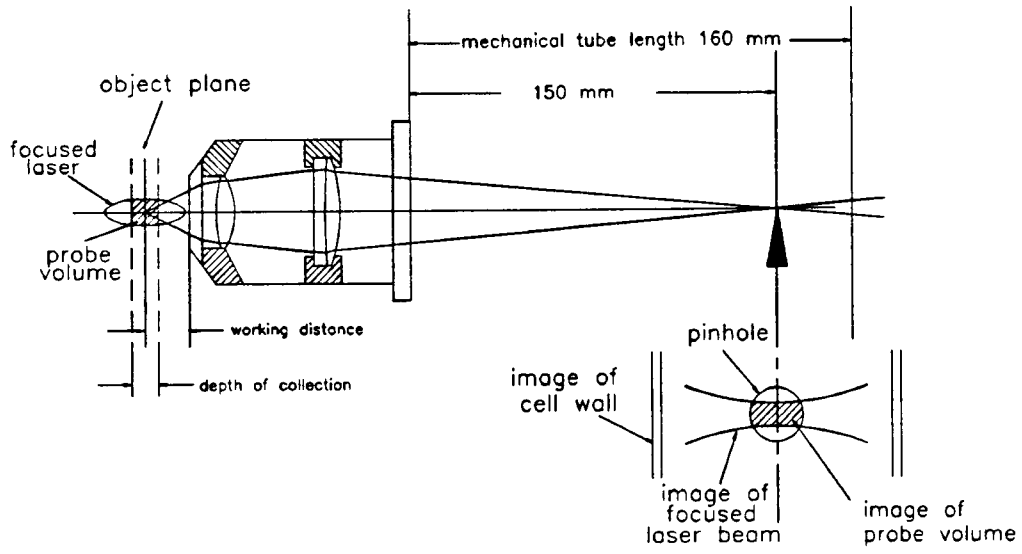



Figure 4.1: Probe volume and collection options

In order to detect all molecules within a sample, the sample stream size must be smaller than both the laser beam and the DOC of the objective. Fluorescence from molecules outside the DOC will appear blurred at the spatial filter and will result in smaller amplitude bursts leading to an overall poor distribution of burst amplitudes.

The next consideration in designing the flow system is that the narrow sample stream cannot be obtained with the use of very high flow velocities, since otherwise the molecule transit time will be too short for SMD. Researchers working with laser-induced fluorescence have been using hydrodynamic focusing, well-known in flow cytometry, to provide a narrow and well-defined sample stream. Although hydrodynamic focusing can yield sample stream sizes of typically $\sim 20\mu m$ in diameter, it does this by significantly increasing the sample flow velocity to overcome the diffusion of sample molecules into the surrounding sheath flow, as discussed in §4.4. This makes the molecule transit time across the laser beam very short. Consequently, there will not be enough fluorescence photons extracted from the molecule for

detection, since the number of cycles the molecule is excited by the pulsed laser decreases linearly as the transit time is decreased.

In order to overcome the above difficulties, a new diffusion-limited sample delivery scheme that gives both a narrow sample stream and a sufficiently slow flow speed is implemented. The sample is delivered from a capillary with a fine opening at the tip, which is positioned near the center of the flow cell, directly above the laser beam. The tip acts essentially as a point source and the molecules pass through the laser beam before diffusional spreading becomes significant, as explained in §4.5 and 4.6.

As a preliminary step to achieving SMD using this scheme, initial experiments were conducted under similar conditions within a simple capillary flow cell. §4.2, 4.3 and 4.4 review laminar flow, the simple flow cell and hydrodynamic focusing. In §4.5, the diffusion-limited sample delivery scheme is described. Calculations on the diffusional spreading of the molecules using a simplified model are presented in §4.6. In §4.7 the differences between the simplified model and the real flow cell are discussed and a qualitative account of the expected flow profile is given. Some experimental results on the sample stream size are given in §4.8.

## 4.2 Laminar flow

Flow in capillary tubes is classified as either laminar or turbulent. Laminar flow is characterized by non-mixing of flow lines inside the flow system. If a dye marker is threaded into the flow system, it will remain undisturbed as a thin dye line within the stream, without significant mixing (i.e., only mixing due to diffusion occurs). In flow cytometry and similar applications such as detection of single molecules in solution, the suspending medium should transport the particles along well-defined paths through the probe volume and thus only laminar flow is desirable.

Whether the flow is laminar or turbulent is indicated by the dimensionless Reynolds number, $R_e$, defined as

$$R_e = \frac{d\overline{v}\rho}{\eta}, \tag{4.1}$$

where $d$ is the diameter of the tube, $\overline{v}$ is the mean flow velocity, $\rho$ is the fluid density and $\eta$ is the viscosity. Above the critical Reynolds number, $R_{ekr} = 2300$, laminar flow is sometimes possible but is highly unstable and any small perturbation will cause the flow to become turbulent. Below this value, the flow is always laminar.

Often Reynolds number is calculated in terms of the kinematic viscosity, $\nu$, which is the viscosity $\eta$ divided by the fluid density $\rho$. At room temperature (293 $K$), water has a viscosity of $\eta = 1.0 \times 10^{-3}\ kgm^{-1}s^{-1}$ (in c.g.s. units $1.0 \times 10^{-2}$ Poise) and a density of $\rho = 1.0 \times 10^{-3}\ kgm^{-3}$ and thus a kinematic viscosity of $\nu = 1.0 \times 10^{-6} m^2 s^{-1}$. For the flow cell used in this work, the diameter of the tube amounts to $d \approx 0.8\ mm$. For the maximum flow speed of $\overline{v} \sim 2 \times 10^{-2} ms^{-1}$, the Reynolds number is

$$R_e = \frac{d\overline{v}}{\nu} = 16,$$

which is considerably less than the critical Reynolds number 2300. In the reported experiments the flow therefore is laminar.

Due to viscous forces, real fluids form boundary layers at the flow cell walls. Fully developed laminar flow in a long straight cylindrical tube of radius $R$ results in a parabolic velocity profile. The flow velocity at a radial distance $r$ from the axis is given by

$$v(r) = 2\overline{v}\left(1 - \frac{r^2}{R^2}\right). \tag{4.2}$$

In Eqn. 4.7, $\overline{v}$ is the mean flow velocity, which is given by

$$\overline{v} = \frac{Q}{F}, \tag{4.3}$$

where $Q$ is the volumetric flow rate and $F$ is the cross sectional area of the tube. Note that the flow velocity along the axis ($r = 0$) is twice the mean velocity.

## 4.3 Simple flow cell

A cylindrical flow tube is not suitable for use in SMD experiments because the curved surface acts as a strong cylindrical lens which introduces large aberrations at the collection optics. To avoid such aberrations a flow cell with flat windows must be used. In the first experiment demonstrating SMD, B. Shera *et al.* [10] used a fused silica cuvette with inside dimensions of $1cm \times 0.4cm$. However, due to the small working distance of the collection objective, it was necessary to arrange for the probe volume to be very close to the cell wall, by focusing the laser beam near the edge of the cell. Since the flow velocity varies steeply near the cell wall, the molecular transit time differs for different regions of the probe volume. In order to place the probe volume at the center of the flow cell where the velocity does not change much with position, the cross sectional area of the flow cell must be small enough to accommodate the small working distance of the microscope objective, as discussed in §5.5.3.

To meet these requirements, the simple flow cell used in the experiments reported in this dissertation consists of a $\sim$ 5 $cm$ length of square-bore borosilicate glass capillary, with inside dimensions (I.D.) of 0.8 $mm \times 0.8$ $mm$ and a wall thickness of 0.16 $mm$. The flow cell (part number WST-8280) is purchased from Wilmad Glass (Buena, NJ). The capillary I.D. tolerance is specified as $\pm 10\%$. The walls of the microcell are not grounded to be optically flat but are formed by a molding process. Consequently the optical quality is somewhat poor, although adequate for use in the SMD experiments.

The sample solution is directed to and from the flow cell through Tygon tubing (Thomas Scientific 9561-583) which is stretched over the ends of the flow cell to make a tight connection. The flow cell is mounted on a translation stage to allow the adjustment of its position so that the focused laser beam is at the center of the cell.

Since the flow cell has a square cross section, the velocity profile is no longer

given by Eqn. 4.7 and the velocity at the center is no longer $2\overline{v}$. The fully developed velocity distribution under laminar flow conditions for a rectangular duct of cross-sectional dimensions $x = 2a$ and $y = 2b$ is given by Lundgren *et al.*[19] as

$$v(x,y) = \frac{3ab^2\overline{v}}{2ab^4 - 12\sum N_n^{-5}\tanh N_n a}\left[b^2 - y^2 + \frac{4}{b}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{N_n^3}\frac{\cosh N_n x \cosh N_n y}{\cosh N_n a}\right],$$

where $N_n = (2n+1)\pi/2b$. The origin of the coordinates is at the center of the duct. For a square duct with $a = b$, the velocity at the center (x=y=0) is found to be approximately 2.09 $\overline{v}$.

It is possible to obtain an estimate of the flow velocity at the center by measuring the volumetric flow rate and using Eqn. 4.3 and the above factor of 2.09, although the uncertainty in the I.D. dimensions and the fact that the square cross section of the capillary has rounded corners would cause the percentage error of the estimate to be $\sim 25\%$. In SMD experiments, the important parameter to know is the transit time of a molecule through the probe volume, which is the focused laser beam diameter divided by the flow velocity. This parameter is more accurately deduced from the width of the autocorrelation function, as discussed in §6.5.1.

## 4.4 Hydrodynamic focusing

In the simple flow cell, the probe volume consists of only a small region at the center of the cross section and consequently only a small fraction of the molecules pass through the probe volume. One possible way to cause all molecules to be confined to a narrow stream smaller than the probe volume is to use hydrodynamic focusing, which is commonly used in flow cytometry for the detection of cells. In this type of flow system, the sample solution is confined within a sheath flow as it passes through the probe volume. The flow system is designed in such a way that the sample stream is focused by the sheath flow.

The principle of hydrodynamic focusing can be explained Fig.4.2, which shows

the flow from a large container into a small diameter tube of radius $R$. According to potential flow theory, all flows are described by flow velocity lines and pressure equipotential lines that are orthogonal to each other. For an incompressible fluid, conservation of mass requires that the amount of fluid crossing a hemi-spherical equipotential of radius $s$ in the large container must equal the amount of fluid crossing an equipotential surface in the tube, so that

$$v_s 2\pi s^2 = \overline{v}\pi R^2,$$

where $v_s$ is the flow velocity at the hemi-spherical potential and $\overline{v}$ is the mean flow velocity in the tube. For laminar flow, the large area, $A$, in the reservoir transforms into a small area, $a$, within the tube. If the velocity profile at the entrance of the tube is uniform the ratio of the areas is given by

$$\frac{a}{A} = \frac{R^2}{2s^2}.$$

In this way, streams of large area are focused into narrow threads.

The usual configuration used for hydrodynamic focusing is illustrated in Fig. 4.3. The sample is introduced through a small capillary at the center of the sample cell into a rapidly flowing sheath flow. As the sample stream accelerates to attain the same velocity as the sheath flow, the cross sectional area of the sample stream decreases. If the volumetric flow rate of the sample is $Q1$ and that of the sheath flow is $Q2$, then downstream from the input capillary the mean flow velocity is

$$\overline{v} = \frac{Q1 + Q2}{\pi R_2^2},$$

where $R_2$ is the effective radius of the cuvette. Eventually a parabolic velocity profile, given by

$$v(r) = 2\frac{(Q_1 + Q_2)}{\pi R_2^2}\left(1 - \frac{r^2}{R_2^2}\right), \tag{4.4}$$

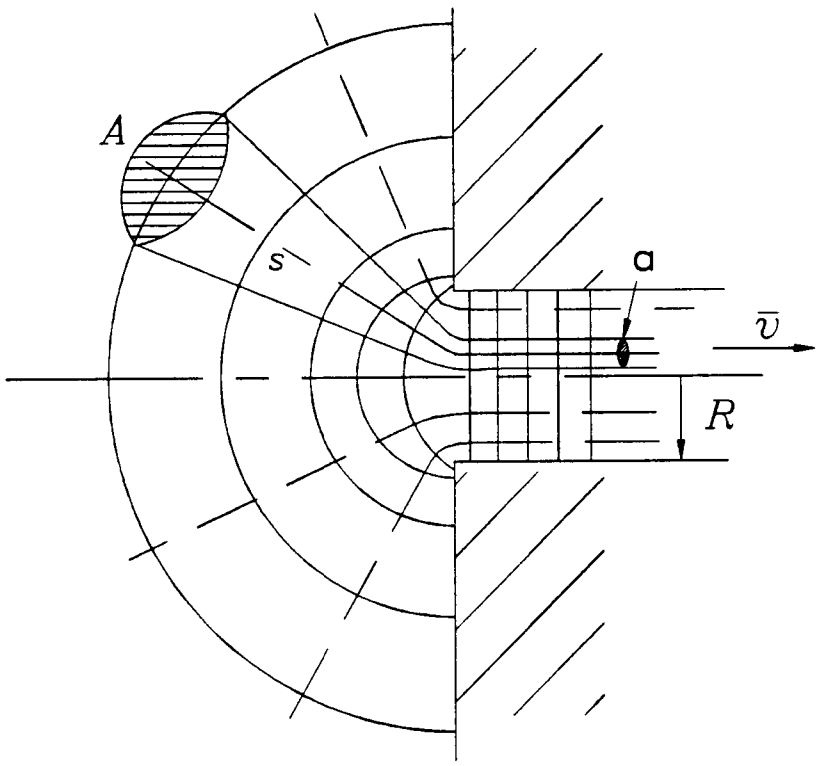



Figure 4.2: Principle of hydrodynamic focusing

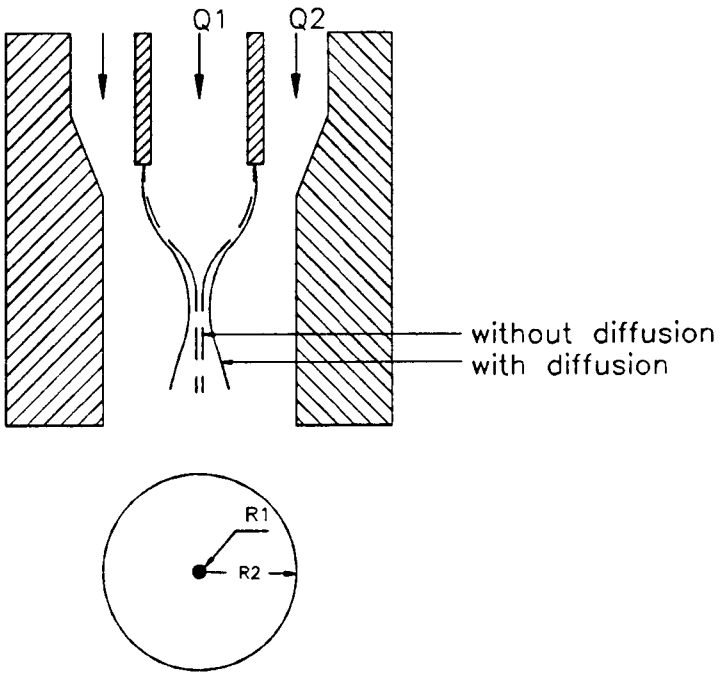



Figure 4.3: Practical hydrodynamic focusing configuration

develops. Due to conservation of volumetric flow, the radius of the hydrodynamically focused sample stream, $R1$, is then found by solving

$$\int_0^{R_1} v(r) 2\pi r dr \;=\; Q_1. \tag{4.5}$$

As given by Zarrin *et al.* [20], the result is

$$R_1 = R_2 \sqrt{1 - \sqrt{1 - \frac{1}{1 + \frac{Q_2}{Q_1}}}}. \tag{4.6}$$

Following the parameters used by D.C. Nguyen *et al.* [8] for sample and sheath flow rates of $Q1 = 4.1 \times 10^{-5}$ cm$^3 s^{-1}$ and $Q2 = 3.9 \times 10^{-3} cm^3 s^{-1}$ and an effective sheath flow radius of $R_2 = 141\ \mu m$, a hydrodynamically focused sample stream of radius $R_1 \approx 10\ \mu m$ is obtained, with a flow speed of about 126 $\mu m/ms$. By comparison, the flow velocity used in the first reported SMD experiments was about 0.24 $\mu m/ms$ [10], which is substantially smaller.

It may seem from Eqn. 4.6 that the same hydrodynamic focusing can be obtained at slower velocities by simultaneously reducing $Q1$ and $Q2$ with the ratio $Q2/Q1$ constant, but this is true only if one ignores diffusion.

Diffusional spreading is determined by the magnitude of the diffusion coefficient $D$. For Einstein-Stokes diffusion of a sphere of radius $r$ through a liquid of viscosity $\eta$ and temperature $T$, the diffusion coefficient is given by

$$D = \frac{kT}{6\pi\eta r}, \tag{4.7}$$

where $k$ is Boltzmann's constant. To a reasonable approximation, dye molecules in solution are surrounded by a spherical hydration region and obey Einstein-Stokes diffusion. The radius of the sphere varies approximately as the cube root of the molecular weight. Large molecules and particles such as cells have considerably smaller diffusion coefficients than single chromophore dye molecules and therefore diffusion is not a serious problem in flow cytometry. Furthermore, the signal from large fluorescently stained cells is considerably stronger so that faster flow velocities may be used.

For single chromophore molecules, the diffusion coefficient in water at room temperature is approximately that of sucrose, i.e., $D \sim 4.5 \times 10^{-6} cm^2 s^{-1}$. However, in general the exact value for specific molecules is not well known and it can vary by up to two to three times this value. [21]

In Fig. 4.3, the sample stream profile has been illustrated both in the case of no diffusion (dashed line) and with diffusion (solid line). A hydrodynamic model, when neglecting diffusion, predicts that the sample stream size will gradually decrease due to the acceleration in flow speed and will asymptotically attain a steady velocity and a constant radius. Diffusion becomes significant for low volumetric flow rates and for molecules with small molecular weight. With diffusion, the edges of the sample stream are no longer sharp but may be defined as that within which a certain fraction of the molecules are contained. The radius of the sample stream, which must be

calculated by including diffusion in the hydrodynamic equations, can be taken to be due to the sum of a hydrodynamic component and a diffusion component, i.e.,

$$r_{sample} = r_{hydrodynamic} + r_{diffusion}, \tag{4.8}$$

where the hydrodynamic radius is given by Eqn. 4.6.

By integrating the one-dimensional diffusion equation[22] it may be shown that the amount of diffusional spreading, $\delta$, after a time $t$ is given by

$$\delta = \sqrt{2Dt} \ . \tag{4.9}$$

Thus, as a rule of thumb, for a diffusion coefficient of $D = 5 \times 10^{-6} cm^2 s^{-1}$, diffusional spreading in each dimension occurs at an 'effective rate' of 1 micron per millisecond, $\sqrt{10}$ microns per 10 milliseconds, 10 microns per 100 milliseconds, etc.

For the cylindrical geometry, the diffusion-induced radius, $r_{diffusion}$, measured after some time $t$ from the capillary exit, also increases in proportion to $\sqrt{Dt}$. Thus to achieve small diffusional spreading, fast flow velocities must be used to reduce $t$. For example, if hydrodynamic focusing occurs $\sim 10^{-3}$ $m$ downstream from the inlet capillary, if $r_{diffusion}$ is to be less than $\sim \sqrt{10}$ $\mu m$, the mean flow velocity during the acceleration must be $\sim 10^{-1}$ $ms^{-1}$, which is about an order of magnitude faster than the velocities used in the flow system of §4.5.

In summary, although hydrodynamic focusing can provide a well-defined sample stream, the size is in practice limited by diffusion. In order to achieve a narrow sample stream, fast flow velocities must be used to overcome diffusion. However, single molecule detection cannot be achieved at such flow velocities, because the residence time in the tightly focused laser beam is not long enough to obtain an adequate fluorescence signal. These considerations led to the new flow system discussed below.

## 4.5 Diffusion-limited flow system

In the diffusion-limited flow system, the sample solution is simply delivered by a capillary with a very small hole at the tip, which is positioned at the center of the microcell and directly above the laser beam. Molecules exiting from the hole will undergo diffusional spreading as they are carried downstream through the laser beam by the sheath flow. If there were no diffusion, the molecules would simply follow the sheath flow and would all pass through the laser beam, provided that the hole of the capillary were smaller than the beam. With diffusion, the size of the sample stream at the beam must be determined from Eqn. 4.8. It is apparent from Eqn. 4.9 that for small diffusional spreading, the time $t$ that it takes for sample molecules to pass from the hole to the beam must be small. A small $t$ can be obtained by placing the capillary tip close to the beam and by using a fast sheath flow velocity. By using a capillary tip with a very small hole as close as possible to the probe volume, it is possible to achieve a sample stream narrower than the probe volume while using a flow velocity considerably slower than that required in hydrodynamic focusing.

Disposable, commercially available capillaries (Eppendorf, Femtotip), intended for microinjection of single cells, are used. An electron microscope photograph of the tip of a capillary, obtained from the product literature, is shown in Fig. 4.4. The capillaries are finely pulled to provide an opening at the tip of $0.5\ \mu m \pm 0.2\ \mu m$ in diameter. With such a small hole, it would take a long time to fill a capillary by drawing sample through the tip. Instead, capillaries are conveniently filled using a soft plastic pipette (Eppendorf, Microloader) which is inserted down into the capillary through the opening at the other end. This large end of the capillary is bonded to a plastic mount with screw thread that allows the capillary to be affixed to a tubular holder (supplied by Eppendorf). The holder is then mounted on a precision 3-dimensional translation stage (Newport 461) which allows the tip to be positioned at the center of the flow cell. The flow cell consists of the same square-bore microcell as discussed in the §4.3.

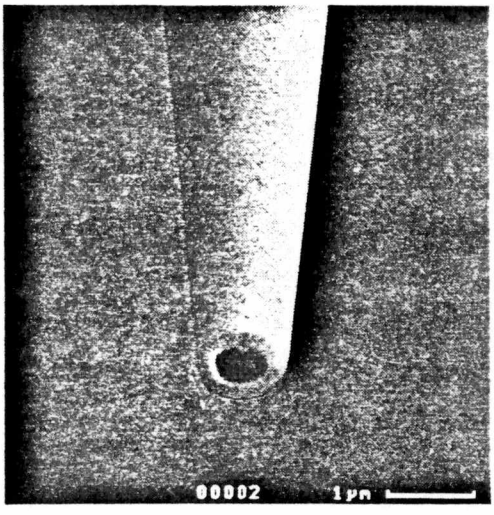



Figure 4.4: Eppendorf Femtotip capillary

Because the capillary is formed from tubing of outside diameter $\sim 1.4\ mm$, it is not possible to completely insert it into the square-bore microcell, which has I.D. $0.8\ mm \times 0.8\ mm$. Since the shoulder of the capillary near the tip is quite abrupt, it is only possible to insert it to a depth of $\sim 2\ mm$ into the microcell, as shown in Fig. 4.5.

Water for the sheath flow is injected into the opening of the microcell through a 33-gauge hypodermic needle placed at one corner of the top of the microcell. At the same time, the flow rate of water out of the microcell due to gravity is adjusted using a manometer in order to maintain a constant level of water at the top of the microcell.

Due to the physical size limitations among the cell, tip and needle, the needle is mounted on a micro-positioning mechanism so its position may be adjusted to avoid interference with the vertical movement of the capillary.

A syringe pump (Harvard Apparatus Infusion Pump Model 975) with a 50 $ml$ disposable syringe and 0.2 $\mu m$ particulate filter (Gelman Sciences, Acrodisc) is con-

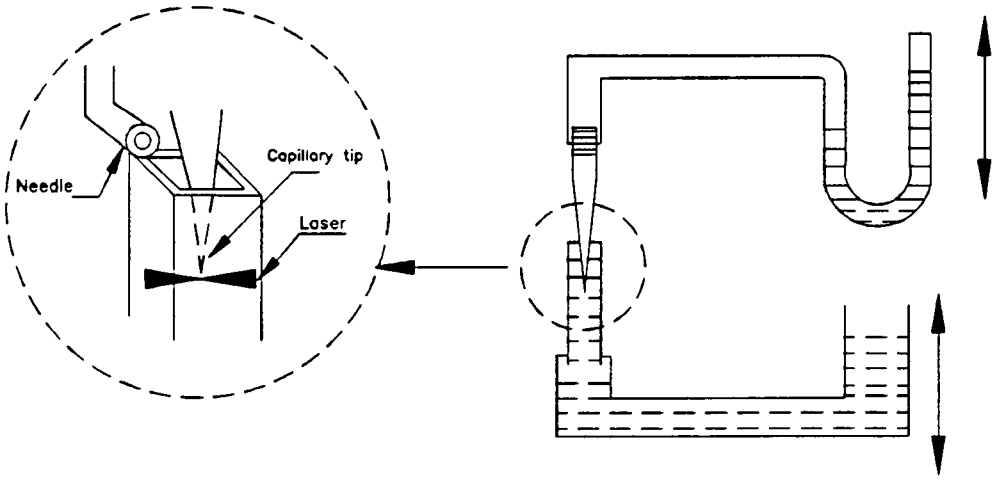



Figure 4.5: Flow cell set-up showing the sample capillary tip

nected to the hypodermic needle to control the volumetric flow of water into the microcell. The fine gauge hypodermic needle requires a high pressure of delivery from the syringe pump which was found to result in occasional surges in the flow. In later experiments (§8.4), it is replaced by a glass capillary with a larger tip glued into position without the micropositioning system. Even at the lower pressure, the syringe pump still produces slight surges in the flow rate. A gravity based delivery system could not conveniently deliver sufficient pressure to allow the use of the particulate filter. Consequently, a bottle pressurized by helium gas was used to provide the sheath flow in the later experiments.

The input side of the capillary is connected via Tygon tubing (Thomas Scientific) to a water filled manometer. The gravity balanced water level pressure from the manometer allows the sample delivery rate to be adjusted.

The set up of the sample cell, including adjustment of the water level and position of the capillary tip, is monitored through a microscope from the opposite side of the flow cell. For this, the flow cell is illuminated from the side by a fiber optic illuminator (Fiber-Lite Model 190, Dolan-Jenner Industries, Woburn, MA).

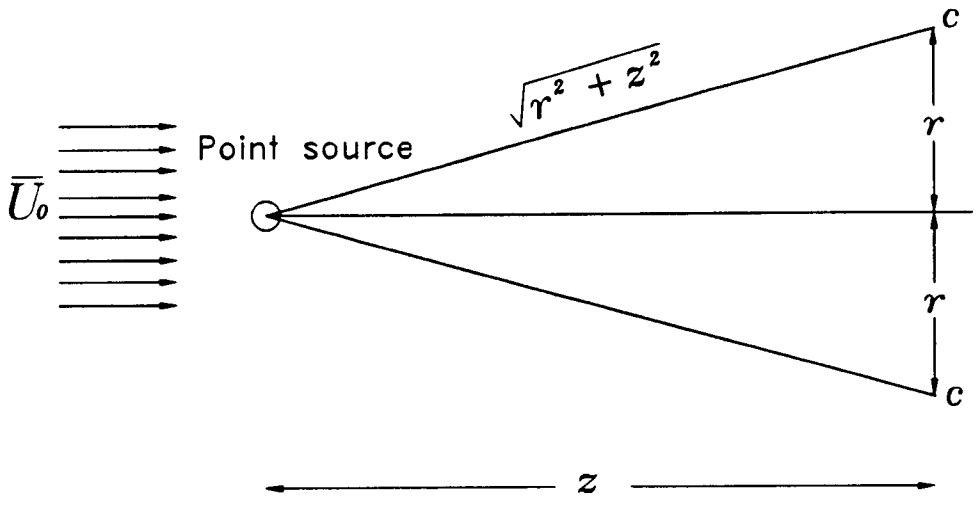



Figure 4.6: Diffusion from a point source in a uniform flow stream

In order to understand the performance of the new flow system, theoretical and experimental investigations are conducted, showing that the size of the diffusion-limited molecule trajectory can be smaller than the probe volume. The effect of diffusion of the dye molecules after they come out of the tip opening into the sheath flow is calculated using a simplified model presented in §4.6.

## 4.6 Calculations of diffusional spreading

Since the size of the tip opening is small compared to the size of the probe volume, it may be treated as a point source. Furthermore, as depicted in Fig. 4.6, the velocity of the sheath flow is assumed to be uniform and the volume of the sample introduced through the tip is negligibly small so that it does not perturb the uniform sheath flow.

This case has many applications for both molecular and particulate diffusion, as in studies of the spread of gases from tall stacks and smoke generators. It may

also be used to measure the diffusion coefficient $D_{AB}$ of molecules of type A into molecules of type B. For example, Walker and Westerberg[23] measured $D_{AB}$ in high temperature gases by observing the radial diffusion from a point source of one gas spreading into a slowly rising stream of a second. If material A is supplied from a point source at the origin into material B, flowing at a uniform velocity $\bar{U}_0$, the concentration as a function of radial position $r$ and position downstream from the source $z$ is given as [24]

$$c(r,z) = \frac{g}{4\pi D_{AB}\sqrt{z^2+r^2}} \exp\left[-\frac{\bar{U}_0}{2D_{AB}}(\sqrt{z^2+r^2}-z)\right]. \tag{4.10}$$

At a particular position $z$ the accumulation of concentration in the radial direction from $r = 0$ to $r = r_F$ is found by integrating Eqn. 4.10. Setting $z^2 + r^2 = \xi^2$, with $2rdr = 2\xi d\xi$, gives

$$\begin{aligned} 2\pi \int_0^{r_F} c(r,z) r dr &= \frac{g}{2D_{AB}} \int_z^{\sqrt{r^2+r_F^2}} \exp\left[-\frac{\bar{U}_0}{2D_{AB}}(\xi - z)\right] d\xi \\ &= -\frac{g}{\bar{U}_0}\left[\exp\left(-\frac{\bar{U}_0}{2D_{AB}}(\xi - z)\right)\right]_z^{\sqrt{z^2+r_F^2}} \\ &= \frac{g}{\bar{U}_0}\left[1 - \exp\left(-\frac{\bar{U}_0}{2D_{AB}}(\sqrt{z^2+r_F^2}-z)\right)\right]. \end{aligned} \tag{4.11}$$

Thus the fraction, $F$, of molecules which pass within a disk of radius $r_F$ located a distance $z$ down stream from the point source is given by

$$F = 1 - \exp\left[-\frac{\bar{U}_0}{2D_{AB}}(\sqrt{z^2+r_F^2}-z)\right]. \tag{4.12}$$

The radial contours within which a fraction, $F$, of the molecules is contained are found by solving Eqn. 4.12 for $r_F$. This yields

$$r_F = \sqrt{\alpha^2 + 2\alpha z}\ , \tag{4.13}$$

where

$$\alpha = -\frac{2D_{AB}}{\bar{U}_0}\ln(1-F).$$

Alternatively, expressing $z$ as a function of $r_F$ gives

$$z = \frac{r_F^2}{2\alpha} - \frac{\alpha}{2}. \tag{4.14}$$

Note that at $r_F = 0$, $z = -\alpha/2$, due to upstream diffusion.

Eqn. 4.14 defines a parabolic concentration contour of down stream position $z$ as a function of radial distance $r_F$, such that a fraction $F$ of the molecules will be confined inside the profile. This parabolic profile is shown in Fig. 4.7 (for different values of $F$), for a flow velocity of $\bar{U}_0 = 10\ \mu m/ms$ and a diffusion coefficient of $D_{AB} = 4.5 \times 10^{-6} cm^2 s^{-1}$. Note that diffusional spreading will be greater if a larger value is taken for $D$. In Fig. 4.7, the origin is shifted to the center of the focused laser beam, which is shown to scale at a distance of 15 $\mu m$ from the capillary tip. The coordinate system is such that the laser beam propagates along $x$ direction (into the page).

Although Fig. 4.7 indicates that $> 95\%$ of the molecules pass through the laser beam, it should be noted that the calculation is based upon assumptions which considerably simplify the real situation. Firstly, individual molecules do not follow smooth parabolic trajectories but may diffuse in and out of the laser beam. §7.5 gives a more sophisticated Monte Carlo simulation model, although again a uniform flow speed is assumed. Secondly, and more importantly, the flow profile near the end of the capillary tip is not uniform. A complete computational fluid dynamics calculation would be necessary to find the detailed flow profile. However, in §4.7 a qualitative discussion on the velocity profile near the tip guided by analytical results for simpler geometries is given.

## 4.7 Velocity profile

The physical presence of the tip changes the flow so that the velocity profile close to the tip is not uniform. As mentioned in the previous section, it is difficult to calculate the exact flow velocity of the molecule as it comes out of the tip and passes

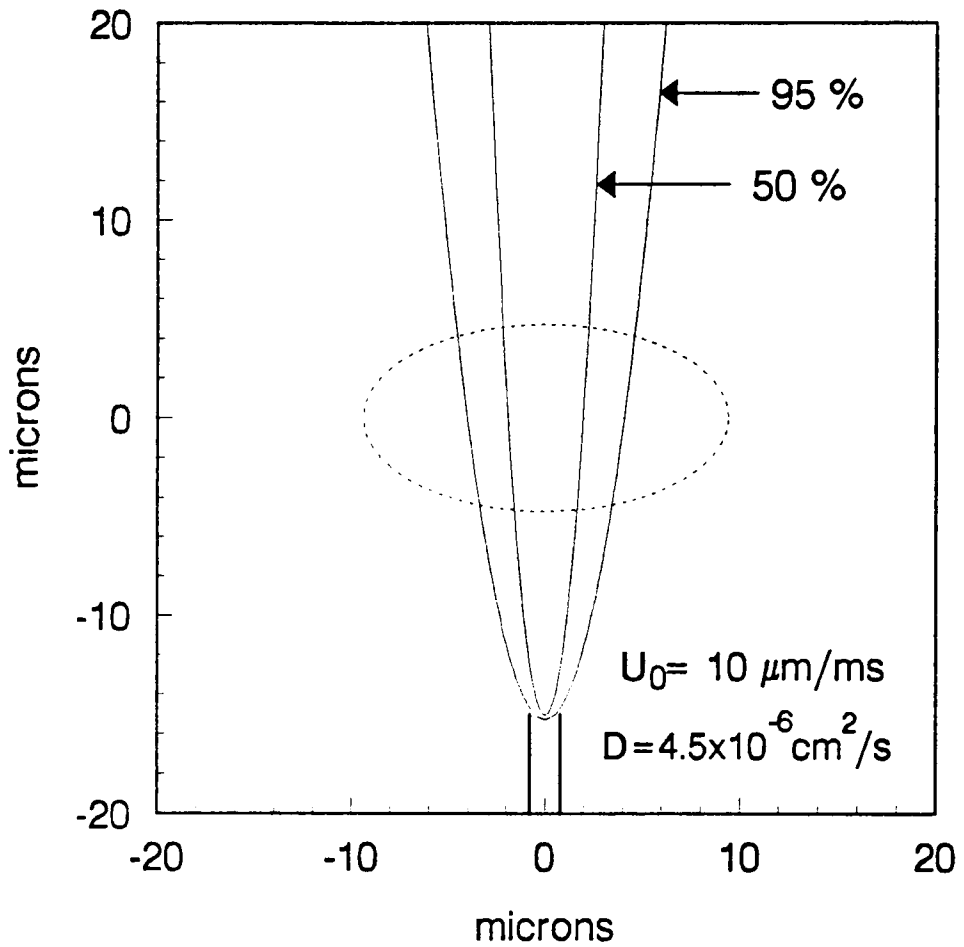



Figure 4.7: Parabolic concentration profile

through the probe volume. However, based on the analytical velocity profile result of a similar but simple geometry, a qualitative estimate of the real flow situation and the velocity profile near the tip can be made.

Upstream of the tip, due to the presence of the sample inlet capillary, the flow situation can be approximated to be that of a round annular tube. Lundgren *et al.* [19] give the analytical solution of a fully developed velocity in this case as

$$\frac{u}{\overline{v}} = 2\left[\frac{r_o^2 - r^2 - 2r_m^2 \ln(r_o/r)}{r_o^2 + r_i^2 - 2r_m^2}\right],$$

where

$$r_m^2 = \frac{r_o^2 - r_i^2}{2\ln(r_o/r_i)},$$

and $r_i$ and $r_o$ are the inner and outer radii of the annulus and $r$ is the radial coordinate. The position of maximum velocity and zero shear is denoted by $r_m$. For the flow cell used in the SMD experiments, if $r_0$ is taken to be $\sim 400\mu m$ and if the taper of the inlet capillary is ignored so that $r_i \sim 1\mu m$, then $r_m \sim 115\mu m$. Due to the boundary conditions, the flow velocity at the wall of the inlet capillary must be zero.

Downstream from the tip, after a distance sufficient for a fully developed flow to have formed, the velocity profile is can be approximated to be that of a round tube as given in Eqn. . In this case, the maximum velocity occurs at the axis. Therefore, immediately after the tip, there is a transition region in which the flow accelerates from almost zero to this maximum value. This behavior is depicted in Fig. 4.8.

Since the velocity near the tip is not uniform, the calculation of §4.6 is only valid in so far as the non-constant velocity can be approximated by an effective velocity of value $\bar{U}_0$. In the actual situation, greater diffusional spreading can be expected near the tip, where the flow velocity is less.

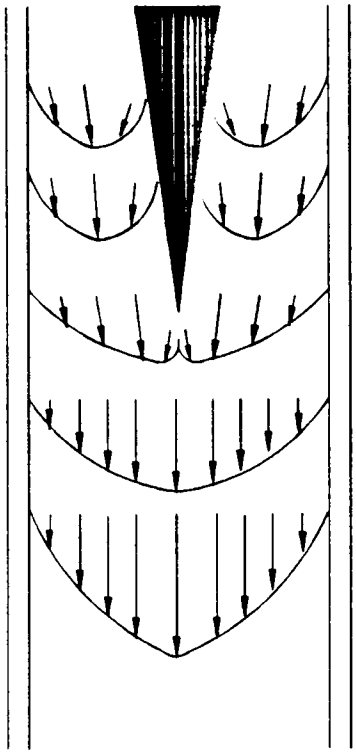


Figure 4.8: Qualitative flow velocity profile in flow cell

## 4.8 Measurements and interpretation

As pointed out in §4.1, the motivation for adopting the capillary tip and diffusion-limited flow system is to cause all or at least most of the molecules to pass through the probe volume and be detected. Due to the complex flow structure near the tip, the theoretical calculations alone do not provide convincing evidence that this occurs. Thus, in order to provide supporting evidence, an experimental estimation of the sample stream size is made.

In practice, it is impossible to accurately measure the sample stream size by simply imaging under high magnifiction. This is because a microscope objective with high magnification has a very shallow depth of focus, as explained in §5.5.3.

Nevertheless, images of the sample stream using a concentrated solution of R6G dye and with illumination by a round laser beam of waist $\omega_0 = 4.7\ \mu m$ (see §5.5.1)

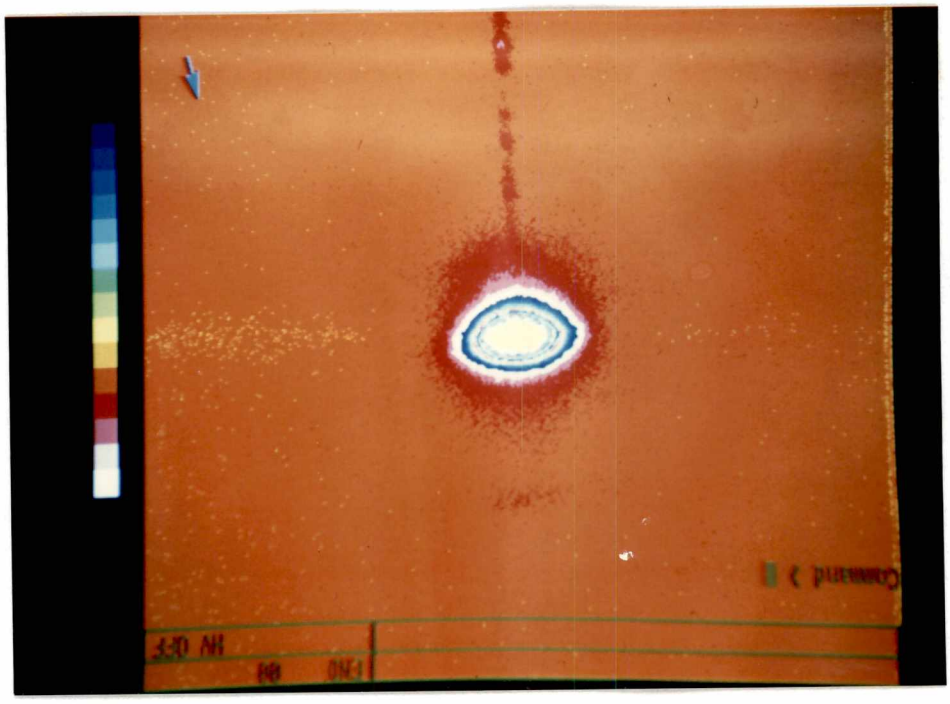


Figure 4.9: CCD image of concentrated dye stream illuminated by laser

are obtained with a 40 × microscope objective and an intensified CCD camera system (Hamamatsu, Argus 100). Fig. 4.9 shows such an image. Although the area of illumination appears elliptical in the horizontal direction, by changing the focus of the microscope slightly the image can be made to appear elliptical in the vertical direction. The depth of the sample stream is $\sim 10\ \mu m$ while the depth of focus is $\sim 0.35\ \mu m$. Regions that are out of focus also contribute to the image.

To obtain an estimate of the sample stream size, the integrated brightness of the image is measured as the sample stream is translated through the laser beam. The relative position of the sample stream is obtained from the differential micrometer of the translation stage that is used to mount the inlet capillary. The total brightness of the image is obtained using software of the CCD camera system.

Fig. 4.10 shows that the brightness versus tip position is approximately Gaussian in shape. The standard deviation of the Gaussian curve fit is 4.1 $\mu m$. This mea-

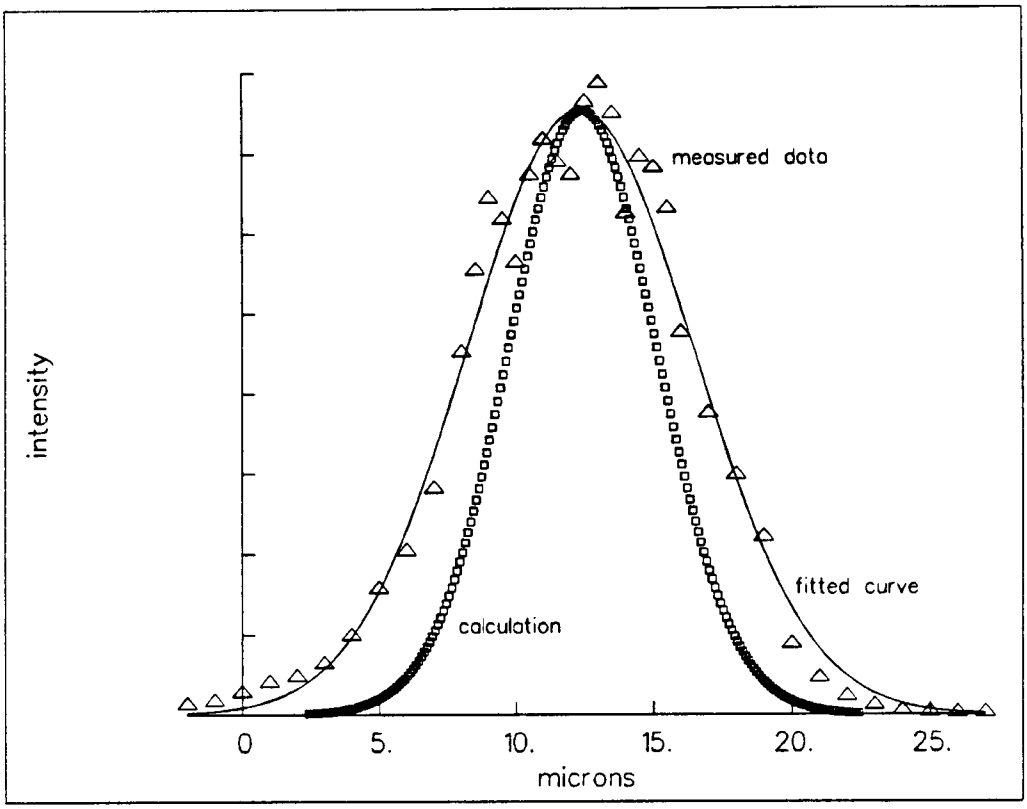



Figure 4.10: R6G sample stream size measurement

surement is obtained with a volumetric flow rate of $Q = 5.0 \times 10^{-9} m^3 s{-}1$, which corresponds to a mean flow velocity of $\bar{v} = 7.8 \times 10^{-3} m s^{-1}$. Note that this is smaller than the actual flow velocity used in the SMD experiment of §8.4 and so provides an upper limit to the diffusional spreading. The distance from the capillary tip to the center of the laser beam cannot be measured precisely but is estimated to be about $z_0 \sim 15 \mu m$. The major source of error in the measurement is caused by micron-sized fluctuations in the position of the tip during translation.

The width of the Gaussian curve is due to both the concentration profile and the Gaussian profile of the laser, which has a standard deviation of $\omega_0/2$, as in Eqn. 5.6. If the concentration profile were also taken to be Gaussian, it would have a standard deviation of $\sim \sqrt{4.1^2 - 2.35^2} = 3.4 \mu m$.

Fig. 4.10 also shows the expected brightness versus relative position of the laser beam and sample stream, $B(\Delta)$, calculated using the theoretical concentration pro-

file of §4.6. This is found by using numerical integration to evaluate

$$B(\Delta) = \int\int\int_{-\infty}^{\infty} I(x-\Delta, z)c(\sqrt{x^2+y^2}, z+z_0)dxdydz, \qquad (4.15)$$

where $c$ is obtained from Eqn. 4.10 using $\tilde{U}_0 = 4.4 \times 10^{-3} ms^{-1}$ and $D = 9 \times 10^{-6} cm^2 s^{-1}$ and the laser intensity profile $I$ is given in Eqn. 5.3.

The calculated curve is narrower than the result of the measurement. This may result from uncertainties in flow speed, the R6G diffusion coefficient and the spacing between the tip and laser. It may also be due to the non uniform flow conditions near the tip, as discussed in §4.7.

While there is still some uncertainty in the size of the sample stream, it is apparent that it is of comparable size to the depth of the probe volume, determined by the DOC of the microscope objective (see §5.5.3). The final test for whether the sample stream is smaller than the probe region is whether the distribution of burst amplitudes from single molecules has a well defined peak. In the experiments of §8.4 a well defined peak is obtained, indicating that most molecules do pass through the probe region.

# Chapter 5

# Experiment realization

In this chapter the experimental implementation of the single molecule detection experiments is discussed. At first, details are given on the photophysical properties of dye molecules used in the experiments and on the procedure used for sample preparation. An overview of the experimental set-up is then presented, followed by more detailed accounts of the lasers and the optics. Lastly, steps involved in setting up of the experiment are discussed.

## 5.1 Photophysical properties of dye molecules

In this work, SMD experiments were performed using aqueous solutions of three different dye molecules– rhodamine 6G (R6G) and B-Phycoerythrin (B-PE), using excitation at 532 $nm$, and sulforhodamine 101 (S-101), using excitation at 585 $nm$. Table 5.1 lists the various photophysical constants and other parameters of relevance for SMD experiments when the dye molecules are in water, as obtained from Refs. [25] and [26].

The initial experiments were conducted with the simple flow cell of §4.2.1 using R6G, which is commonly used as a laser dye, due to its high fluorescence quantum efficiency and photostability in ethanol and ethylene glycol. However at concentrations less than $\sim 10^{-4}$ $M$ in water, when dimerization is negligible, the fluorescence quantum efficiency is only 0.45. The absorption and emission spectra of R6G are

Table 5.1: Photophysical quantities of R6G, B-PE and S-101

| Molecule species | R6G | B-PE | S-101 |
|---|---|---|---|
| Absorption cross section ($\times 10^{-16} cm^2$) | 2.2 | 91.6 | 3.7 |
| Extinction ($cm^{-1}M^{-1}$) | 57,900 | 2,410,000 | 97,368 |
| Molecular weight($g/mol$) | 479.02 | 240,000 | 606.71 |
| Absorption maximum($nm$) | 530 | 545&565 | 585 |
| Emission maximum($nm$) | 550 | 575 | 620 |
| Fluorescence quantum yield | 0.45 | 0.98 | 0.35 |
| Fluorescence lifetime($ns$) | 3.0 | 2.1 | 4.0 |
| Photobleaching quantum efficiency ($\times 10^{-5}$) | 1.9 | 1.1 | 5.5 |

shown in Fig. 5.1. While the emission spectrum is that of the dye at low concentrations in water [11], the absorption spectrum is that reported in the dye catalog [27] using ethanol solvent, although the spectrum in water is expected to be quite similar. For the SMD experiments, the chloride salt of R6G, obtained from Exciton Chemical (Dayton, OH), is used to prepare the solutions.

In order to advance to experiments with the diffusion-limited sample delivery system described in §4.3, experiments were then conducted using the multichromophore molecule B-PE, since this allows SMD at higher flow rates and also reduces the effects of diffusion due to the large molecular weight, as explained in §4.4. B-PE is a member of the phycofluor class of fluorescent reagents, which are conjugates of intensely fluorescent alagal phycobiliproteins. B-PE has 34 bilin chromophores between which efficient energy transfer may occur, resulting in an antenna-like effect that gives rise to a high absorption coefficient and high fluorescence quantum yield. The fluorescence intensity of a molecule of phycoerythrin has been estimated to be more than 20-fold greater than that of a typical single chromophore molecule such as fluorescein. The red emission and large Stokes shift of phycoerythrin, shown in Fig. 5.2 [26], are also advantageous in discriminating fluorescence from Rayleigh scattering and background fluorescence. Furthermore, B-PE

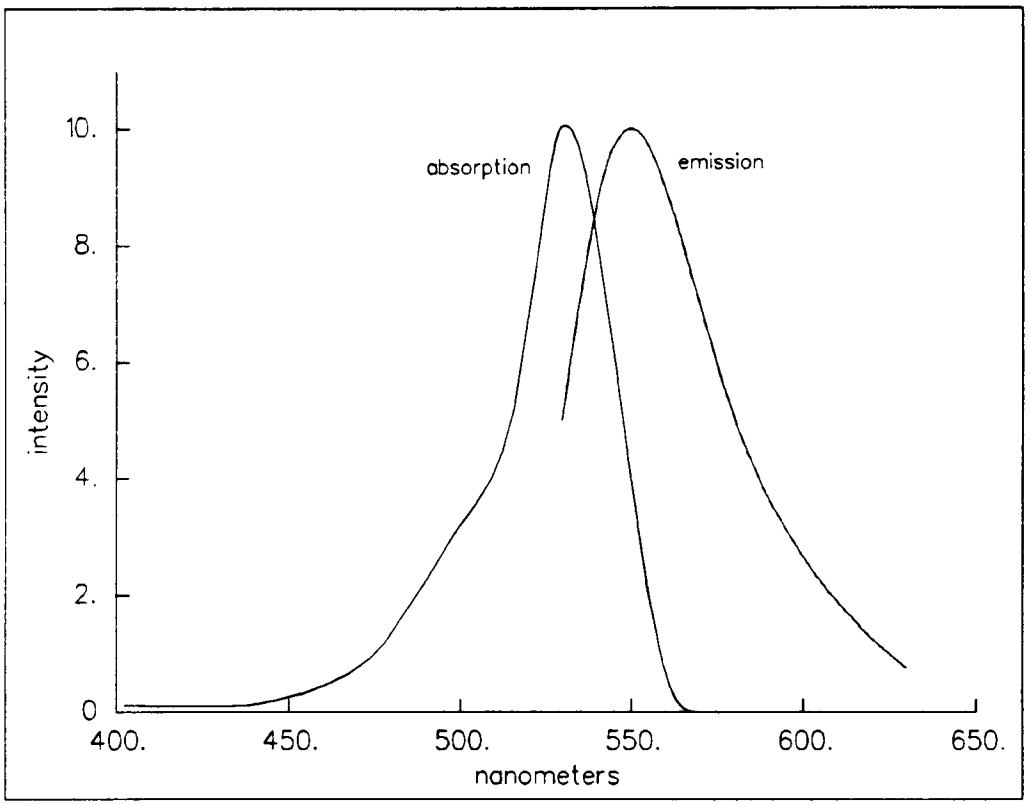



Figure 5.1: Absorption and emission spectra of R6G

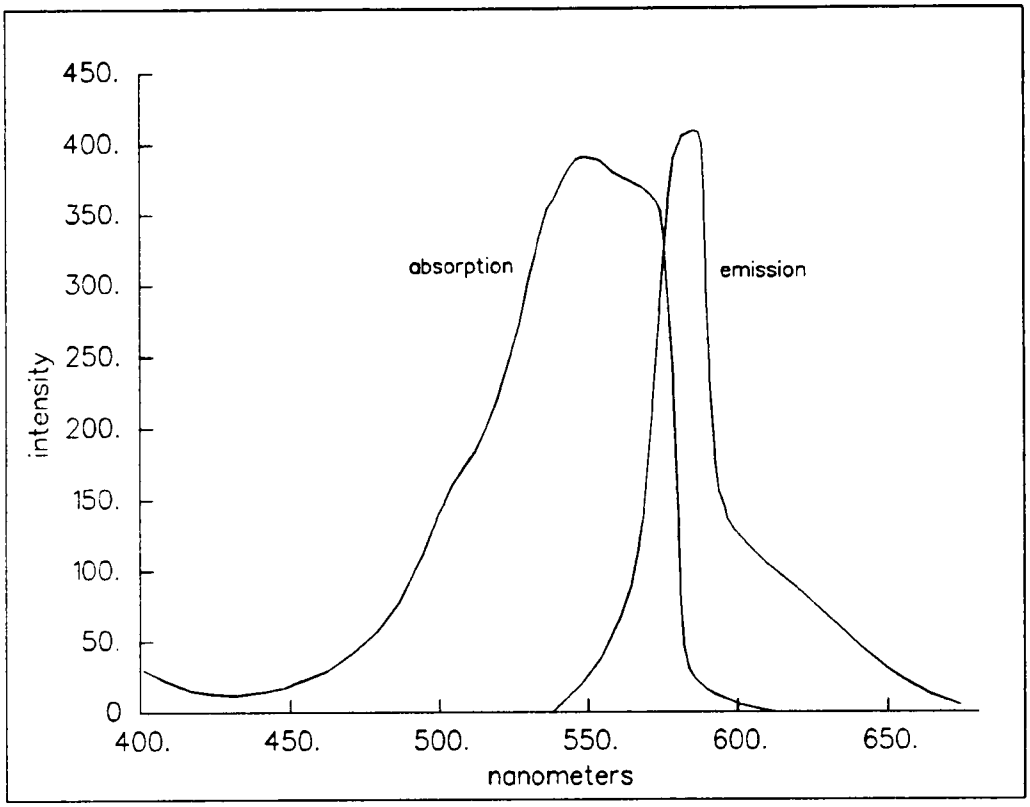



Figure 5.2: Absorption and emission spectra of B-PE

is stable on storage and its emission is constant over a broad range of pH and not quenched by most molecules in solution. For the SMD experiments, a 2 $mg$ sample of 4 $mg/ml$ solution in 60% ammonnium sulfate buffer (pH 7) was purchased from Molecular Probes (Eugene, OR).

After further improvements were made in the sensitivity of the SMD apparatus, experiments were attempted using R6G at higher flow velocities and higher laser excitation intensities both in the new diffusion-limited sample cell and in the simple flow cell. However, due to unidentified impurity molecules, even the ultra-clean water gave photon bursts under these conditions. As explained in § 8.1, a longer excitation wavelength (585 $nm$) was then used to avoid excitation of the impurities. This permitted experiments to be successfully conducted using the single chromophore molecule S-101 in both cells.

As may be noted in Table 5.1, the S-101 molecule exhibits a smaller fluorescence quantum efficiency and poorer photostability than R6G. The emission spectrum of S-101 in water [11] and the absorption spectrum in ethanol obtained from the dye catalog [27] are shown in Fig. 5.3. S-101 solutions are prepared from the solid powder, purchased from Sigma Chemicals (St. Louis, MO).

## 5.2 Ultradilute solution preparation

Ultradilute solutions of the dyes in triply distilled water, purchased from Cabisco Chemicals (Burlington, NC), are prepared by serial dilution of a stock solution of known concentration. In the case of R6G and S-101 the stock solution is made from the solid dry powder. In the case of B-PE, since the purchased solution contains solids, it is first passed through a sterile Acrodisc 0.2 $\mu m$ low protein binding non-pyrogenic filter (Gelman Sciences) in order to make the stock solution.

In order to minimize errors, the number of steps in the serial dilution process is kept to a minimum. For example, in order to prepare the ultradilute solution of R6G, 7.75±0.04 $mg$ of R6G is weighted into a clean weigh funnel and transferred to a

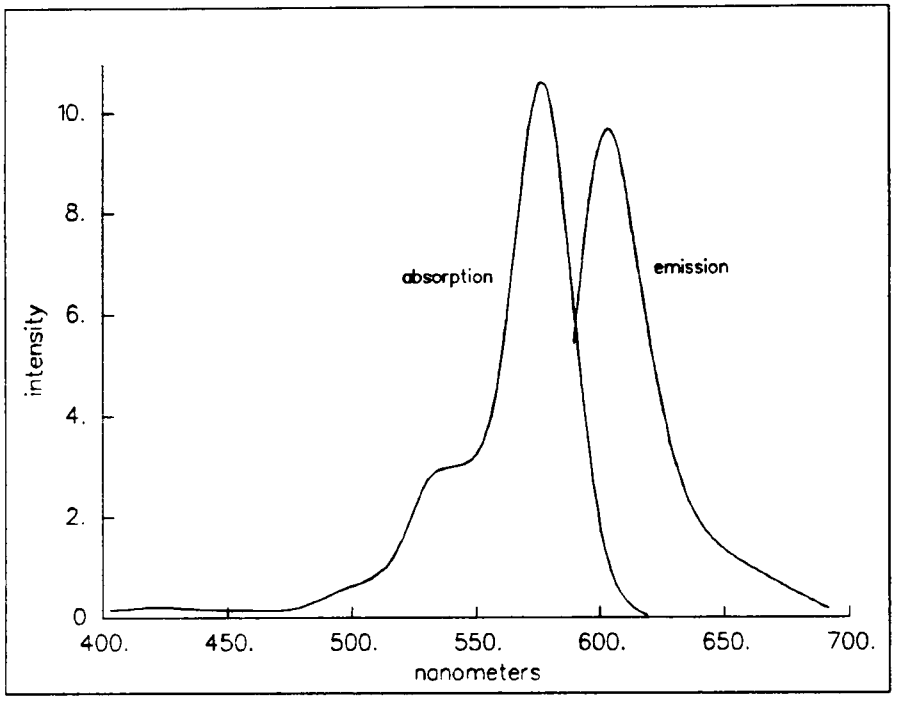



Figure 5.3: Absorption and emission spectra of S-101

100 $ml$ volumetric flask to create a R6G solution of the concentration of 1.618± 0.8% $\times 10^{-2}$ $M$. A 19.0 ±6% $\mu l$ sample from the above flask is transferred to another 100 $ml$ flask giving a dilution factor of $1.9 \times 10^{-4}$, resulting in a concentration of $3.074 \times 10^{-6} M$. This step is repeated twice in a similar fashion, to generate the final concentration of $1.0 \times 10^{-14} M$. The total random percentage error involved in the process is thus

$$\%\sigma = \sqrt{0.8^2 + 3 \times 6^2} = 10.4\%.$$

In the serial dilution process, care must be taken to avoid cross contamination by using clean new disposable micro-pipette tips. Nevertheless, it is quite possible that adsorption of dye molecules onto surfaces leads to systematic errors of comparable magnitude to the random error, particularly in the case of B-PE, for which the filtering process may alter the concentration of the stock solution. However, for the SMD experiments, the exact concentration of the ultradilute sample is not of

major importance, so long as the concentration is sufficiently dilute that the chance of having more than one dye molecule in the probe volume at any given time is negligible, as explained in §7.1.

On the other hand, contamination in the serial dilution process can lead to uncertainty as to whether detected photon bursts are due to the dye molecules of interest or to other objects, which could possibly be much brighter. To avoid outside contaminants, the flasks are thoroughly cleaned before use by rinsing them with ultraclean water until the rinse water shows no fluorescence bursts when processed by the single molecule detector. Furthermore, in order to test for the possibility of contamination in the process of serial dilution, a blank solution is prepared by performing the same serial dilution procedures as in the dye solution preparation, but beginning with ultraclean water. The result shows that contamination can be avoided if one takes extra care in handling the solution and flasks.

## 5.3 Overview of the experimental set-up

A schematic drawing of the experimental set-up used for the detection of single molecules is shown in Fig. 5.4. The system consists of a laser, focusing optics, a sample cuvette or other flow system, collection optics, spatial and spectral filters, a single photon detector, electronics modules for time gated photon counting and a personal computer for analysis of photon bursts and data storage.

The design of the system is guided by two points. Firstly, the system should be configured to maximize the detection of photons from the single molecules while minimizing the background noise. Secondly, the system should be simple, user-friendly and should use the minimum number of components. This guideline is the first step toward the development of a commercial instrument that requires low cost, reliability, stability and ease of use.

As discussed in § 5.4, a Coherent Antares Model 76-s frequency-doubled mode-locked Nd:YAG laser, producing 70 *ps* pulses at a wavelength of 532 *nm* and a

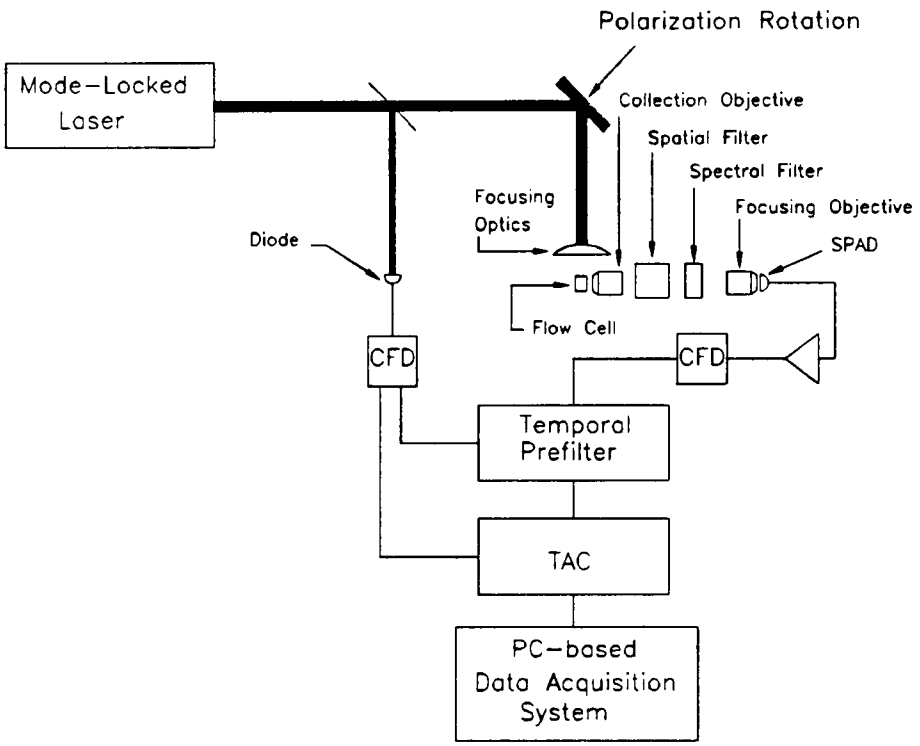



Figure 5.4: Experimental set-up

repetition rate of 76 $MHz$, is used as one of the excitation sources. For the detection of S-101 molecules, a Coherent 702-1 synchronously pumped picosecond dye laser, which operates at 585 $nm$ and produces pulses of a few picoseconds at 76 $MHz$, is used. The polarization of the output beams from each of these lasers is vertical, i.e., normal to the plane of the optical table. In order to minimize the Raman and Rayleigh scattering in the direction of observation (see Appendix B for details), the polarization direction of the excitation beam was rotated by 90° using a beam steering device. The laser excitation power is adjusted to an appropriate value (see § 7.6) by a series of neutral density filters. The laser power can be measured using a power meter (Laser Precision model RK–5720) placed before the focusing lens. The excitation beam is focused into the sample cell using the optics discussed in § 5.5. The flow cell, which is constructed from a glass square-bore capillary, as discussed in §4.2, is mounted in an upright (i.e., vertical) position.

Fluorescence light is collected at right angles, as discussed in § 5.5.3. A pinhole with a diameter of either 600 $\mu m$ or 800 $\mu m$, located at the image plane of the objective, serves as the spatial filter and limits the size of the probe volume, as discussed in § 5.5.4. After the spatial filter, there is one or more spectral filter, chosen to match the emission spectrum of the particular dye molecule under investigation, as discussed in § 5.5.5. The light from the pinhole is then imaged by a 10× microscope objective (Newport M-series) onto the SPAD detector. Both the collection objective and the SPAD detector are mounted on precision 3-dimensional translation stages (Newport 461, 462 series) for alignment.

The SPAD output signal is amplified by a $20dB$ gain, noninverting, $2.2GHz$ bandwidth amplifier (B&H Electronics AC2010H). The signal is then conditioned by a CFD (Tennelec TC454) and processed by the temporal prefilter discussed in § 2.9. The following electronics are basically those of a time-correlated single photon counting apparatus with a TAC (Ortec model 567) configured in reverse mode, as discussed in § 2.6. In the initial experiments, data acquisition and analysis are

accomplished by a 286 PC based multichannel scalar (MCS), as discussed in § 6.2. In later experiments, the TAC output pulses are digitized by a fast analog-to-digital converter (ADC), which passes the data to a 486 PC for analysis, as described in § 6.3.

## 5.4 Laser system

The desirable characteristics for a laser excitation source for use in SMD experiments include high repetition rate ($> 50MHz$) sub-nanosecond pulses, stable power in the range of $\sim 10mW$ and good beam quality for tight focusing.

A Coherent Antares 76-S mode-locked frequency doubled Nd:YAG laser, which provides these characteristics at a wavelength of 532 $nm$, is used for the detection of R6G and B-PE molecules. The absorption spectra of these molecules peak near 532 $nm$, as can be seen from Figs. 5.1 and 5.2.

The intracavity components of the Nd:YAG laser include the laser head containing the Nd:YAG rod and dual krypton arc lamps, an aperture for selecting the $TEM_{00}$ mode, a Brewster quartz plate to force linear polarization of the optical field inside the cavity and an acousto-optic mode-locker assembly. During the course of this dissertation, regular arc lamp replacement and several other major repairs were performed on the Nd:YAG laser. Leakage of light through the high-reflector is used to continually monitor the intracavity 1064 $nm$ pulses with a picosecond photodiode (Antel Optronics AR-S2) and a sampling oscilloscope (Tektronics S-4 sampling head and 7704A scope). The output beam is frequency doubled using a KTP doubling crystal to give $\sim 70$ $ps$ pulses at 532 $nm$ with an average power of $\sim 2$ $W$.

The output beam is split by a 50% dielectric beam-splitter and one branch is used as the pump source for a dye laser. The dye laser (Coherent 702-1), which can be tuned in wavelength from 560 $nm$ to 610 $nm$, is used to detect S-101 molecules. For these experiments, it is tuned to the peak absorbency of S-101 at 585 $nm$. Since the dye laser is synchronously pumped by the Nd:YAG laser it operates at the same

repetition rate. The dye laser can produce output pulses of $< 1ps$, but since the resolution of the time-gated detection is limited by the SPAD response to $> 100\ ps$, such short pulses were not necessary. Hence the dye laser was operated without the usual saturable absorber, yielding a pulse width of several picoseconds, as measured using a background free autocorrelator (Femtochrome Research Model FR-103).

## 5.5 Optical design

### 5.5.1 Laser focusing

Fig. 5.5 illustrates the set-up of the focusing optics.

The laser beam must be tightly focused into the sample cell with a beam waist of a few micrometers. Due to well corrected aberrations, high-quality microscope objectives are often the best off-the-shelf option not only for efficiently collecting light from a small region but also for focusing a laser beam to a small diffraction-limited spot. However, the small working distances of such lenses together with their physical size makes it impossible to place two objectives at right angles, each within the required distance of the flow cell. Consequently, a high numerical aperture objective is used only for light collection, while the laser is tightly focused into the cell using a low f-number achromatic doublet lens with a focal length of 47.625 $mm$ and a clear diameter of $\phi = 15mm$. Since the focused laser spot size is inversely proportional to the incoming beam diameter, the beam from each laser is expanded to a diameter of $\sim 6.0mm$ before it reaches the focusing lens. For the Nd:YAG laser, this is accomplished simply by using a long path between the laser output and the focusing lens. For the dye laser, a beam expander consisting of a $10\times$ microscope objective ($f = 14.8\ mm$) and an achromatic doublet lens ($f = 160\ mm$), indicated by L1 and L2 in Fig. 5.5, are used.

As discussed in §4.5, it is necessary to place the capillary tip as close as possible to the center of the laser beam so that the sample stream passes through the laser before diffusional spreading becomes significant. However, laser light scattering from

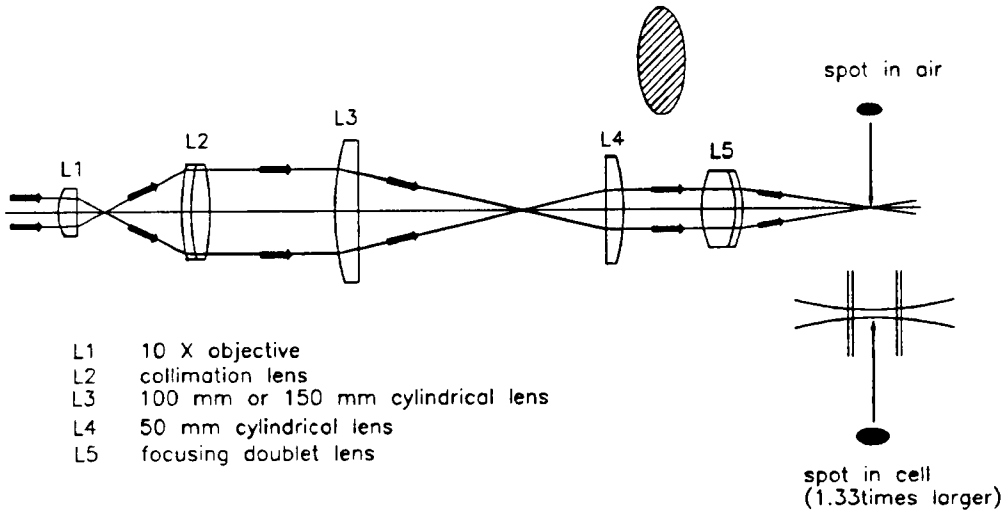



Figure 5.5: Focusing optics

the tip becomes intolerable if it is closer than about three times the beam waist. Therefore, an elliptical laser beam, narrower in the vertical direction along the axis of the flow cell (the $x$ direction), is used. This serves two purposes. Firstly, it allows the tip to be positioned closer to the center of the beam. Secondly, it allows the beam size along the light collection direction (the $y$ direction) to be larger than the depth of collection, thus ensuring strong molecule excitation throughout the probe volume, as discussed in §5.5.3.

To form an elliptical beam waist, a pair of cylindrical lenses is used to form a collimated elliptical spot before the focusing lens, as illustrated by L3 and L4 in Fig. 5.5. The cylindrical lenses increase the size of the waist in the horizontal $y$ direction without changing the waist size along the vertical $x$ direction. The focused elliptical beam waist has a horizontal to vertical size ratio of 3 to 1 throughout the experiment with B-PE. Later, guided by the computer simulations (see §7.6), this ratio is changed to 2 to 1 when detecting S-101 molecules.

### 5.5.2 Laser beam size

The laser beam size must be known in order to calculate the peak excitation intensity from the average power, to estimate how close the beam is to the capillary tip, to ensure that the beam waist in the $y$-dimension is sufficiently large, to calculate the volume of probe region and to estimate the flow velocity from the molecular transit time.

Without the cylindrical lenses in place, the beam waist is axially symmetric. Thus a measurement of the waist in the $y$ direction with the cylindrical lenses removed gives the size of the waist in the $x$ direction with or without cylindrical lenses. The size of the waist is measured only without the cylindrical lenses in place. The increase in size of the waist in the $y$ direction with the cylindrical lenses in place may then be calculated using the ratio of focal lengths of the cylindrical lenses.

To measure the laser beam waist, a simple method involving gradual eclipsing of the laser beam by a sharp knife-edge is used. As the knife-edge intersects the beam in a direction perpendicular to the propagation axis of the beam, a photodetector immediately behind the knife-edge measures the intensity of the unblocked portion of the beam. If the spatial profile of the laser beam is Gaussian, the signal measured by the detector is represented by an error function, as discussed below.

The cross-sectional spatial intensity profile of a Gaussian laser beam of power $P$ propagating in the z-direction is given by

$$I(r) = \frac{2P}{\pi\omega_0^2} \exp\left[-\frac{2r^2}{\omega_0^2}\right], \tag{5.1}$$

where $\omega_0$ is the beam waist. The intensity takes a maximum value of

$$I_{max} = \frac{2P}{\pi\omega_0^2} \tag{5.2}$$

on axis at $r = 0$ and falls to $1/e^2$ times this value at $r = \omega_0$. For a Gaussian beam of elliptical cross-section, the intensity profile consists of a product of Gaussians for each Cartesian coordinate $x$ and $y$, i.e.,

$$I(x,y) = \frac{2P}{\pi\omega_x\omega_y} \exp\left[-\frac{2x^2}{\omega_x^2}\right] \exp\left[-\frac{2y^2}{\omega_y^2}\right]. \tag{5.3}$$

If the beam size in the $x$ dimension, $\omega_x$ is the same as that in the $y$ dimension, $\omega_y$, Eqn. 5.3 reduces to Eqn. 5.1.

As the knife edge is translated along the $y$ axis, the detector signal due to the transmitted portion of the beam is given by

$$D(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} I(x',y')dx'dy'. \tag{5.4}$$

Substituting Eqn. 5.3 into the above equation and performing the integral over $x$ yields

$$D(y) = P \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{y'^2}{2\sigma^2}\right] dy', \tag{5.5}$$

where

$$\sigma = \omega_y/2. \tag{5.6}$$

The integral in Eqn. 5.5 is a normalized Gaussian error function. Thus, as the knife-edge is translated through the beam the detector signal is proportional to an error function with standard deviation equal to $\omega_y/2$.

A piezo-electric inchworm translator controlled by a Burleigh motion controller (Model 7000) is used to translate the knife-edge along the y-direction at an uniform speed. The error function waveform generated by the detector is captured on a digital oscilloscope (LeCroy 9400) and transferred to a PC through GPIB. The waveform is analyzed by a nonlinear curve fitting algorithm based on the down-hill simplex routine from "Numerical Recipes" [28]. From the fitted curve, the $\sigma$ parameter and hence the beam waist $\omega_y$ are deduced.

The measured scan of the dye laser beam at the focal plane without the cylindrical lens is shown in Fig. 5.6. The inchworm translator translation speed is $4.78\mu m/s$,

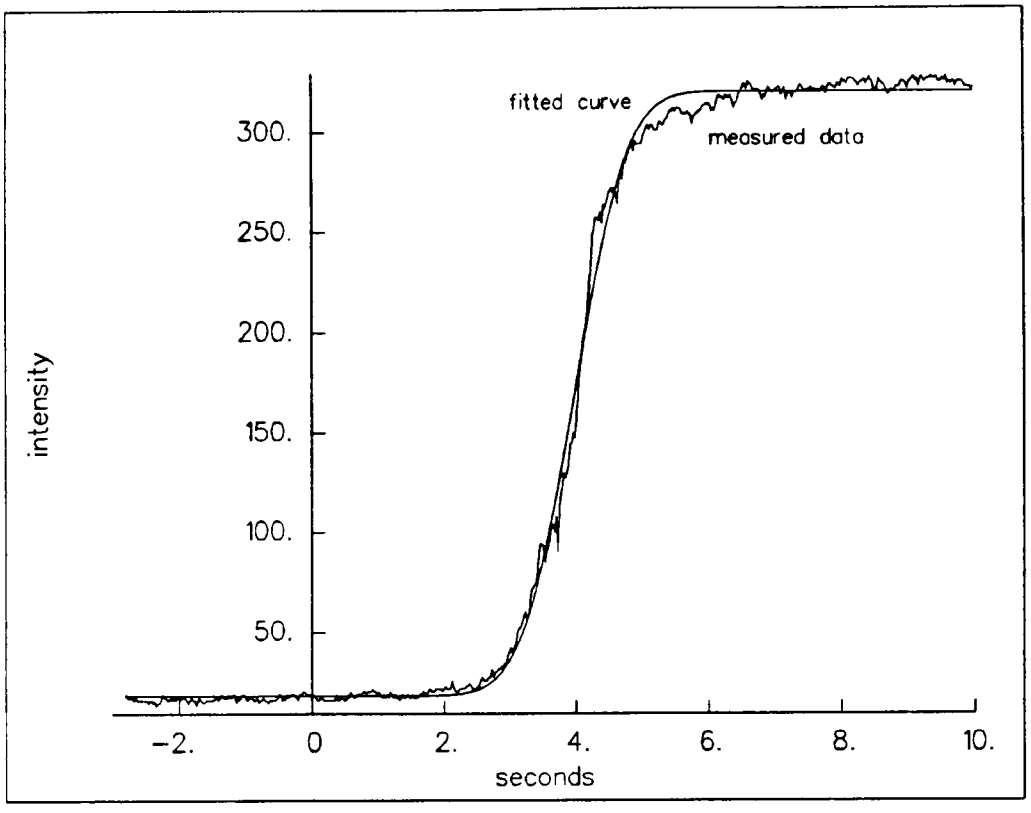



Figure 5.6: Laser beam waist measurement in air ($\lambda = 585\ nm$)

which yields a beam waist in air of $3.6\mu m$. The beam waist of the 532 $nm$ laser measured in the same way gives a waist of $3.5\mu m$.

In each case, the incoming beam at the focusing achromat is collimated and has a spot diameter of approximately 6 $mm$. Theoretically, a collimated beam of diameter $D$ will be focused by a lens of focal length $f$ down to a beam waist of

$$\omega_0 = \frac{2\lambda f}{\pi D}. \tag{5.7}$$

From Eqn. 5.7, using $D = 6\ mm$, $f = 47.6\ mm$ and $\lambda = 0.585\mu m$, the expected beam waist is $\omega_0 = 3\ \mu m$. The error in this theoretical value is about 25% due to the error in $D$. Nevertheless, the measured and theoretical values are very close. Error in the measured value, due largely to diffraction effects and uncertainty in the location of the focal plane, is estimated to be 10%.

The beam waist when the flow cell is in place will be different from that measured in air. Since the Rayleigh range of the focused beam is much smaller than half the width of the cell, geometric optics can be used to calculate the increase in the beam waist. For the 585 $nm$ beam, the Rayleigh range is $z_0 = \pi\omega_0/\lambda = 65\ \mu m$, while half the width of the cell is 400 $\mu m$. Due to refraction as the beam enters the cell, the beam spot size focused in the water is larger than that in air by a factor of 1.33, which is the ratio of refraction indices of water and air. This yields a beam waist of 4.8 $\mu m$.

In summary, the beam waists in the sample cell are $\omega_x = 4.7/4.8\ nm$ for the 532/585 $nm$ beams respectively and $\omega_y = 2$ or 3 times $\omega_x$.

### 5.5.3 Collection optics

Since the fluorescence intensity generated by a single molecule is very weak, it is necessary to use high numerical aperture optics to collect as much fluorescence as possible. Also, the design of the collection optics should prevent scattered light reaching the detector, even though it cannot be completely eliminated.

A microscope objective is used to collect the fluorescence. The collection efficiency, $\eta$, for an objective of numerical aperture N.A., is given by

$$\eta = sin^2\left[\frac{sin^{-1}(N.A./n)}{2}\right], \tag{5.8}$$

where $n$ is the index of refraction of the medium surrounding the object.

The collection efficiency versus numerical aperture in the case of water, for which $n = 1.33$, is plotted in Fig. 5.7. Although oil or water immersion objectives can have a numerical aperture greater than 1, most objectives have smaller values. In any event, the maximum light collection efficiency is 0.5, since the solid angle of collection is at most $2\pi$.

In the initial SMD experiments, a Newport M series 40× objective is used. This has $N.A. = 0.65$ and thus a collection efficiency of $\eta = 0.06$, as indicated by the

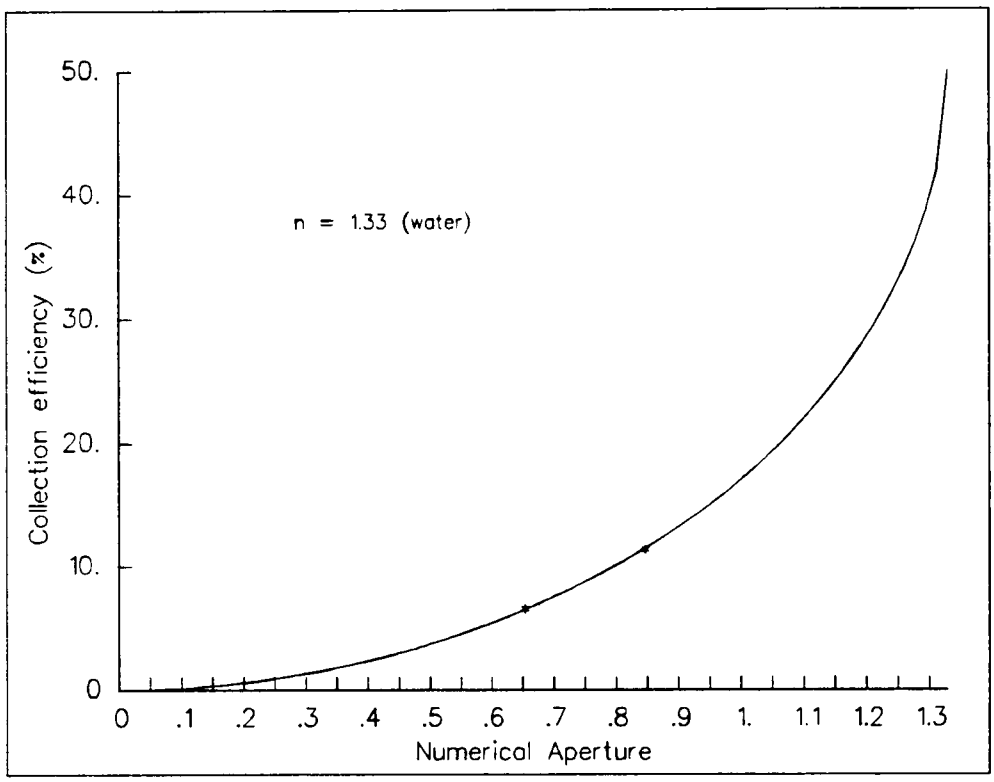



Figure 5.7: Collection efficiency versus numerical aperture

asterisk to the left in Fig. 5.7. Later experiments used a Nikon CF Plan Achromat (model 79173) 60× objective. This lens has 0.85 N.A. and therefore gives a collection efficiency of $\eta = 0.12$, indicated by the asterisk to the right in Fig. 5.7. The cone of light collection has an full angle of 79.4°, which can still be accommodated within the square-bore microcell, although the rounded corners of the microcell decrease the acceptance angle to less than 90° and hence may cause some light loss.

The depth of focus (DOF), which can be extremely shallow for a high N.A. objective, is an important concern in system design. It is approximately given by

$$DOF = \frac{\lambda}{4 \times N.A.^2}. \tag{5.9}$$

For $\lambda = 585\ nm$ and the 60 × objective (N.A.= 0.85), the DOF is about 0.2 $\mu m$. For pure imaging applications, such a shallow DOF means that an object translation of $\sim 0.2\ \mu m$ away from the object plane would result in a fuzzy image. However, for light collection, a larger amount of defocusing is tolerable, as long as the spatial filter does not significantly block the enlarged fuzzy image.

The depth of collection (DOC) of the 0.85 N.A. microscope objective is estimated by imaging laser scattering from a single $0.5\mu m$ diameter latex microbead, using an intensified CCD camera system (Hamamatsu Argus 100). The microbead is suspended in ethylene glycol to keep its position stable. The microscope objective is defocused by a small amount using the Burleigh piezo-electric inchworm translator and the degradation of the image is observed. By noting the fall-off in intensity contained within a disk of the same size as the pinhole used in the spatial filter, the DOC is estimated to be $\pm \approx 6\mu m$. This value is used in the Monte Carlo simulation of the experiment discussed in §7.6.

In addition to the light collection efficiency and DOF, there is another constraint that limits the choice of the objective– the collection objective must have a sufficiently large working distance to collect light from the center of the sample stream. If the distance from the center to the inside wall of the cuvette is $r$ and the thickness

Table 5.2: Working distances for different N.A. lenses

| Model | Focal length(mm) | Numerical aperture | Working distance(mm) |
|---|---|---|---|
| M-5× | 25.5 | 0.1 | 15 |
| M- 10× | 14.8 | 0.25 | 6.0 |
| M- 20× | 8.3 | 0.40 | 1.9 |
| M- 40× | 4.3 | 0.65 | 0.4 |
| M- 60× | 2.9 | 0.85 | 0.3 |

of the cell wall is $d$, then the working distance must be larger than $r/n + d/n_2$, where $n$ and $n_2$ are the refractive indices of the water and glass respectively. For $r = 400\ \mu m, d = 160\ \mu m$, $n = 1.33$ and $n_2 = 1.5$, the minimum working distance is $407\ \mu m$.

Table 5.2 lists the numerical apertures and working distances for the Newport M series microscope objectives. They are corrected for rear conjugates at a standard tube length of 160 $mm$. For the 40× objective used in the initial experiments, the working distance is not quite large enough. Consequently it is necessary to move the probe volume slightly off center, by focusing the laser closer to the cell wall.

Few lenses are designed with both a long working distance and a high numerical aperture because this requires large diameter lenses which add significantly to the cost. However, the Nikon 60× objective has a high numerical aperture as well as a working distance of 0.41–0.45 $mm$, which is slightly larger than the required minimum. Its focal length is 2.46 $mm$ and it is also corrected for a conjugate working distance of 160 $mm$.

Many objectives, especially those with higher numerical apertures, are designed for biological applications in which a thin (0.17 $mm$) cover glass is placed over the object being viewed. To compensate for this layer of glass, a slight negative spherical aberration term has been included in the objective design. When such high numerical aperture objectives are used without a cover glass or with a non-standard cover glass between the object and the objective, a slight deterioration of

image quality should be expected. The Nikon 60× objective provides an adjustable aberration correction for cover glass thicknesses from 0.11 to 0.23 $mm$. In the experiment, the correction was set at the maximum value, because both the cell wall of thickness 0.16 $mm$ and the 0.4 $mm$ depth of water contribute to spherical aberration.

### 5.5.4 Spatial filter

A pinhole placed at the image plane of the collection objective serves as a spatial filter or field stop to block light scattered from the capillary walls and from regions outside the probe volume. The size of the pinhole determines the diameter of the region from which light is efficiently collected.

For the simple flow cell of §4.3, light should be collected from a region slightly larger than the diameter of the focused laser in the vertical or $x$ direction. For example, in the experiment reported in §8.3, with a laser waist of $\omega_x = 4.8\ \mu$ and a 60 × microscope objective, the pinhole diameter should be $> 576\ \mu m$. However, in order to test whether SMD is achievable under conditions similar to those of §8.4, a 800 $\mu m$ pinhole is used.

For the diffusion-limited sample delivery system of §4.5, light should be collected from a region larger than the diameter of the sample stream, which in turn should be smaller than the DOC of the objective. For the experiments of §8.2, with an estimated sample stream radius of $< 6\ \mu$ and a 40 × objective, a pinhole of diameter $> 480\ \mu m$ must be used. A 600 $\mu m$ pinhole is used. For the experiments of §8.4, with an estimated sample stream radius of $\sim 6\ \mu$ and a 60 × objective, a pinhole of diameter $> 720\ \mu m$ must be used. A 800 $\mu m$ pinhole is used.

To ensure that the collected light passes through the spatial filter, detergent solution is passed through the flow cell and the image of the laser beam and probe volume at the pinhole is viewed through a view port with a magnification lens mounted at an angle to the optical axis. The view port is used in this way for

the initial optical alignment of the experiment. Also, when clean water passes through the cell, the image of light scatter from the capillary tip allows the tip to be positioned immediately above the probe volume.

### 5.5.5 Spectral filters

High quality spectral filters are used to effectively block the Rayleigh scattered light from the solvent while passing a large proportion of the weak fluorescence signal.

Since the Stokes shift of the fluorescence band of the dye molecules is not large, spectral filters which have a sharp cut on edge are required. The transition from blocking to transmission is not sufficiently sharp for common band pass filters. Instead, interference filters with a six cavity design are used to achieve the sharp edges on the transmission curve. The filters are custom designed by Omega Optical at affordable costs. They provide high optical blocking density at the excitation wavelength (532 $nm$ or 585 $nm$), i.e., transmission $T < 10^{-4}$, and a relatively high throughput over the fluorescence band, i.e., $T > 60\%$.

To maximize the signal, the filter's transmission band should cover the fluorescence emission spectrum of the dye molecule. However, this may result in considerable overlap with other sources of noise. After the Rayleigh scattered light is largely eliminated by the filter, Raman scatter from the solvent becomes the major source of background.

Fig. 5.8 gives the Raman spectrum of water[29]. There are two major bands, one at $\sim 1650 cm^{-1}$, corresponding to the symmetric bending intramolecular vibrational mode and the other extending from 3100 to 3700 $cm^{-1}$ and peaked at 3400 $cm^{-1}$, corresponding to the symmetric O-H stretching vibrational mode. Note that the Raman stretching band is much stronger than the bending band. For an excitation wavelength of 532 $nm$, the bending band is located at 583 $nm$, while the stretching band extends from 637 $nm$ to 662 $nm$. For an excitation wavelength of 585 $nm$, the bending band is located at 647 $nm$, while the stretching band extends from 715 $nm$

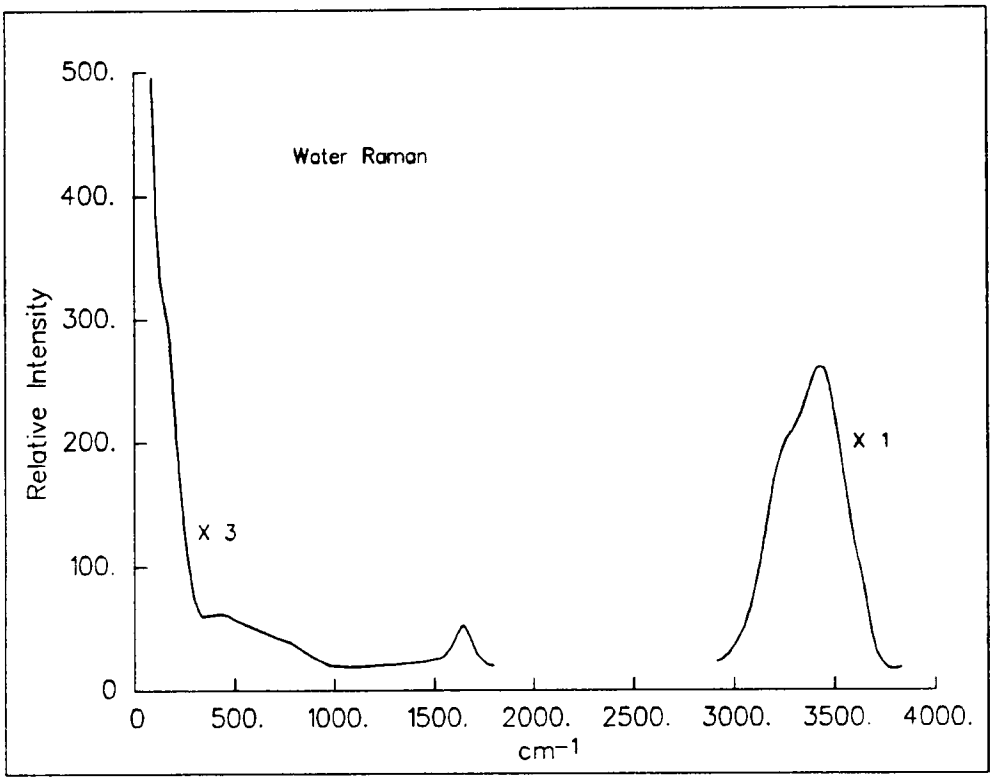



Figure 5.8: Water Raman bands

to 747 *nm*.

Although the Raman noise that passes the spectral filter is eventually discriminated by the temporal prefilter, if the rate of Raman counts is very high there will be a loss of signal due to the finite dead time of the detector, as discussed in §3.4. Therefore, the pass bands of the spectral filters are chosen to avoid collecting the Raman stretching band. This reduces the Raman counts considerably but it may also reduce the fluorescence signal. A trade-off factor has to be considered in choosing the pass band of the filter.

Also, the substrate of the filter should be chosen so that it is not intrinsically fluorescent. Otherwise Rayleigh scattered light will induce stray fluorescence that contributes to the background. Interference filters are usually constructed from absorptive color glasses. In setting up the experiments, intrinsic fluorescence was observed from two filters, which were replaced by filters constructed using a clear

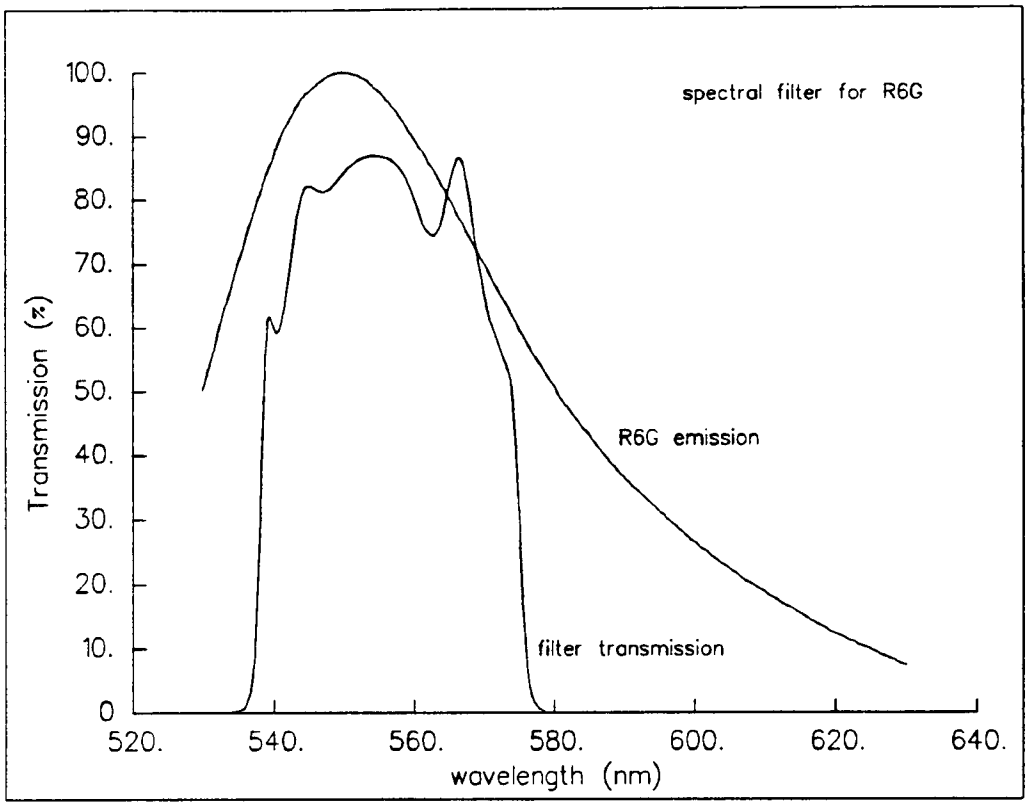



Figure 5.9: Transmission curve of the spectral filter for R6G

glass substrate.

Fig. 5.9 and Fig. 5.10 give the transmission curves of the spectral filters that are used for the experiments to detect R6G and B-PE respectively. For the detection of S-101 molecules, two band pass filters are used. The transmission curves of these two filters together with the emission spectrum of S-101 are shown in Fig. 5.11.

## 5.6 Single molecule detection experiment realization

In order to make the experiment successful, several factors need to be considered.

Firstly, laser power fluctuations should be minimized before running the experiment in order to maintain the stringent timing provided by the photodiode signal triggered by the laser pulse. Also, a high stability laser is necessary to achieve the

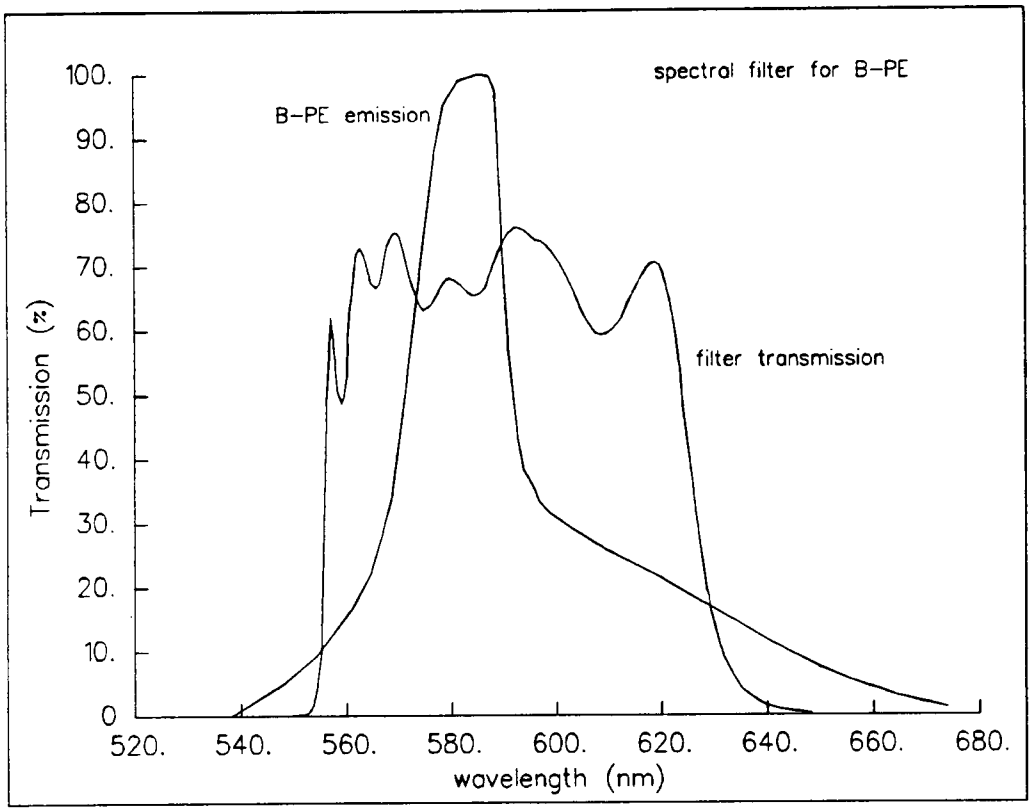



Figure 5.10: Transmission curve of the spectral filter for B-PE

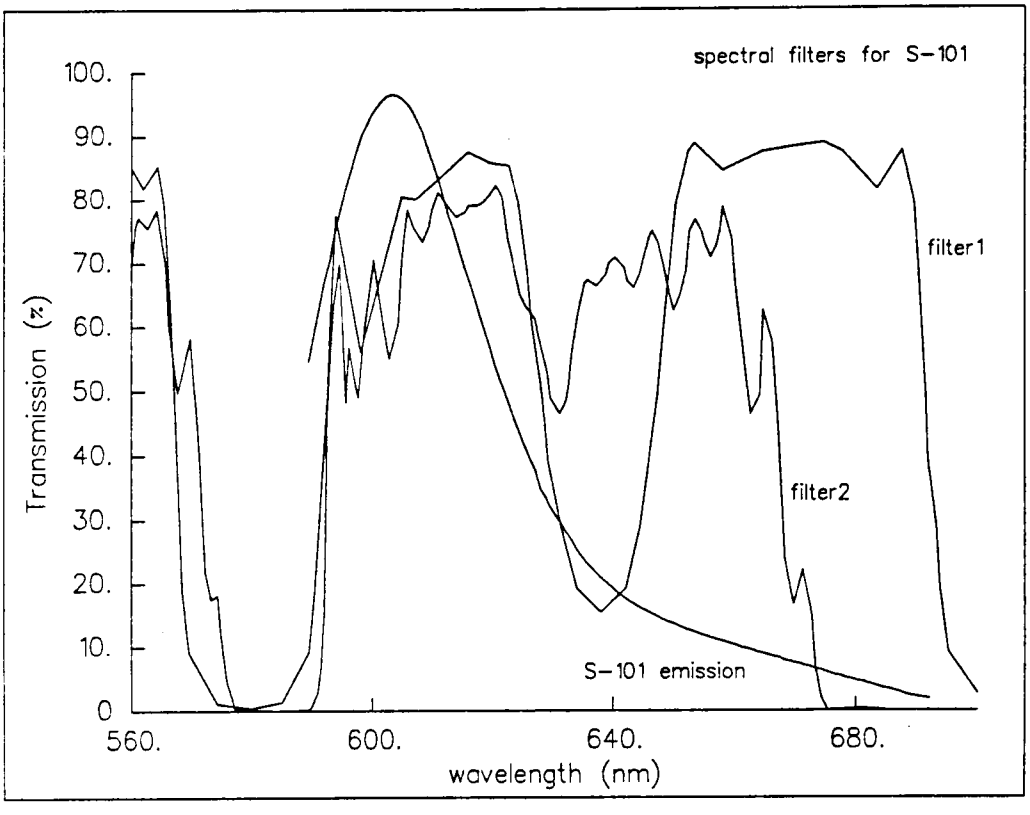



Figure 5.11: Transmission curves of the spectral filters for S-101

same excitation intensity for all molecules so that the distribution of photon burst amplitudes is not affected by laser fluctuations.

Secondly, it is essential that the SMD apparatus be light-tight from the flow cell all the way to the detector. Since the SPAD detector is very sensitive, even a small amount of light leakage will generate an intolerable amount of background counts. In order to provide ready access to adjustments, it is preferable not to completely enclose the fluorescence detection arm but rather to provide a light tight housing around the optical path only. Because of the various structures and shapes of the optical components and holders, rather than machining a custom sized housing from pipes, the light-tight housing is constructed using black vinyl electrical tape (3M Super 33+).

Thirdly, optical alignment is crucial for the success of the experiment. The height of the excitation laser beam and the SPAD detector should be at the same level relative to the optical table. Since the fluorescence signal is too weak to be used for alignment purposes, a He-Ne laser is used to initially align the collection and refocusing objectives, the spatial filter, the spectral filter and the lateral and focusing positions of the SPAD. Alignment is checked when the flow cell is in position. The flow cell is filled with detergent solution ('micro' from International Products Corporation, Trenton NJ) so that the laser beam may be viewed using light scattering from bubbles in the detergent. This procedure is used to focus the laser beam in the middle of the cell and to adjust the focus and position of the collection objective so that the image of the probe region passes through the spatial filter.

Fourthly, it is essential that the flow system be thoroughly clean before running the experiment. The outside walls of the flow cell are cleaned using lens tissue and methanol. The inside of the flow cell is cleaned by flowing the 'micro' detergent, which is also used for the optical alignment as described above, and then by flushing with clean water for $\sim$ 20 minutes. When the system is clean, the time spectrum from clean water should consist of a narrow prompt without a tail and the count

rate within the window of the time gate should be low. Also, the autocorrelation function should be flat, as discussed in §8.3.

Lastly, in setting up the experiment one must pay extra care not to break the flow cell when adjusting the collection objective into focus. The very short working distance of the high numerical aperture microscope objective makes it difficult to find the optimum focus position. The objective cannot be translated closer to the probe region to observe defocusing of the image of the laser beam. A high N.A. value microscope objective with an even longer working distance, if existing in the market, would greatly facilitate the operation of the experiments.

# Chapter 6

# Data acquisition and on-line analysis

## 6.1 Overview

For measuring fluorescence decays by time-correlated single photon counting, the output from the TAC is directed to a MCA, which gives the full time spectrum. For SMD, the time spectrum is still needed so that the window for time-gated detection can be selected. However, the data acquisition scheme must also process the TAC output pulses to provide the necessary information for the detection of the photon bursts that indicate the passage of individual molecules through the laser beam.

In Fig. 5-4, the data acquisition part of the experimental set-up is illustrated simply by a box. Fig. 6.1 gives a detailed schematic of the data acquisition hardware within this box. In Fig. 6.1, two separate data acquisition schemes are illustrated. To the left, within the dashed lines, is the scheme used in the early experiment of Ref. [15]. The later scheme, used to obtain the experimental results of Chapter 8, is illustrated within the solid line to the right.

The full time spectrum is necessary in both schemes. The biased amplifier (Tennelec TC 252) is used to match the range of the TAC output pulses so that it covers the full range of the collected time spectrum. However, the biased amplifier is not essential since the correct placement of the decay curve in the time spectrum can be achieved by means of variable delay lines. Nevertheless, it greatly facilities this

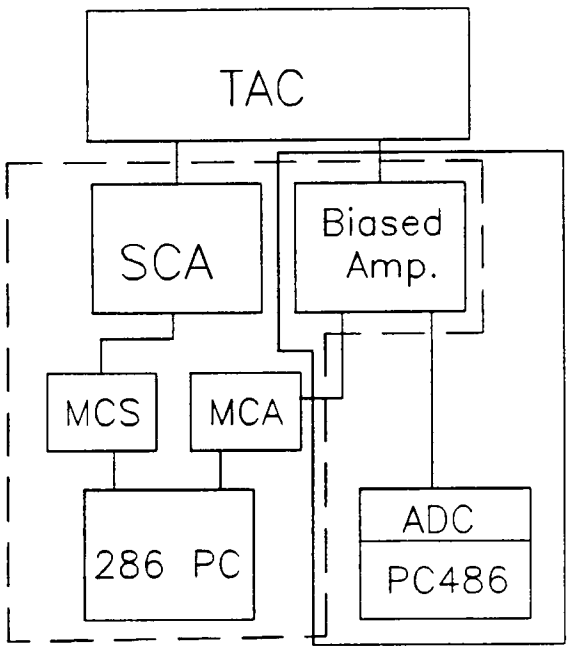



Figure 6.1: Data acquisition hardware connection

operation and furthermore allows the TAC output pulses to be amplified to yield increased resolution.

From the prompt in the time spectrum one can judge whether the optical alignment is correct and from the tail of the prompt one can judge whether the blank solution is clean. The time window is set by viewing the time spectrum. In the early scheme, this is accomplished through hardware, i.e., by adjusting the upper and lower limits of the single channel analyzer (SCA), which is contained within the TAC module. In the later scheme, the TAC output pulses are digitized and the time window is set using software. The new data acquisition scheme also provides more flexibility and versatility in data display and analysis.

## 6.2 Initial data acquisition scheme

The early data acquisition scheme involves the use of a 286 PC-based multichannel scalar (MCS) (EG&G Ortec Ace-MCS). The MCS counts the output pulses from the

SCA, which indicate photon pulses within the time window of the TAC. For a dwell time of 2.5 $ms$, the MCS collects $\sim$ 1 second of data in 4096 channels. The data is then saved to the hard disk of the PC. A Fortran program then reads in the file generated by the MCS software, processes it with the simple sliding sum algorithm and plots the result on the computer screen. The simple sliding sum advances one channel for each time step and is calculated at each time step by adding the photon counts over a fixed number of adjacent channels. The number of channels is chosen to be the full transit time of the molecule divided by the dwell time of the MCS. For the experiment of Ref. [15], the simple sliding sum is performed over 12 channels.

The main disadvantage of this scheme is that the data acquisition is not continuous. Since the analysis of the data is not performed at the same time as it is acquired, the display of the photon bursts is delayed and must be intermittently interrupted for the collection of new data. For every one second of data collection, the acquisition process is inactive for $\sim$ 3 seconds. This inevitably causes loss of data. Also, the autocorrelation function, which is essential for determining the molecular transit time for data analysis, is obtained by off-line analysis. The data acquisition software does not incorporate other important information such as counting statistics. For system optimization, the autocorrelation function, the prompt spectrum as well as counting statistics should be generated and displayed in real time.

## 6.3 Hardware of improved scheme – EISA-A2000 board

For the later experiments, a new data acquisition scheme is implemented. The TAC pulses are processed by the biased amplifier as before and then directed to a fast analog-to-digital converter (ADC). The ADC consists of a high speed analog input board (National Instruments EISA-A2000), which is connected into one slot of the EISA (extended industry standard architecture) bus of a 486 personal computer (PC). Software on the PC is used for selecting the window for the time-gated de-

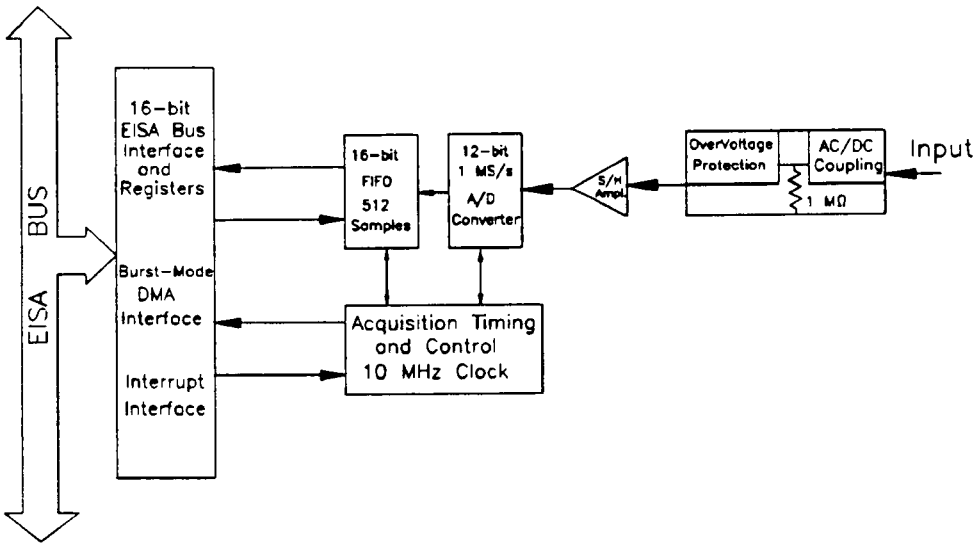



Figure 6.2: Simplified block diagram of EISA-A2000 analog to digital converter

tection, for processing the stream of photon counts in order to recognize bursts of photons due to single molecules, for real time graphical presentation of the photon bursts and for data storage.

The board has 4 channels with independent sample-and-hold circuits. Multiple data acquisition boards can be synchronized via the National Instruments Real-Time System Integration (RTSI) bus. However, the SMD experiment uses only one channel, for which a simplified block diagram is shown in Fig. 6.2.

The acquisition timing and control circuitry generates the necessary timing signals for the data acquisition process. An on-board programmable 16-bit sample interval timer generates the timing for the analog to digital conversions. The timer has software selectable time bases of 200 $ns$, 1 $\mu s$, 10 $\mu s$, 100 $\mu s$, 1 $ms$ and 10 $ms$. The minimum sample interval for single-channel data conversion is 5 clock cycles at the shortest time base, i.e., 1 $\mu s$. Thus the board can digitize at a sample rate of up to $10^6\ s^{-1}$ for one channel, or at reduced rates for more than one channel by multiplexing.

The ADC provides 12-bit resolution over a range of $-5$ to 5 $V$, which is equivalent to an analog voltage resolution of 2.44 $mV$. Since the output pulses from the biased amplifier are in the range of 0 to 10 $V$, they cannot be passed directly to the ADC input. Furthermore, the output impedance of the biased amplifier and the impedance of the cable are each 50 $\Omega$ but the input impedance of the ADC board is $1M\Omega$. By terminating the 50 $\Omega$ cable with a $50\Omega$ resistor at the ADC input, cable reflections are avoided and the input pulses seen by the ADC are reduced to the range of 0 to 5 $V$.

The EISA-A2000 uses the PC system DMA controller to transfer the digitized data directly to the extended memory (XMS) of the computer using EISA burst-mode (DMA type C) transfers. A 512-sample first-in-first-out (FIFO) buffer prevents data loss under conditions when the data cannot immediately be transferred to memory. In burst mode, DMA requests are asserted whenever the FIFO contains a block of 1, 16, or 32 samples. The DMA request remains asserted until the DMA controller transfers the samples to memory at full-burst speed– 16.5 *Mbytes/s*. The transfer of data is performed without intervention by the computer CPU and consequently the digitized data can be saved and analyzed at the same time as new data is being acquired.

A 32-bit sample counter keeps track of the total number of conversions and stops the acquisition process when the desired number of samples has been acquired. Alternatively, the board can operate in a continuous acquisition mode, in which case the sample counter is disabled and the data is written into the PC XMS in a cyclic manner.

## 6.4 Software and data structure

Fig. 6.3 shows the block diagram of the software hierarchy. The user-interface data acquisition program is developed in our laboratory using C-language. This program calls for driver software also written in C (National Instruments NI-DAQ). The

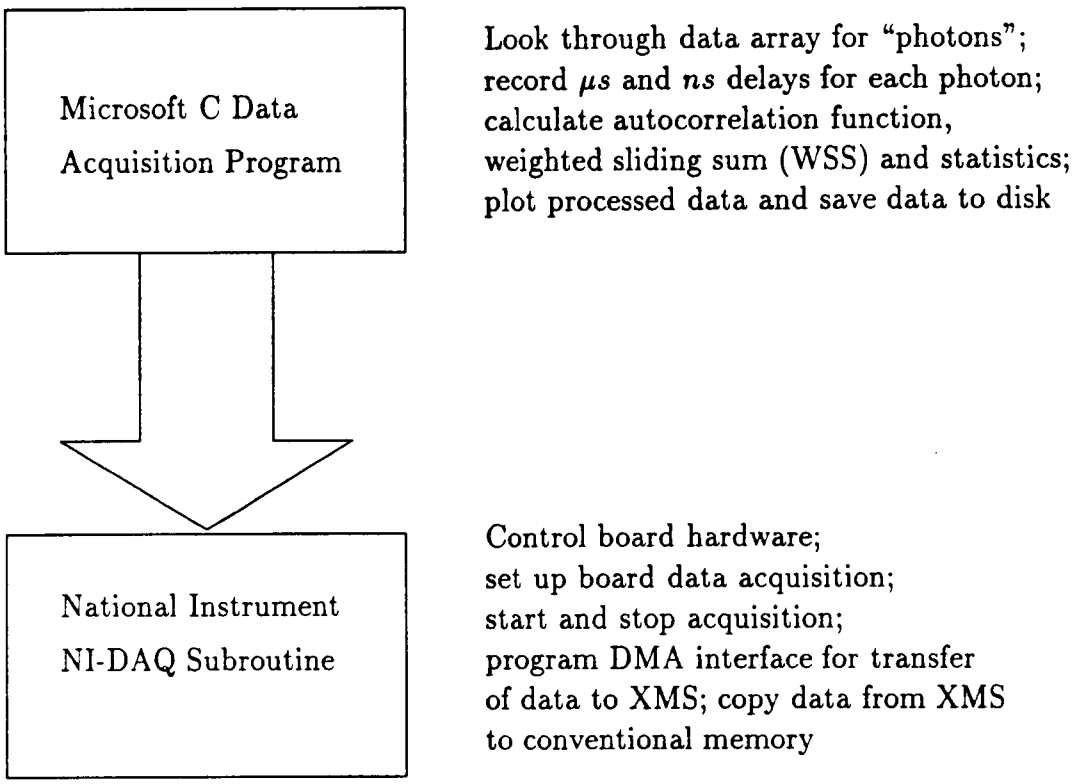



Figure 6.3: Software hierarchy

driver software includes routines which initialize the board and PC DMA controller, specify the rate and number of samples to collect, begin the data acquisition and transfer the data from XMS to conventional memory, where it may be accessed by other software. At the bottom of the hierarchy is the National Instruments's assembly-level software which programs the registers on the hardware including DMA control.

Output pulses from the TAC give rise to pulses at the ADC input in the range of 0 to 5 $V$ which occur at random times. The software records the number of

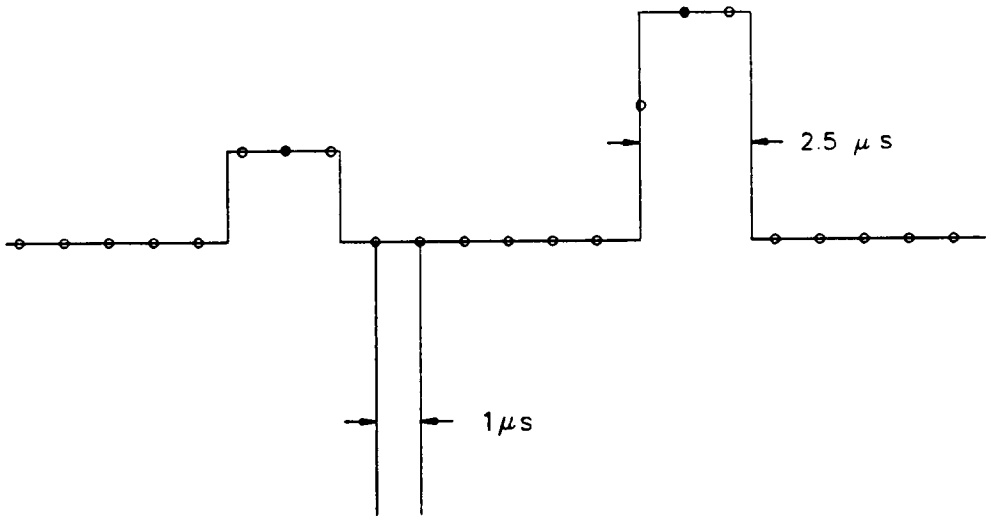



Figure 6.4: TAC output pulses and sampling points

microseconds between successive pulses, $m_i$, as well as the digitized amplitude of each pulse, $n_i$. Thus for each TAC pulse, a pair of data $(m_i, n_i)$ is created and saved to disk and into an array for further analysis. The data acquisition program uses this data array to construct the time spectrum, the autocorrelation function and the photon burst signal.

Since the ADC samples the input at equally spaced points at a rate of $10^6\ s^{-1}$ and the TAC generates pulses at a considerably lower rate, most of the digitized points give a zero voltage reading. Furthermore, since the TAC output occurs at random times, some of the digitization points may fall on the rising or trailing edges of the pulses. If this happens, the voltage read out will not be the full amplitude. To obtain a digitization point that is always at the flat top of the TAC pulse, the width of the pulse is adjusted to $\sim 2.5\ \mu s$. This ensures that at least two points are sampled from each TAC pulse. The second sample point will never be located at the rising or trailing edges, as illustrated in Fig. 6.4. The value at this point is then used as the pulse amplitude, $n_i$, and the number of digitizations since the last such point gives the time interval, $m_i$.

## 6.5 Digitized data stream

As shown in Fig. 6.4, the raw data consists of a stream of digitized values, either 0 or some positive value. The data rate is 1 million points per second and each data point is stored as a 2-byte integer directly into the XMS of the PC. For a total XMS of 7 Mbytes, approximately 3.5 seconds of data can be stored at one time. While data is being acquired, execution of the program loaded in the conventional memory of the PC allows the raw data stream stored in the XMS to be copied segment by segment to the page frame of the operating system. The data acquisition program then searches the page frame for peaks to obtain the data pairs $(m_i, n_i)$.

After all the raw data in the XMS memory are analyzed, another cycle of acquisition begins. Overall, due to the fact that the DOS program uses only 16-bit addressing, the XMS cannot be directly accessed but the the data must be copied down to the page frame by the XMS driver (himem.sys). The copy process slows the software so that it cannot quite keep up with continuous data acquisition. For every 3.5 seconds of acquisition, it requires $\sim$ 1 extra second for analysis. Thus data is not acquired continuously, although the acquisition process and display is continuous. It is possible that by using a 32-bit C-compiler one can directly search for peaks within the XMS memory.

After processing of the raw data, a pair of data $(m_i, n_i)$ is obtained for each TAC pulse. The first value, $m_i$, contains the time delay since the last pulse in $\mu s$, while the second value, $n_i$, contains the digitized TAC pulse amplitude. Recall that the TAC pulse amplitude is proportional to the time interval between the START and STOP which are supplied by the SPAD and excitation laser respectively. Thus, $n_i$ contains the arrival time (in $ns$) of the detected photon relative to the laser pulse. A histogram of the frequency of occurrences of the nanosecond delays, $n_i$, gives the time spectrum, which is the same as that obtained by the MCA. A graph of the time spectrum is used to observe the setting of the upper and lower limit of the time filter for time-gated detection, which is controlled by software. If a pulse does not

fall within the window of the time gate, the microsecond delay, $m_i$, is added to that of the following pulse. This process generates a sequence of numbers, $\bar{m}_i$, which give the time intervals between the detected photons that do pass the time gate. These values are then used for the further data analysis explained in §§6.6 and 6.7.

## 6.6 Time averaged autocorrelation function

The time averaged autocorrelation function gives a very sensitive indication of photon bursts. Previous researchers have used the autocorrelation function as a criterion to determine whether a detector has sufficient sensitivity for detecting single molecules.[9] Furthermore, references on dynamic light scattering show that the autocorrelation function can be used to obtain useful information about the average motion of small particles in solution.[30]

For a signal of amplitude $d(t_i)$, defined at fixed time intervals, the time averaged autocorrelation function in digital form is given by

$$g(\tau_j) = \sum d(t_i)d(t_i + \tau_j). \tag{6.1}$$

The sum is taken over the full duration of the experiment. For a photon counting experiment, the signal amplitude $d(t_i)$ is the number of photons within a fixed interval at time $t_i$. In the SMD experiment, $d(t_i)$ may be obtained by the data acquisition software from the sequence of numbers, $\bar{m}_i$, which give the time intervals between photons accepted by the time gate.

A typical sequence of data, $\bar{m}_i$, may have the following values:

$$23, 123, 34, 13, 12, 10, 6, 56, 342, 70, ....$$

This means that after the first photon, the second photon comes 23 $\mu s$ later, and the third photon comes 123 $\mu s$ after the second, etc. From this set of $\bar{m}_i$, one can easily construct the fluorescence photon burst signal, $d(t_i)$, which is shown in Fig. 6.5.

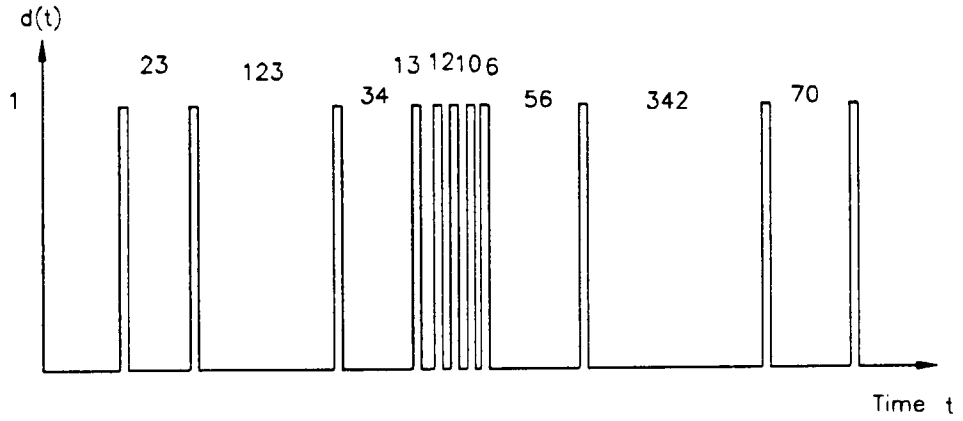



Figure 6.5: Reconstructed fluorescence signal $d(t_i)$ constructed from $m_i$

In this case, note that most of the time $d(t_i) = 0$, but whenever a photon is detected, $d(t_i) = 1$. This fact is used in the algorithm for the calculation of the autocorrelation function in Eqn. 6.1, directly from the $\bar{m}_i$.

If no molecules pass through the laser and all detected photons occur at random times so that the signal corresponds to white noise, then the time averaged autocorrelation function is expected to consist of a delta function at $\tau_j = 0$ superimposed upon a completely flat background. Thus a flat autocorrelation function is used as a test to indicate that the system is completely clean.

However, if molecules pass through the laser to give photon bursts, the autocorrelation function will contain a broadened peak at $\tau_j = 0$. Since the autocorrelation function is collected over the full duration of the experiment, all molecules contribute to the peak, which will be present even if the individual photon bursts are very weak. Thus the peak in the autocorrelation function is a very sensitive indication of the

passage of molecules.

The width of the peak is proportional to the mean duration of the photon bursts, which in turn is determined by the mean transit time of molecules through the laser beam. The laser intensity has a Gaussian profile given by Eqn. 5.3. If molecules travel at a constant linear velocity through the center of the laser beam ($y = 0$), the time-dependent excitation intensity is found by substituting $x = vt$ into Eqn. 5.3, giving

$$I(t) = \frac{2P}{\pi\omega_x\omega_y} \exp[-2v^2t^2/\omega_x^2]. \tag{6.2}$$

For excitation below saturation and no photodegradation, the average fluorescence emission rate is proportional $I(t)$, as explained in §7.2. In this case, the fluorescence emission rate, $F(t)$, should be Gaussian, i.e.,

$$F(t) \sim \exp\left[-\frac{2t^2}{\delta_t^2}\right], \tag{6.3}$$

where $\delta_t = \omega_x/v$ is the half-transit time of the molecule, i.e., the time required for the molecule to travel from the center of the laser to the $1/e^2$ intensity point.

Since the autocorrelation function of a Gaussian is another Gaussian with a width of $\sqrt{2}$ times that of the original, the autocorrelation of the emission intensity is proportional to

$$A(\tau) \sim F(t) * F(t) \sim \exp\left[-\frac{\tau^2}{\delta_t^2}\right]. \tag{6.4}$$

Thus the standard deviation of the autocorrelation is

$$\sigma_a = \delta_t/\sqrt{2}. \tag{6.5}$$

The half width at half maximum of the autocorrelation function, which is more easily estimated by eye, is $\sqrt{\ln 2} = 0.83$ times $\delta_t$.

## 6.7 Photon burst detection

Although the autocorrelation function is a sensitive indication of single molecule detection, it cannot be used to indicate the passage of *individual* molecules.

To generate the photon burst signal, $S(t_i)$, that signifies the passage of individual molecules, it is necessary to process the raw data $\bar{m}_i$. The simplest data processing scheme is a simple sliding sum in which, at each time step, $t_i$, all photon counts within a fixed interval are summed to obtain the photon burst signal as

$$S(t_i) = \sum_{\tau=0}^{2\delta_t} d(t_i + \tau). \tag{6.6}$$

Note that the interval over which counts are summed is chosen to be the full transit time of the molecule, $2\delta_t$. The simple sliding sum exhibits a peak whenever the photons are bunched together in a burst, i.e., whenever the raw data $\bar{m}_i$ contains a sequence of small numbers.

Instead of the simple sliding sum, it is possible to use other processing schemes, which employ digital filters to enhance the signal. For example, Shera *et al.*,[10] in their first single molecule detection experiment, adopted a weighted quadratic filter defined as

$$S(t) = \sum_{\tau=0}^{k-1} w(\tau)(d(t + \tau))^2, \tag{6.7}$$

where the range, $k$, covers a time interval of the order of the residence time of the molecule. The weights, $w(\tau)$, were chosen as an asymmetric triangular ramp, i.e.,

$$w(\tau) = (\tau + 1)/k, \tag{6.8}$$

to account for the sudden cessation of the signal due to photobleaching in their experiment. Their later simulations indicated that the quadratic nature of the above filter does not improve the detectability.

In the experiments in this work, most molecules do not photodegrade and on average the fluorescence emission is expected to exhibit a Gaussian temporal profile, given by Eqn. 6.3. Hence a digital filter based upon a weighted sliding sum, with Gaussian weights directly proportional to Eqn. 6.3, is adopted, i.e.,

$$S(t_i) = \sum_{\tau=-\delta_t}^{\delta_t} w(\tau) d(t_i + \tau), \tag{6.9}$$

where

$$w(\tau) = 2\sqrt{2/\pi} \exp\left(-\frac{2\tau^2}{\delta_t^2}\right). \tag{6.10}$$

By applying the weighted Gaussian filtering, fluorescence bursts characteristic of the transit of individual molecules are enhanced. Though random fluctuations in the background may occasionally produce a burst of photons in a short time, they are less likely to duplicate the characteristic time profile of the single molecule fluorescence bursts. $S(t)$ is large whenever the temporal pattern of incoming photons is typical of that from a molecule. Thus the presence of a molecule is recorded when the value of $S(t)$ exceeds a set threshold.

The optimum value for the threshold may be determined from the histogram of burst amplitudes, which is also used to evaluate the detection efficiency of single molecules. In generating the burst amplitude histogram, fluctuations at the leading or trailing edge of the burst due to shot noise are ignored in order to avoid counting a single burst more than once. As can be seen in the experimental results of Fig. 8.9, the distribution of burst amplitudes from pure water consists only of small amplitude bursts due to noise fluctuations. However, the distribution due to single molecules contains an additional peak at a finite burst amplitude. The threshold for counting single molecules is chosen to be at the valley before the peak in order to maximize the probability of detecting a molecule while also minimizing false detections.

# Chapter 7

# Calculations and simulations

In this chapter calculations and simulations of several details of the experiments are presented. This serves to select operating conditions that are appropriate for successful experiments, to examine whether the measurements generate expected resultsand also to form a better understanding of the experiments, including limitations and possible improvements.

§7.1 explains how to select a suitable concentration for the ultradilute solution such that the probability of two molecules occupying the probe volume at the same time is negligible. In §7.2, the dependence of the fluorescence photon yield on the molecule transit time, laser power and beam size is given. The expected photon yield for the conditions of the experiment in §8.4 is calculated. The dependence of the background noise on the experimental parameters and the expected photon statistics are described in §7.3. In §7.4, the effects of photodegradation are discussed and the fraction of molecules that photodegrade in the experiment of §8.4 is calculated. §7.5 presents the factors that must be considered in optimizing experimental conditions. In §7.6, a simple Monte Carlo simulation model is developed to generate the mean count rate versus time as molecules transit across the laser beam. The simple simulation, which includes the random location of molecules and random photodegradation, led to the development in our laboratory of a more sophisticated simulation, which also includes diffusional spreading, detector and electronics

dead time. Although the programming for this simulation was not a part of this dissertation, the results of the simulation guided the choice of parameters used in the experiment and gave confidence in the validity of the results. Therefore, to give a comprehensive presentation, results of the sophisticated simulation under the conditions of the experiment of §8.4 are given in §7.7.

## 7.1 Single molecule concentration

In the simple flow cell, the concentration of the ultradilute solution must be adjusted so that most photon bursts are due to single molecules rather than two or more molecules passing through the probe region simultaneously. Since molecules pass through the probe region at random times, the probability of having $m$ molecules at any given time has a Poisson distribution with

$$P(m) = \frac{(\overline{m})^m e^{-\overline{m}}}{m!}, \tag{7.1}$$

where $\overline{m}$ is the mean number of molecules in the probe volume.

In order for most photon bursts to be caused by single molecules, the probability of having one molecule in the probe region must be considerably larger than the probability of having two molecules. If the maximum fraction of peaks due to two molecules that may be tolerated in an experiment is $f$, then from Eqn. 7.1, the ratio of probabilities must be

$$\frac{P(2)}{P(1)} = \frac{\overline{m}}{2} = f \ll 1. \tag{7.2}$$

Thus the maximum concentration, $C$, of the ultradilute sample solution in moles per liter is determined as

$$C = \frac{\overline{m}}{V_p N_A} = \frac{2f}{V_p N_A}, \tag{7.3}$$

where $N_A = 6.023 \times 10^{23}$ *molecules/mole* is Avagadro's number and $V_p$ is the volume of the probe region in liters.

If $\overline{m} = .02$, the fraction of peaks due to two molecules passing simultaneously through the probe volume is $f = 0.01$ or only 1%. For the experiment in the simple flow cell of §8.3, the volume of the probe region is

$$V_p = 9.6 \times 13.3 \times 12 \ \mu m^3 = 1.5 \ pl \tag{7.4}$$

and the concentration must then be $C < 2.2 \times 10^{-14} M$.

In the new diffusion-limited flow cell, the mean number of molecules inside the probe volume depends on the sample delivery rate as well as the concentration of the sample solution in the capillary tip. In practice, a small concentration $\sim 10^{-13} M$ is used and the sample delivery rate is then adjusted so that condition 7.2 is satisfied. Since the exact sample delivery rate from the capillary tip is not accurately known, the mean number of molecules inside the probe volume, $\overline{m}$, is determined from the observed mean rate of photon bursts $\overline{M}$ and the molecule transit time $2\delta_t$ as

$$\overline{m} = \overline{M} \times 2\delta_t. \tag{7.5}$$

If the observed number of bursts is $\overline{M}$, then the fraction of peaks due to two molecules passing through the probe region simultaneously is

$$f = \overline{M}\delta_t. \tag{7.6}$$

For the experiment of §8.4, $\delta_t \sim 0.5 \ ms$, $\overline{M} \sim 150 s^{-1}$ so that $f \sim 0.07$.

## 7.2 Signal strength estimation

The mean number of fluorescence photons that are detected when a molecule passes through the laser beam, $\overline{n}_f$, is determined by the number of times the molecule is

excited, $N$, the fluorescence quantum efficiency of the molecule, $\Phi_f$, and the overall photon detection efficiency of the instrument, $E$, i.e.,

$$\overline{n}_f = N \times \Phi_f \times E. \tag{7.7}$$

Since the molecule can be excited only once for each laser pulse, the maximum number of times that it may be excited is given by the full transit time multiplied by the repetition rate, $R$, of the laser (see §2.5), provided that photodegradation does not occur. However, typically the laser intensity is such that the molecule is excited by less than one pulse in ten. Also, as the molecule passes through the Gaussian profile of the focused laser, the intensity and hence the probability of excitation vary. For a molecule with absorption cross-section $\sigma_a$, the probability of excitation per laser pulse is

$$p_{excite}(t) = 1 - \exp\left[-\frac{\sigma_a I(t)}{3RE_\gamma}\right], \tag{7.8}$$

where $E_\gamma$ is the photon energy. The factor of 3 accounts for the random distribution of the dipole moment of the molecule with respect to the polarization of the laser excitation beam.

When excitation is far below saturation, Eqn. 7.8 gives

$$p_{excite}(t) \approx \frac{\sigma_a I(t)}{3RE_\gamma} \tag{7.9}$$

and the number of excitations per unit time is

$$\begin{aligned} N_{excite}(t) &= R\, p_{excite}(t) \\ &= \frac{\sigma_a I(t)}{3E_\gamma}. \end{aligned} \tag{7.10}$$

If the molecule passes through the center of the laser beam, $I(t)$ is given by Eqn. 6.2. The total number of laser excitations experienced by a molecule that passes through the center without photodegradation is therefore given by

Table 7.1: Parameters for S-101 used for calculation of signal strength

| | | |
|---|---|---|
| laser power | $P$ | $14\ mW$ |
| half transit time | $\delta_t$ | $0.50\ ms$ |
| photon energy at 585 $nm$ | $E_\gamma$ | $3.4 \times 10^{-19} J$ |
| absorption cross section | $\sigma_a$ | $3.7 \times 10^{-16} cm^2$ |
| laser beam waist in flow dirn. | $\omega_x$ | $4.8\ \mu m$ |
| laser beam waist in transverse dirn. | $\omega_y$ | $9.4\ \mu m$ |
| laser repetition rate | $R$ | $76 \times 10^6\ s^{-1}$ |
| peak excitation probability per pulse | $P_{excite}$ | 0.096 |
| mean number of excitations | $N$ | 4220 |
| collection objective efficiency | | 0.12 |
| SPAD quantum efficiency | | 0.55 |
| throughput due to SPAD dead time | | 0.80 |
| throughput of collection objective | | 0.93 |
| throughput of spatial filter | | 0.80 |
| throughput of spectral filter | | 0.40 |
| throughput of time filter | | 0.60 |
| overall detection efficiency | $E$ | $9.6 \times 10^{-3}$ |

$$\begin{aligned} N &= \int_{-\infty}^{\infty} N_{excite}(t)\, dt \\ &= \sqrt{\frac{2}{\pi}} \frac{\sigma_a P \delta_t}{3 E_\gamma \omega_x \omega_y}. \end{aligned} \tag{7.11}$$

The values for the various parameters in Eqn. 7.11 for the experiment of §8.4 are listed in Table 7.1. The peak excitation probability per pulse is calculated from the listed parameters to be 0.096, indicating that excitation is well below saturation. Using the parameters listed in the table gives

$$N = 4.22 \times 10^3.$$

In the lower part of the Table 7.1 are the efficiencies of the various filters and components of the instrument that contribute to the overall efficiency of photon detection. Also included is an additional 20 % loss of signal throughput due to the

dead time of the SPAD, as discussed at the end of §3.4. Using the value of $E$ from the table and a fluorescence quantum yield of $\Phi_f = 0.35$ for S-101 in water, the mean number of detected photons as calculated from Eqn. 7.7 is expected to be

$$\overline{n}_f = 14.2. \tag{7.12}$$

Due to the uncertainty in the efficiencies of the various components, the error in the overall detection efficiency and hence in the expected number of fluorescence photons is estimated to be about 25 %. Also, the expected signal will be less for molecules that do not pass through the exact center of the laser profile.

## 7.3 Background and photon statistics

The bursts of fluorescence due to single molecules passing through the beam are superimposed upon a continuous background (See Fig 2.2). The background is linearly proportional to the laser power and the collection time. In the time interval $2\delta_t$ during which a molecule passes through the beam, the mean number of background photons, $\overline{n}_b$, is given as

$$\overline{n}_b = K\ P\ 2\delta_t. \tag{7.13}$$

The empirical proportionality constant, $K$, accounts for both Rayleigh and Raman scattering from the solvent as well as stray fluorescence. It is dependent on the number of solvent molecules in the probe volume and on a factor which accounts for the collection and throughput of the background noise through the various filters.

In the experiment of §8.4, $K$ is found to be $\sim 4.6 \times 10^5\ s^{-1}W^{-1}$ and for a laser power of 14 $mW$, the mean number of background counts during an interval of $2\delta_t = 1.0\ ms$ is $\sim 6$. Thus, during the transit of a single molecule, the total mean number of detected photons, $\overline{n}$, is expected to be

$$\overline{n} = \overline{n}_f + \overline{n}_b \tag{7.14}$$

$$= 14.2 + 6 = 20.2. \tag{7.15}$$

When there are no molecules in the beam, the probability $P_0$ of having $n$ counts in a time interval $2\delta_t$ is given by the Poisson distribution, i.e.,

$$P_0(n) = \frac{\bar{n}_b^n e^{-\bar{n}_b}}{n!}. \tag{7.16}$$

For the transit of a single molecule through the center of the laser without photodegradation, the probability is

$$P_1(n) = \frac{\bar{n}^n e^{-\bar{n}}}{n!}. \tag{7.17}$$

Since molecules in the experiment do not necessarily pass through the exact center of the laser, the mean number of detected photons, $\overline{n}$, will be smaller. Furthermore, the probability density function $P_1(n)$ will be due to an ensemble average over different trajectories and hence will be broader. However, a well defined peak in the histogram of burst amplitudes should still be obtainable provided that most molecules pass close to the center of the laser. In the case of experiments conducted with the simple flow cell, molecule trajectories that barely pass through the laser will degrade the probability density function so that a well defined peak is no longer obtained.

## 7.4 Photodegradation

The results of §§7.2 and 7.3 apply only to molecules that do not photodegrade as they pass through the laser. However, this is the case for only a certain fraction of the molecules. The probability that a molecule has not photodegraded decreases exponentially with the number of excitations. The mean number of excitations a

molecule can endure before photobleaching is the reciprocal of the photodegradation quantum efficiency, $\Phi_D$.

For S-101 in water solution, as listed in Table 5.1, $\Phi_D = 5.5 \times 10^{-5}$ and thus a S-101 molecule can on average be excited 18,180 times before it photodegrades. In the experiment of §8.4, for a molecule that passes through the center of the laser, the mean number of excitations, as given in Table 7.1, is $N = 4220$. The probability that photodegradation occurs before the molecule exits from the laser beam is therefore

$$P_{photodegradation} = 1 - \exp[-N\Phi_d] = 0.207. \tag{7.18}$$

Even for such a low number of excitations, 20 % of the S-101 molecules photodegrade. This percentage would be decreased for a molecule with a lower photodegradation quantum efficiency such as R6G. By comparison, in the prior SMD experiments using R6G, [10] almost all molecules photodegrade.

The number of excitations that are required for SMD should be minimized to reduce the probability of photodegradation. However, even if a molecule photodegrades, it is still possible that it may have generated enough photons to be detected. Conversely, it is always possible that a molecule may photodegrade the very first time it is excited, before any fluorescence is detected. For those molecules that do photodegrade, the number of detected photons approximately follows an exponential probability density function, while for those molecules that do not photodegrade, the photon probability density function is as discussed in §7.3.

Whether a molecule photodegrades, the exact trajectory of a molecule through the focused laser and how many fluorescence photons are detected are dependent on random processes. In order to form a more complete understanding of the experiment, a Monte-Carlo simulation program is developed.

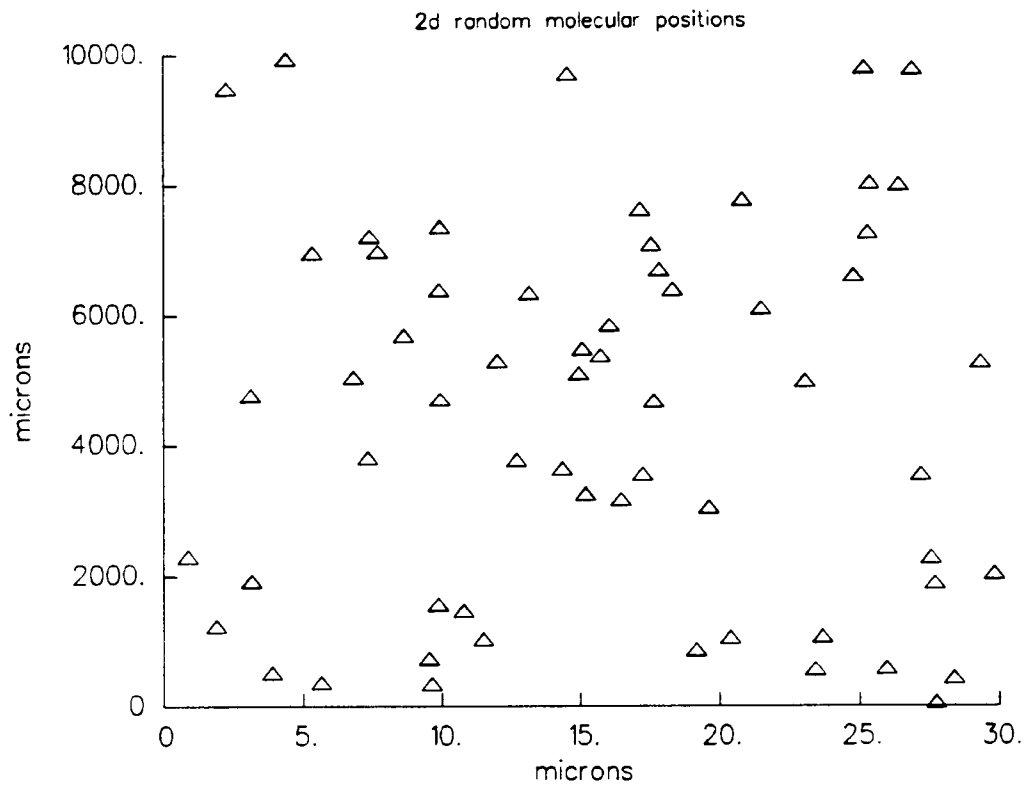



Figure 7.1: 2-D molecule map

## 7.5 Simple simulation model

At first, a simple Monte Carlo simulation program is developed in Fortran. This simulation models the experimental condition of the experiment to detect single molecules of R6G in the simple flow system, as reported in Ref. [15]. The dye solution fills the complete volume of the square-bore flow cell and only molecules that happen to pass through the laser beam are detected.

To begin with, a two-dimensional map of molecule positions is generated using random numbers as shown in Fig. 7.1. The depth of the probe region is assumed to be 10 $\mu m$. The area of the region which is simulated is $1000\mu m \times 30\mu m$, which results in a total volume of $3 \times 10^{-10} l$. For a sample solution concentration of $3 \times 10^{-13} M$, the number of molecules in this volume is 54.

Next, the focused Gaussian laser beam is uniformly translated across the molecule field at a speed equal to the flow velocity of the solution. At time $t$, the laser intensity at the coordinates of molecule $i$, $I(x_i, y_i)$ is given by

$$I(x_i, y_i) = \frac{2P}{\pi\omega_0^2} \exp\left[\frac{-2(x_i^2 + (y_i - y_0(t))^2}{\omega_0^2}\right].$$

The probability that the molecule is excited by the laser pulse is given by Eqn. 7.8. Therefore, the total number of fluorescence photon counts per second from all $M$ molecules inside the probe volume that have not photodegraded is

$$\overline{n}_f(t) = \sum_{i=1}^{i=M} \left(1 - \exp\left[-\frac{\sigma_a I(x_i, y_i)}{3E_\gamma}\right]\right) \times \Phi_f \times E, \tag{7.19}$$

where $\Phi_f$ is the fluorescence quantum efficiency and $E$ is the overall detection efficiency.

At each time step, photobleaching is considered for each molecule in the following way. The probability that a molecule undergoes photobleaching per time step is

$$P_D(t) = P_{excite}(t) \times \Phi_D \times R \times \Delta t, \tag{7.20}$$

where $R$ is the laser repetition rate and $\Delta t$ is the time step of the simulation. Note that $\Delta t$ must be small enough for the probability to be considerably less than unity. At this point, a random number, $\gamma$, between 0 and 1 is generated. If $\gamma < P_D$, the molecule is assumed to photodegrade, otherwise it is assumed that the molecule survives.

Background noise is added as in Eqn. 7.15 following the simple linear relation with laser power given by Eqn. 7.13. In this way, a plot of mean count rate versus time is generated, as shown in Fig. 7.2.

The simple simulation model only outputs the mean photon count rate versus time. Detailed photon statistics, effects caused by diffusion and dead time of the detector and electronics are not considered. Nevertheless, the simple simulation

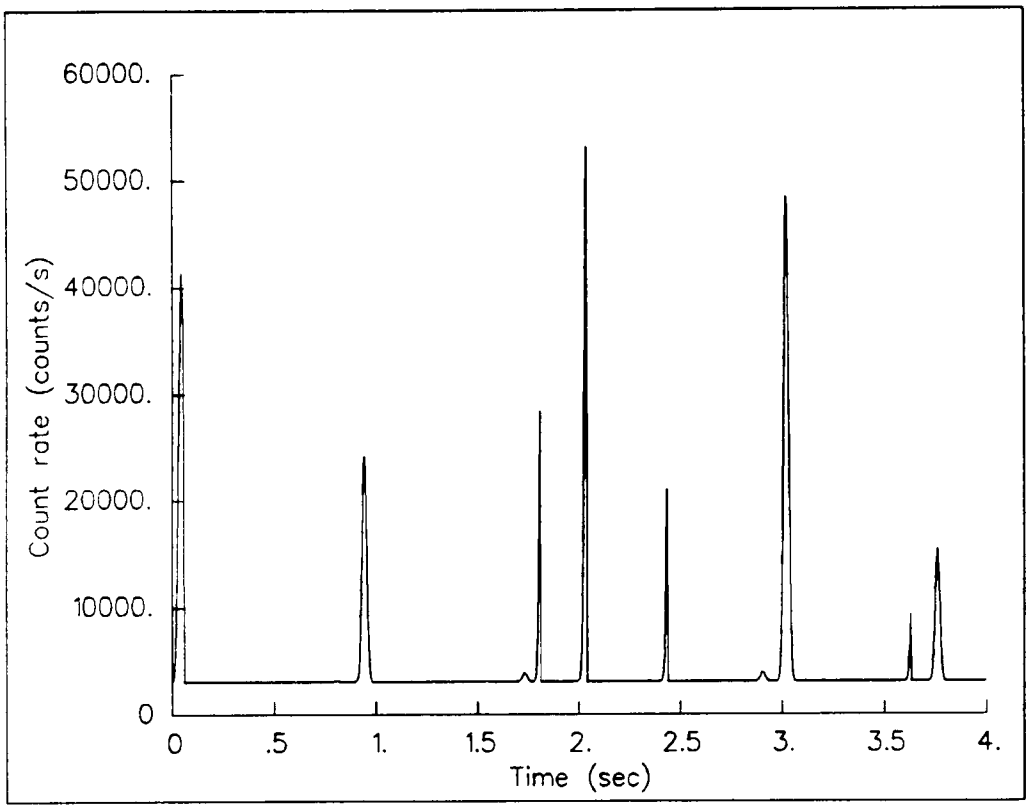



Figure 7.2: Mean count rate versus time

shows that the amplitudes of the photon bursts vary widely because many molecules in the simple flow cell pass near the edge of the Gaussian laser profile.

## 7.6 Complete simulation program

Based on the simple simulation model, a more complete Monte-Carlo simulation program was later developed in our group. [31] The new Monte-Carlo simulation program models the new sample delivery system employed in the later experiments for detecting B-PE and S-101 molecules. By contrast with the simple model, the laser beam is taken to be stationary and for each time step, the molecules are advanced. The new program considers the full three- dimensional geometry of the experiment including the DOC of the collection objective as discussed in §5.5.3. In addition to the Monte Carlo step for photodegradation, the new program includes Monte Carlo steps to account for molecular diffusion and for the photon statistics, although the same overall system detection efficiency as in Table 7.1 is assumed. At each time step the simulation generates either a 1 if a photon is detected or otherwise a 0. The train of photons is processed by the same sliding sum algorithm discussed in § 6.7. In this way, the simulation outputs the photon burst signal corresponding to the transverse of individual molecules through the laser beam.

The two most important factors that are not included in the simulation are laser fluctuations and the real flow velocity profile near the tip, which is unknown, as pointed out in §4.7. Since the simulation is not the main emphasis of this dissertation and much of the programming was performed by another in the group, details on the algorithm are not discussed. Further information is given in Ref. [31]. Only the main results pertaining to the experimental results of Chapter 8 are presented here.

Fig. 7.3 shows a brief segment of the photon burst data generated by the simulation for S-101 single molecules under the same experimental conditions as in §8.4, namely for a laser power of 14 $mW$ and a flow rate of $10\mu m/ms$. For this segment of data, as indicated in the insert, four molecules come out from the capillary and

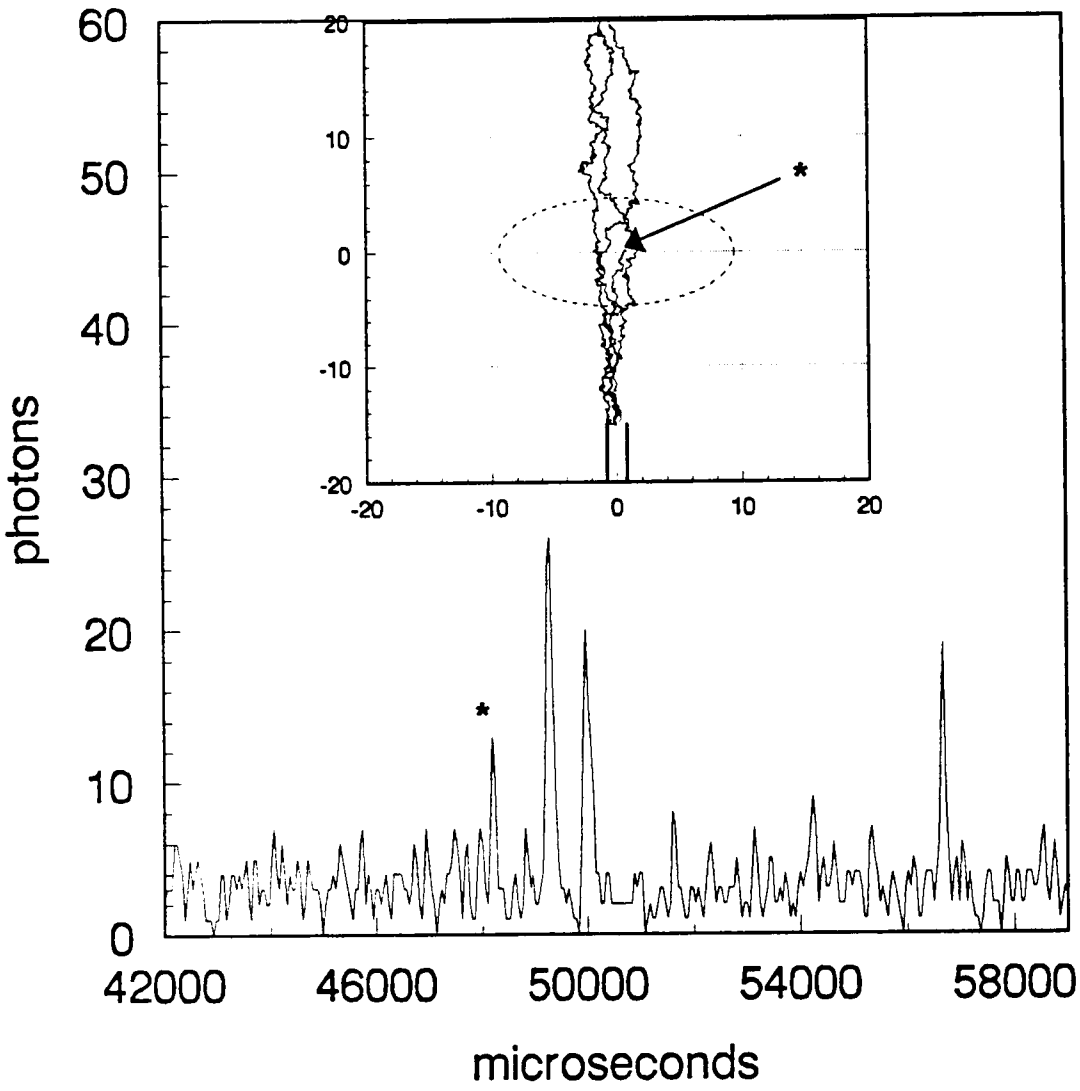



Figure 7.3: Molecule trajectories and corresponding photon bursts for S-101

pass through the focused laser, which is shown by the dashed oval line. It is interesting to note that even though the first molecule (marked by *) photodegraded while passing through the laser, it still gave a detectable photon burst signal. However, if many molecules photodegrade, the overall detection efficiency will be decreased and the burst amplitude distribution will be degraded. The other three molecules give bursts clearly distinguishable from the background fluctuations.

Another feature of the new simulation program is that one can accumulate statistics such as the histogram of the amplitudes of bursts due to molecules, the histogram due to false bursts caused by fluctuations in the background and the overall number of molecules and false bursts counted. Fig. 7.4 shows the histograms of the signal and false counts obtained after the passage of 1000 molecules under the same conditions as in Fig. 7.3. Most of the bursts due to background fluctuations are below 11 photons while most bursts due to single molecules of S-101 are greater than 11. To maximize the probability of detection minus the probability of false detection, the threshold is set at 11, which is at the valley between the two distributions. Note also that about 10 % of the bursts due to S-101 molecules are below the threshold and so are not detected. These bursts are largely due to molecules that photodegrade.

The program is interactive and allows many parameters such as the laser excitation intensity and flow speed to be adjusted from the keyboard. Useful knowledge on the optimum experimental operating conditions can be gained by adjusting of the various parameters and observing the changes in the statistics of the bursts.

In the prior experiments of Ref. [10], the laser power was selected to avoid saturation and a slow flow speed was used to extract as many photons as possible from each molecule. Consequently, almost all molecules would photodegrade while passing through the laser. These operating conditions were selected in order to generate the maximum *mean* signal-to-noise ratio, $n_f/\sqrt{n_b}$. [32] However, once a molecule is detected, there is no point in extracting more signal, as this will only lead to a

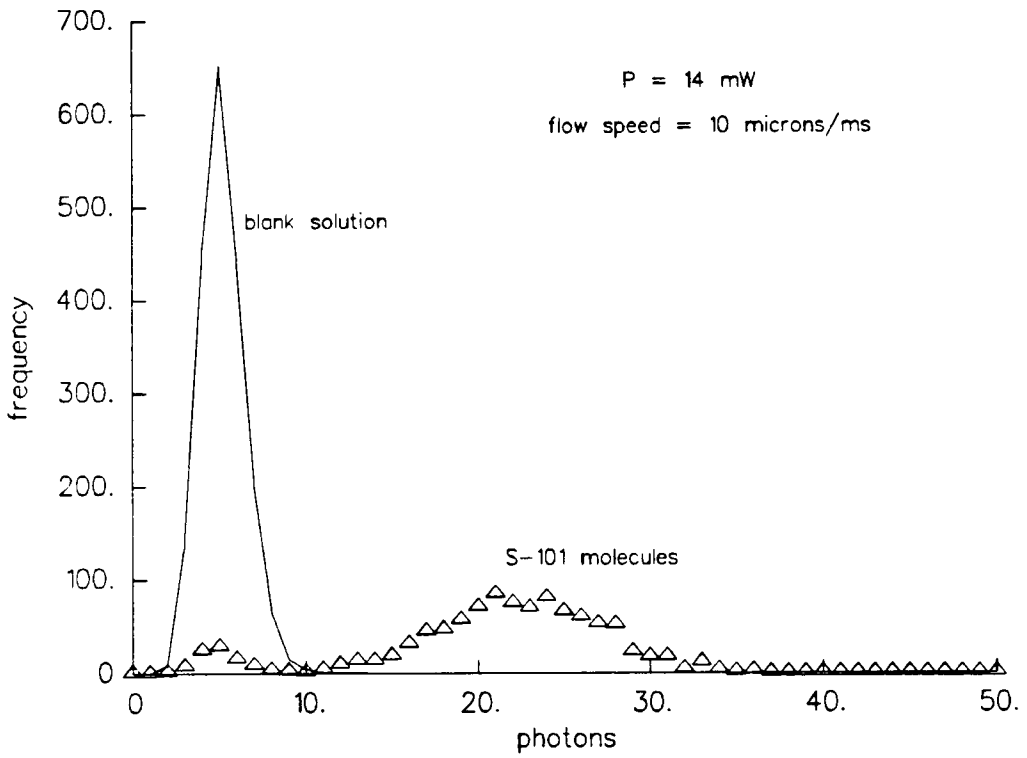



Figure 7.4: Photon Statistics built-up through the simulation

larger background. In contrast, in the experiments in this work, fast flow velocities are selected so that the majority of molecules do not photodegrade. This maximizes the probability of detection minus the probability of false detection.

# Chapter 8

# Results and discussions

## 8.1 Overview of experimental results

During the course of this project, experiments were conducted using different molecule species and different sample flow cell systems. Three types of fluorescent molecules were investigated—rhodamine 6G (R6G), sulforhodamine 101 (S-101) and the multi-chromophore protein molecule B-phycoerithrin (B-PE). Experiments were performed using both the simple flow cell described in §4.3 and the diffusion-limited sample delivery system described in §4.5.

In the first experiment, single R6G molecules were successfully detected at a concentration of $10^{-13}$ M in the simple flow cell. A significant increase in the amplitude of photon bursts is obtained compared to the prior experiments,[10] largely due to the use of the high quantum efficiency SPAD detector. The experiment is an important step as it shows that single chromophore molecules can be individually detected using the SPAD detector. Since details on this portion of the dissertation are reported in a published paper,[15] experimental results using R6G in the simple flow cell will not be repeated in this chapter.

Following the successful R6G experiment, efforts were focused on improving the overall molecule detection efficiency, i.e., on causing all molecules to pass through the probe volume. Also, the data acquisition and analysis scheme was improved, as discussed in §6.3. Preliminary experiments were performed to characterize the

sample stream size in the new diffusion-limited sample delivery system using concentrated dye solutions, as discussed in §4.8. SMD experiments were then conducted in the new diffusion-limited flow cell using B-PE molecules. As discussed in §5.1, the reasons for choosing B-PE are firstly, its large molecular weight, which reduces diffusional spreading, and secondly, its high yield of photon emission, which allows fast flow velocities to be used to further reduce diffusional spreading. Indeed, single B-PE molecules were detected at relatively low laser excitation power ($\sim 2\ mW$) and flow velocities almost ten times higher than in the R6G experiments. This experiment is reported in §8.2.

At that time, improvements in the sensitivity of the SMD apparatus were made by replacing the 0.65 N.A. collection objective with a new Nikon objective with 0.85 N.A. Encouraged by the successful results using B-PE in the new diffusion-limited flow cell, experiments were then attempted using the single chromophore molecule R6G. However, for the higher flow velocity required to overcome diffusional spreading, a higher laser excitation power ($\sim 20\ mW$) is required in order to achieve a sufficient fluorescence signal for single molecule detection. Although photon bursts of the anticipated amplitude could be detected using a R6G solution, similar bursts were also detected when using the clean triple-distilled water. Experiments with R6G were also attempted using the simple flow cell, which could be more easily cleaned by flushing, as it is a closed system with fewer potential contamination sources. Despite attempts to reduce impurities in the water by photodegradation before use and by using water from another vendor, a clean blank could still not be obtained for these experimental conditions.

In order to eliminate photon bursts from the unidentified impurities, new experiments using longer excitation wavelengths were attempted. The reasoning behind using longer wavelengths is that all fluorescent molecules have a cut-off wavelength in the absorption spectrum, beyond which fluorescence cannot be excited. Consequently, most fluorescent impurity molecules are only excited by short wavelengths.

Therefore the dye laser, operating at 585 $nm$, was used for an experiment to detect single molecules of S-101, in the hope that this wavelength would be beyond the absorption cut-off wavelength of the impurities.

Fortunately, the unidentified impurities did not generate photon bursts when excited using this wavelength and the SMD experiments using S-101 were successful. §8.3 presents the results of the SMD experiments using S-101 in the simple flow cell. This experiment was conducted using a spatial filter pinhole of the same size as required for experiments with the diffusion-limited flow cell, in order to evaluate if SMD experiments would be feasible under the same experimental conditions.

Experiments to detect single molecules of S-101 with the diffusion-limited sample delivery system were then carried out. The results are presented in §8.4. The experiment was successful in that it demonstrated that single chromophore molecules could be detected with a high overall detection efficiency ($> 75\%$) at relatively fast flow velocities. Also, a reasonably quantitative photon signal is obtained from each detected molecule, as indicated by the peak in the histogram of burst amplitudes.

The chapter ends with conclusions presented in §8.5 and suggestions for future work in §8.6.

## 8.2 B-PE in the diffusion-limited flow cell

Peck *et al.* have previously reported an experiment using B-PE to demonstrate single molecule sensitivity from the non-random feature in the autocorrelation function.[9] They claim (without presenting data) that $\sim 15\%$ of the molecules that pass through the laser beam are detected. However, in their experiments most molecules do not pass through the laser beam. By contrast, this section presents direct evidence for the detection of individual B-PE molecules using the new diffusion-limited flow cell, which causes almost all molecules to pass through the laser beam. The non-random feature in the autocorrelation function serves mainly to determine the transit time of the molecule and is not the only criterion of success of the experiment.

There are two considerations in choosing B-PE protein molecules for testing SMD in the new flow cell. Firstly, its fluorescence intensity is equivalent to $\sim$ 25 single chromophore molecules and thus it is relatively easy to detect. Secondly, its molecular weight is $\sim$ 500 times that of a typical single chromophore molecule, such as R6G. As discussed below Eqn. 4.7 on page 55, the diffusion coefficient is approximately proportional to the cube root of the molecular weight. Consequently diffusional spreading is smaller than that of a single chromophore molecule by a factor of about 8. It should therefore be easier to keep the sample stream size small and hence cause most molecules to pass through the laser beam.

Some of the experimental results using B-PE are shown in the next three figures. The experimental conditions are given in Table 8.1.

Table 8.1: Experimental conditions used for B-PE in the new flow cell

| | | |
|---|---|---|
| laser power | $P$ | $\sim 2\ mW$ |
| excitation wavelength | $\lambda$ | $532\ nm$ |
| laser beam waist in flow dirn. | $\omega_x$ | $4.7\ \mu m$ |
| laser beam waist in transverse dirn. | $\omega_y$ | $9.4\ \mu m$ |
| sample concentration in tip | $c$ | $\sim 10^{-11}\ M$ |
| pinhole size | $\phi$ | $600 \mu m$ |
| collection objective | | $\times 40$ |
| transit time | $2\delta_t$ | $1.8\ ms$ |
| flow velocity | $v$ | $5.2\ \mu/ms$ |

Fig. 8.1 shows the autocorrelation function, which exhibits a well defined peak at the origin due to the photon bursts from individual molecules. A Gaussian fit to the peak gives a full transit time of $2\delta_t = 1.8\ ms$, as discussed in §6.6. Note that at the origin, superimposed on the Gaussian peak, is another small peak with a duration of $\sim 150\ \mu s$. This feature indicates that, on average, the photon bursts are not completely Gaussian in shape. It may possibly be caused by photodegradation or by other fluctuations for example due to laser noise. Also shown in Fig. 8.1 is the autocorrelation function of the blank, which is obtained by translating the sample

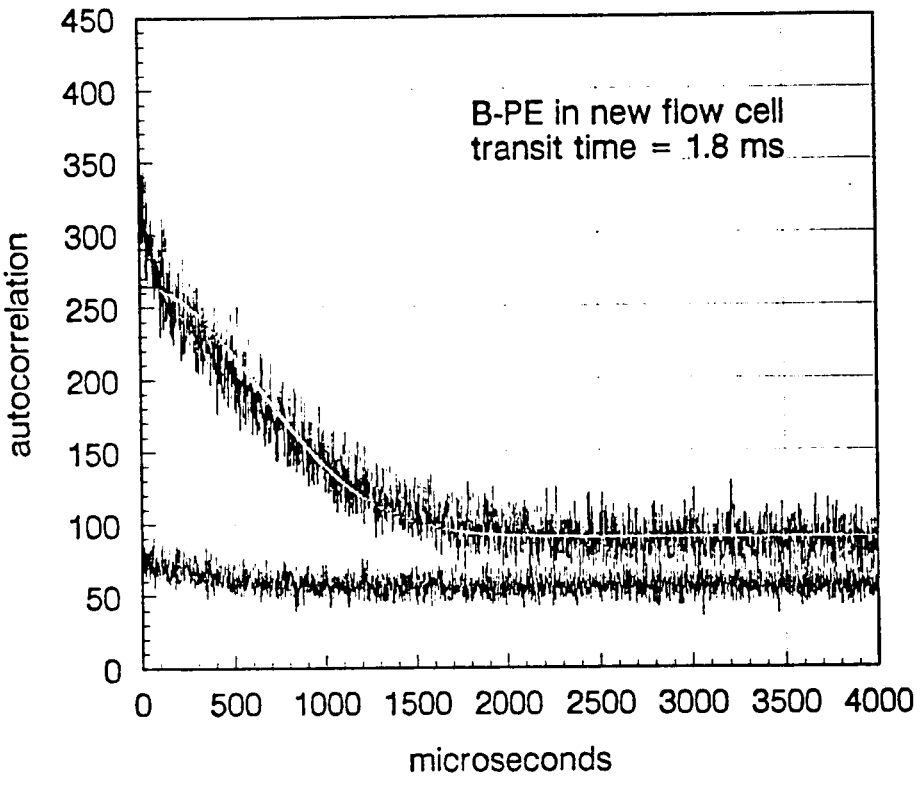



Figure 8.1: Autocorrelations from B-PE and blank in the new flow cell

inlet capillary away from the probe region out of the field of view of the spatial filter. Due to occasional small photon bursts from impurity molecules, the autocorrelation function from the blank is not completely flat.

Fig. 8.2 shows a small segment of the photon burst signal as single B-PE molecules transit across the probe volume. The data is processed using the simple sliding sum algorithm with a bin time of 0.7 $ms$. Since the mean observed rate of molecule detection is $\sim 8\ s^{-1}$ and the transit time is $\sim 1.8\ ms$, the probability that a burst is caused by two molecules passing through the probe volume simultaneously is 0.72 %, as calculated by Eqn. 7.5. In the lower portion of the figure is a segment of the photon burst signal due to the blank. Note that there are no discernible photon burst peaks even though the autocorrelation function from the blank is not completely flat. This demonstrates the increased sensitivity obtained from the time averaged autocorrelation function.

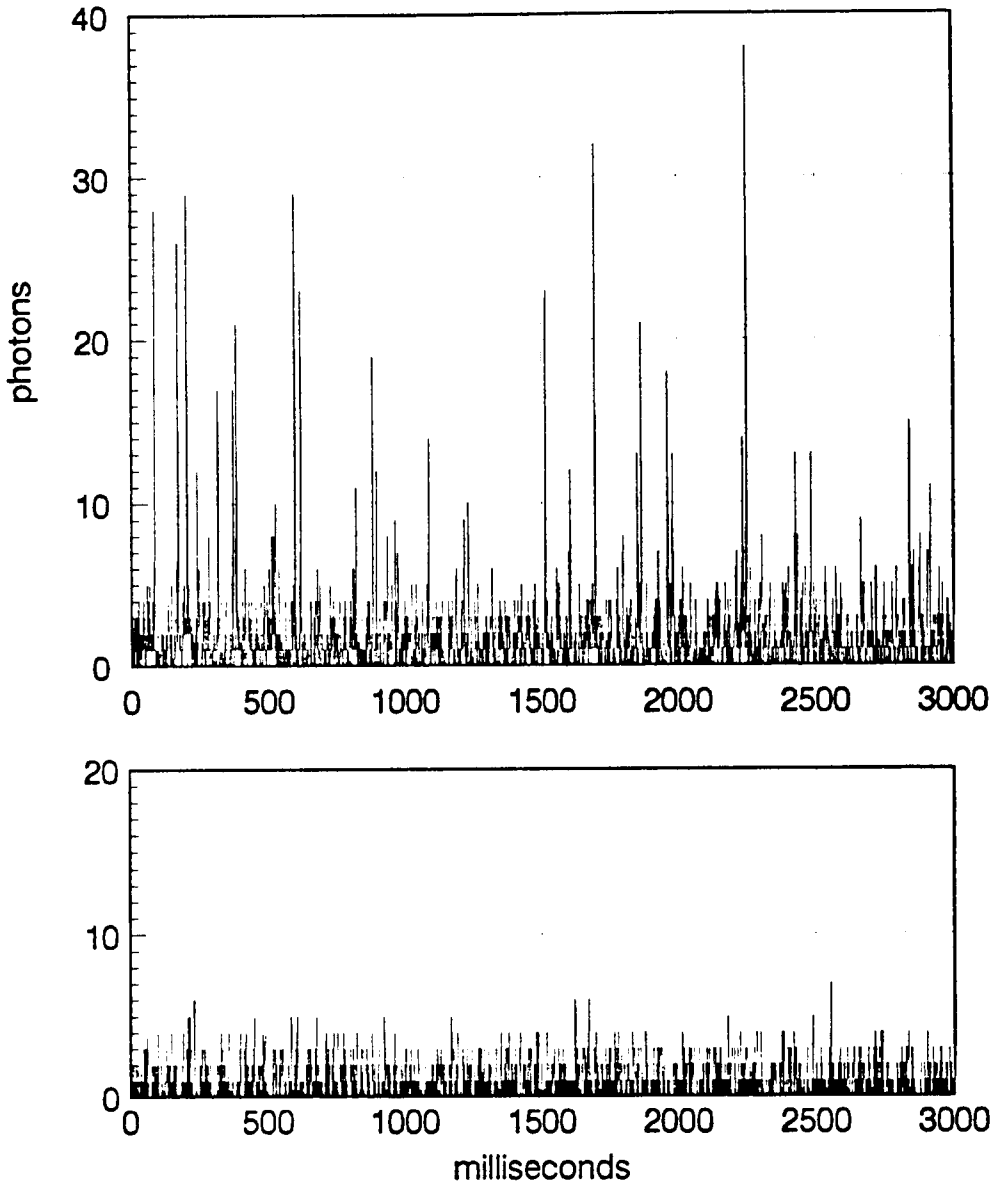



Figure 8.2: Photon burst signals from B-PE and blank in the new flow cell

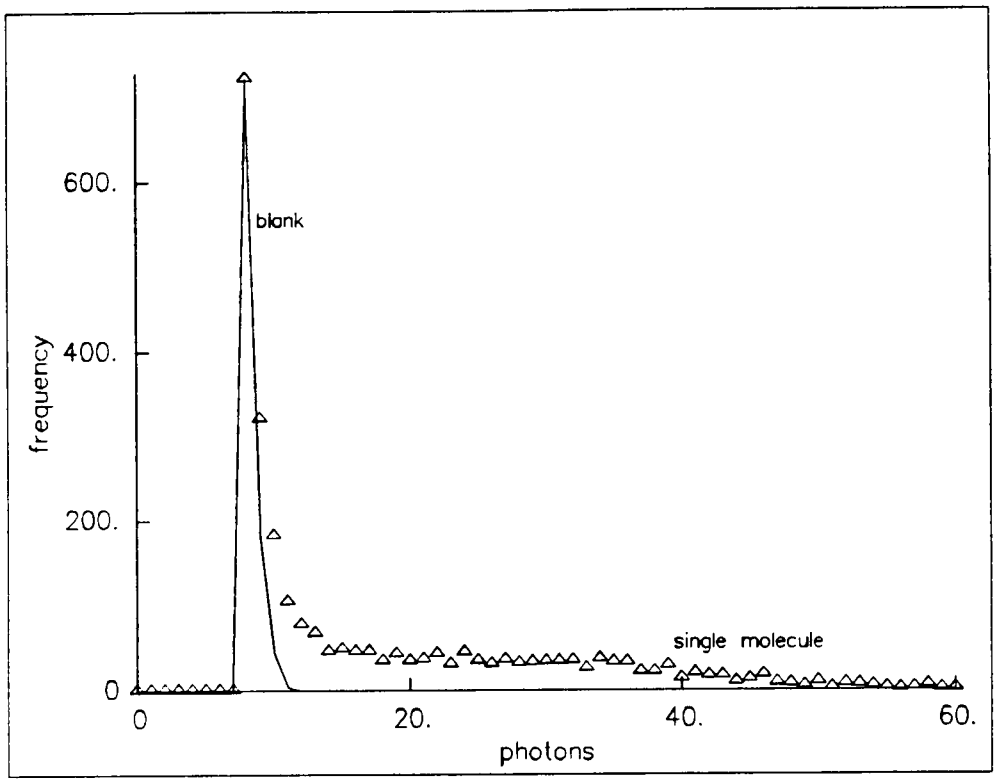



Figure 8.3: Histograms of burst amplitudes from B-PE and blank in the new flow cell

Fig. 8.3 shows the photon burst amplitude histograms from B-PE single molecules and the blank. For the blank solution, the frequency of occurrence of bursts decreases monotonically with amplitude. There are no bursts above 10 photons. When B-PE molecules pass through the probe volume, additional bursts of amplitude greater than 10 appear. The mean amplitude of bursts due to B-PE molecules is $\sim$ 30 photons, although there are occasional bursts of amplitude $> 60$. However, there is no well-developed peak in the distribution, possibly due to imperfect optical alignment and photobleaching. Nevertheless, the majority of the bursts have amplitudes $> 10$ indicating that the molecule detection percentage has been improved significantly compared to prior SMD experiments.

## 8.3 S-101 detection in simple flow cell

In order to detect single chromophore molecules at the high flow velocities required in the new diffusion-limited flow cell, the experimental set up is modified to use excitation by the dye laser at longer wavelengths, a higher numerical aperture collection objective, a larger spatial filter pinhole and spectral filters suitable for S-101. Experiments are first conducted using the simple flow cell, since it is easier to clean and handle than the new diffusion-limited cell. The optimum experimental conditions in the simple flow cell, such as the optical alignment, adjustment of the electronics and laser power, can be referenced when setting up the experiment in the diffusion-limited flow cell.

Some of the experimental results using S-101 in the simple flow cell are shown in the next three figures. The experimental conditions are given in Table 8.2. The experiment is similar to the R6G experiment conducted in the simple flow cell[15] except that the sample flow velocity is $\sim 7.5$ times higher.

Table 8.2: Experimental conditions used for S-101 in the simple flow cell

| | | |
|---|---|---|
| laser power | $P$ | $\sim 12\ mW$ |
| excitation wavelength | $\lambda$ | $585\ nm$ |
| laser beam waist in flow dirn. | $\omega_x$ | $4.8\ \mu m$ |
| laser beam waist in transverse dirn. | $\omega_y$ | $9.6\ \mu m$ |
| sample concentration | $c$ | $\sim 6 \times 10^{-14}\ M$ |
| pinhole size | $\phi$ | $800 \mu m$ |
| collection objective | | $\times 60$ |
| transit time | $2\delta_t$ | $1.6\ ms$ |
| flow velocity | $v$ | $6.1\ \mu/ms$ |

Fig. 8.4 shows the autocorrelation functions due to S-101 molecules transiting through the laser beam and due to clean water. The flat autocorrelation function of the blank solution shows that for an excitation wavelength of 585 $nm$ the pure water gives no impurity fluorescence. A Gaussian fit to the peak of the autocorrelation

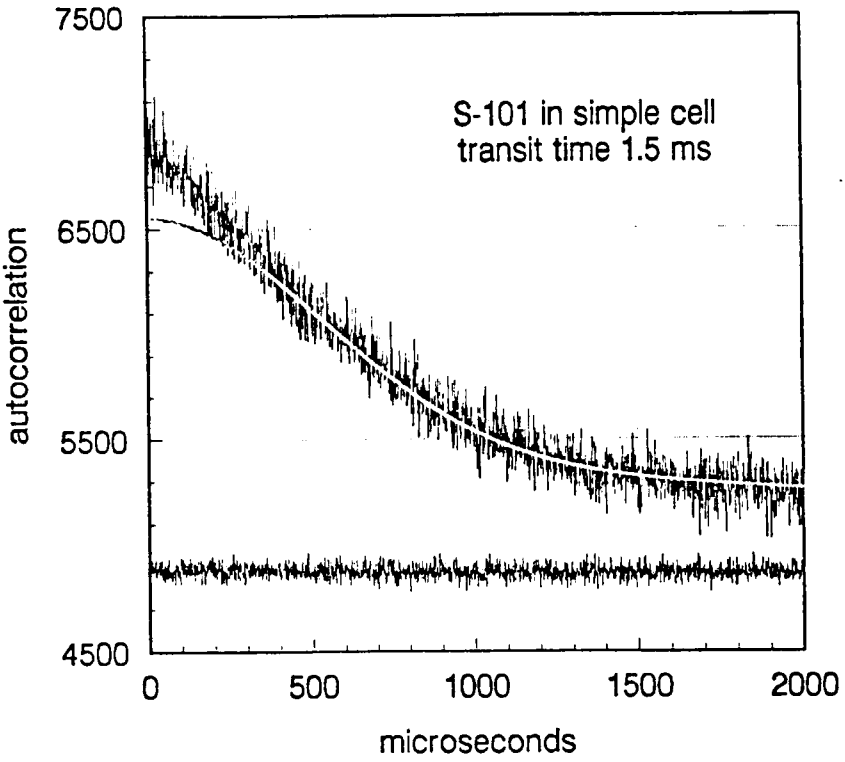



Figure 8.4: Autocorrelations from S-101 and water in the simple flow cell

function of the S-101 solution is also shown in the figure. The fit is used with Eqn. 6.5 to determine the molecular transit time as $2\delta_t = 1.6\ ms$. As in the case of B-PE, there is a non-Gaussian feature near the origin superimposed upon the autocorrelation function. In this case, the duration of the additional peak is $\sim$ 200 $\mu s$. As before, this feature indicates that the mean behavior of the photon bursts is not completely Gaussian in shape, due possibly to photodegradation or laser fluctuations.

The molecule transit time across the laser beam is used to determine the Gaussian weights for the weighted sliding sum filter, as given in Eqn. 6.10. Fig. 8.5 shows a 300 $ms$ segment of the photon burst signal, as calculated by the weighted sliding sum filter. In this segment of data there are 3 bursts greater than 10 photons in amplitude. In the lower part of the figure is another 300 $ms$ segment of data obtained using clean water. In this case there are no bursts greater than 10. The figure shows

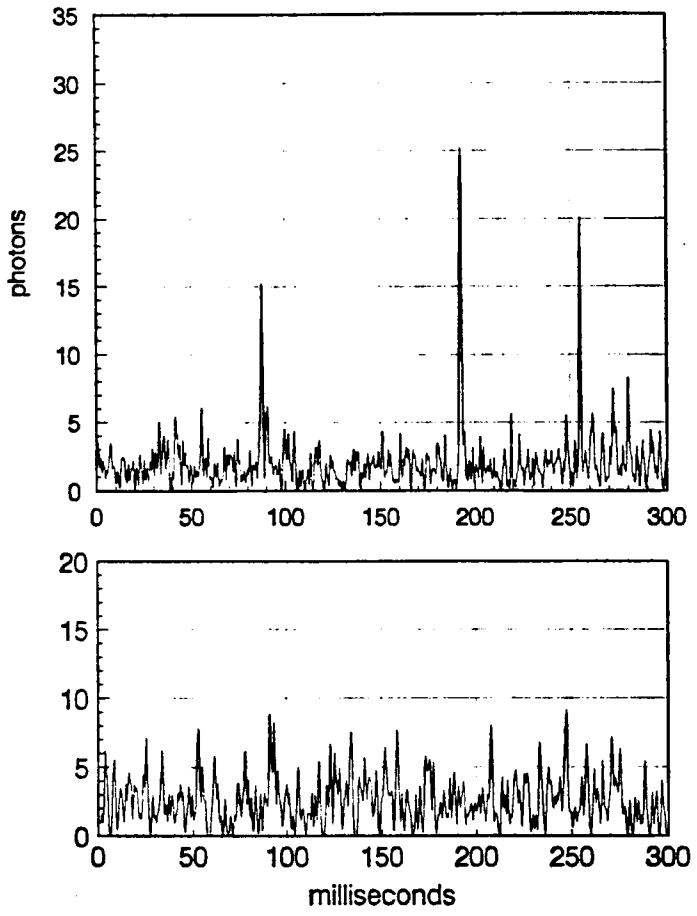



Figure 8.5: Photon burst signals from S-101 and water in the simple flow cell

that single chromophore molecules can be detected at relatively high flow velocities.

Fig. 8.6 gives the histograms of photon burst amplitudes due to S-101 single molecules and clean water respectively. In the simple flow cell, molecules are not confined to a narrow stream but completely fill the cell. Consequently, many molecules pass near the edges of the laser beam and give smaller amplitude photon bursts. As a result, the distribution is monotonically decreasing with photon number.

## 8.4 S-101 in the diffusion-limited flow cell

To cause all molecules to pass through the center of the laser beam, the simple flow cell is replaced by the new diffusion-limited flow cell. Table 8.3 lists the experimental parameters used in the detection of S-101 single chromophore molecules in

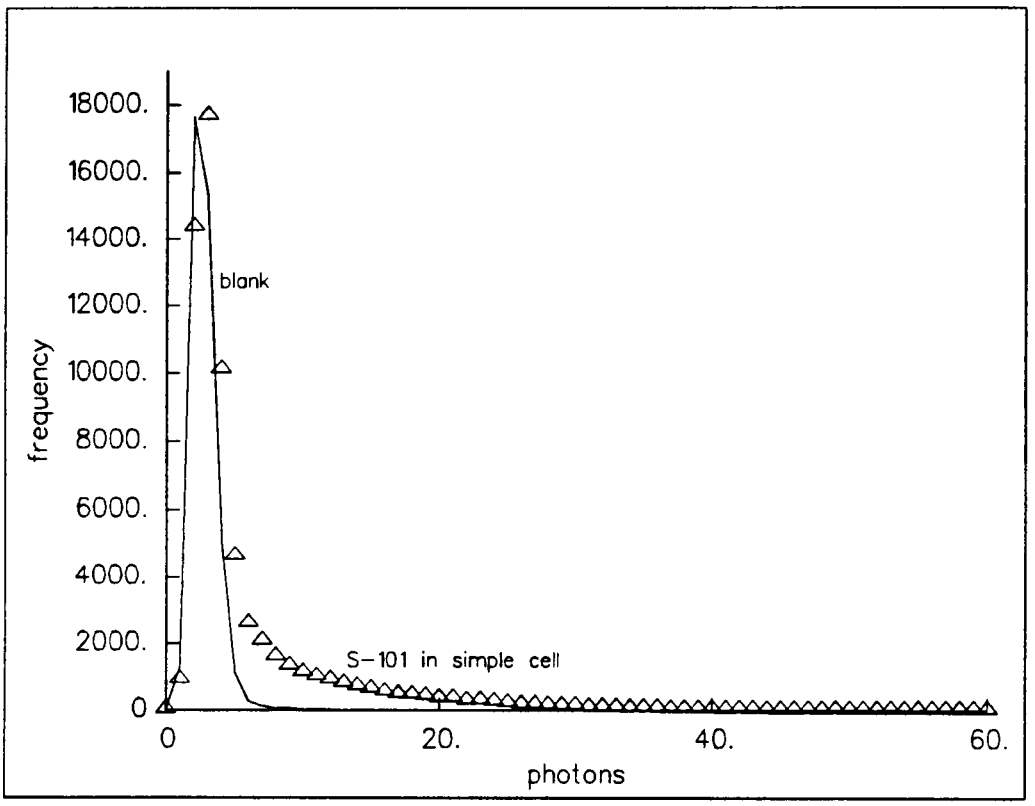



Figure 8.6: Histograms of burst amplitudes from S-101 and water in the simple flow cell

Table 8.3: Experimental conditions used for S-101 in the new flow cell

| | | |
|---|---|---|
| laser power | $P$ | $\sim 14\ mW$ |
| excitation wavelength | $\lambda$ | $585\ nm$ |
| laser beam waist in flow dirn. | $\omega_x$ | $4.8\ \mu m$ |
| laser beam waist in transverse dirn. | $\omega_y$ | $9.6\ \mu m$ |
| sample concentration | $c$ | $\sim 2 \times 10^{-11}\ M$ |
| pinhole size | $\phi$ | $800 \mu m$ |
| collection objective | | $\times 60$ |
| transit time | $2\delta_t$ | $1.0\ ms$ |
| flow velocity | $v$ | $9.6\ \mu/ms$ |

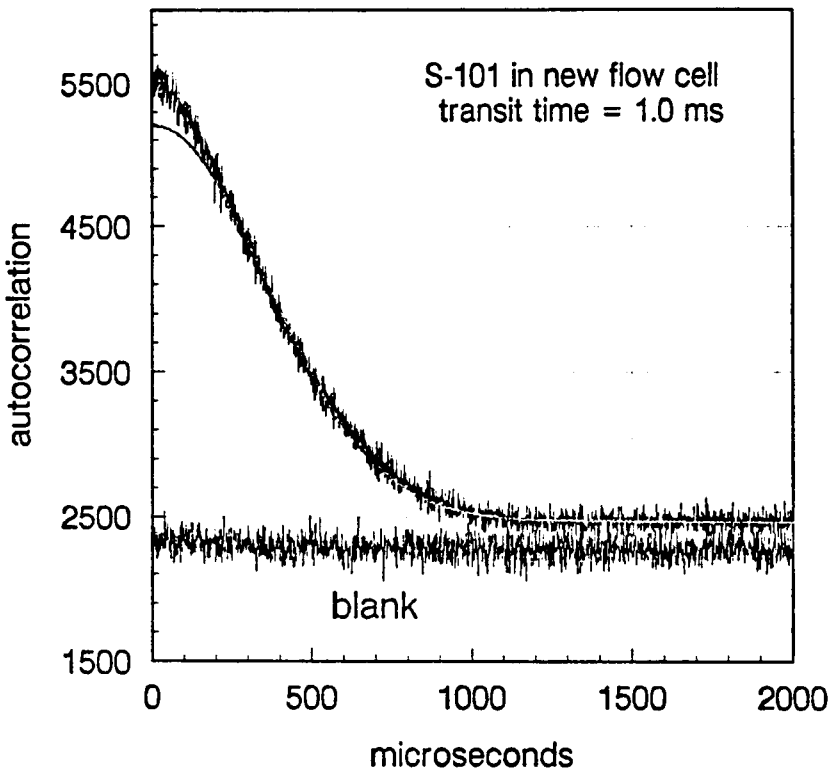



Figure 8.7: Autocorrelations from S-101 and blank in the new flow cell

this case.

Fig. 8.7 gives the autocorrelation functions for S-101 molecules and for blank solution respectively. The blank is obtained by translating the tip of the sample delivery capillary away from the probe region out of the field of view of the spatial filter. An alternative way of obtaining a blank without moving the tip is to stop the molecule delivery by applying negative pressure at the manometer that is connected to the sample delivery capillary, as explained in §4.5. This procedure gives a similarly flat autocorrelation function, indicating that there are no impurity fluorescence bursts. When S-101 molecules are delivered from the tip, the autocorrelation function exhibits a distinct peak. The mean molecular transit time is determined as $2\delta_t = 1.0\ ms$ from the Gaussian fit to the peak, which is also shown in the figure. Again, a non-Gaussian discrepancy is observed near the origin, indicating that the mean shape of the photon bursts is not perfectly Gaussian.

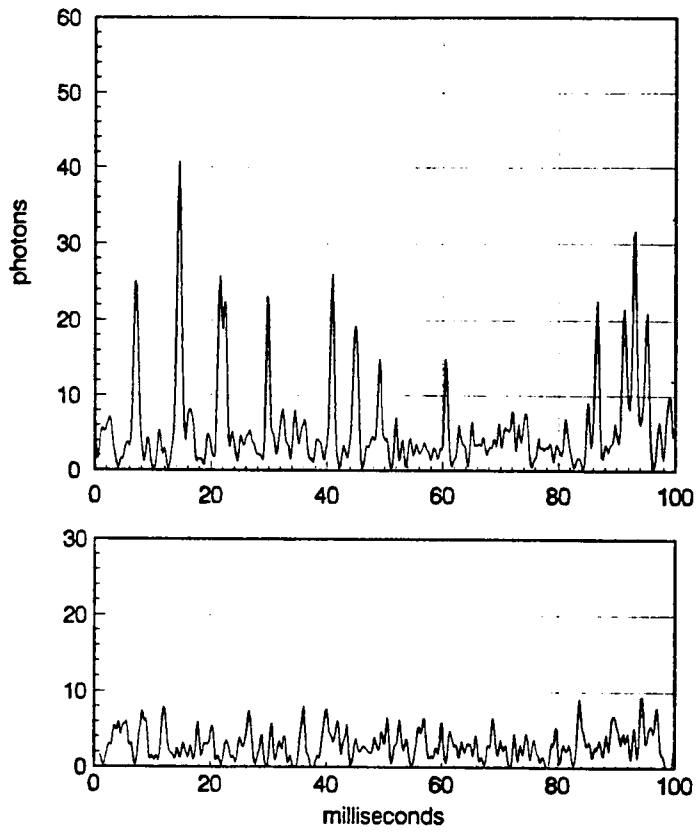



Figure 8.8: Photon burst signals from S-101 and blank in the new flow cell

Nevertheless, to obtain the photon burst signals, the sequence of photon counts is processed by the weighted sliding sum filter, with Gaussian weights given by Eqn. 6.10. Fig. 8.8 shows 100 $ms$ segments of the photon burst signals from the S-101 molecule sample and from the blank, which is obtained by stopping the molecule delivery. In the case of the blank, as shown in the lower part of the figure, there are no peaks of amplitude greater than 10. However, as shown in the top part of the figure, the S-101 sample gives bursts that are clearly above the background noise level. Thus Fig. 8.8 demonstrates that the single chromophore molecule S-101, which has a fluorescence quantum efficiency of only 0.35, can be detected in the new diffusion-limited flow cell at flow velocities as fast as 10 $\mu m/ms$.

Note that there are 13 photon bursts of amplitude $> 10$ in the 100 $ms$ interval

shown in the top part of the figure. By observing a longer period of time than that shown, the mean rate of peaks is observed to be $\overline{M} \sim 150\ s^{-1}$. From Eqn. 7.6, the fraction of bursts that are due to two molecules passing simultaneously through the probe volume is $f \sim 7\ \%$. Thus on average about 1 in every 14 peaks are due to two molecules. The second peak, which has an amplitude of $\sim 40$ photons, is probably due to two molecules. In some cases, the data acquisition can resolve molecules that are in the probe region at the same time. For instance, immediately following the second peak are two peaks in close succession due to two closely spaced molecules. Each of these peaks is separately counted by the software.

Fig. 8.9 shows the histograms of the photon burst amplitudes for S-101 molecules and the blank respectively. For the blank, there are no bursts of amplitude greater than 13 and most bursts are smaller than 11. The mean burst amplitude is $\sim 6$. For the dye molecule sample, there are still bursts of amplitude $\sim 6$, but most bursts have an amplitude greater than 11. Furthermore, there is a distinct peak in the histogram at an amplitude of 18 photons and a distinct valley at an amplitude of 11. For the dye molecule sample, the distribution of burst amplitudes is expected to be due to the sum of a component due to noise fluctuations and a component due to the passage of S-101 molecules. Bursts of amplitude greater than or equal to 11 are attributed to S-101 molecules while bursts less than 11 are attributed to random fluctuations in the noise background. S-101 molecules that generate bursts smaller than 11 are undetected. By extrapolating the component due to molecules only, the fraction of molecules that escape detection is estimated to be $\sim 20\ \%$.

The mean burst amplitude for molecules that pass through the center of the laser without photodegradation is estimated in §7.3 to be 20.2. Since most molecules are not expected to pass exactly through the center of the laser beam and some molecules photodegrade resulting in smaller bursts, the distribution is expected to peak at a value smaller than 20.2. Note that the peak of the distribution in Fig. 8.9 is at 18, in good agreement with the estimation.

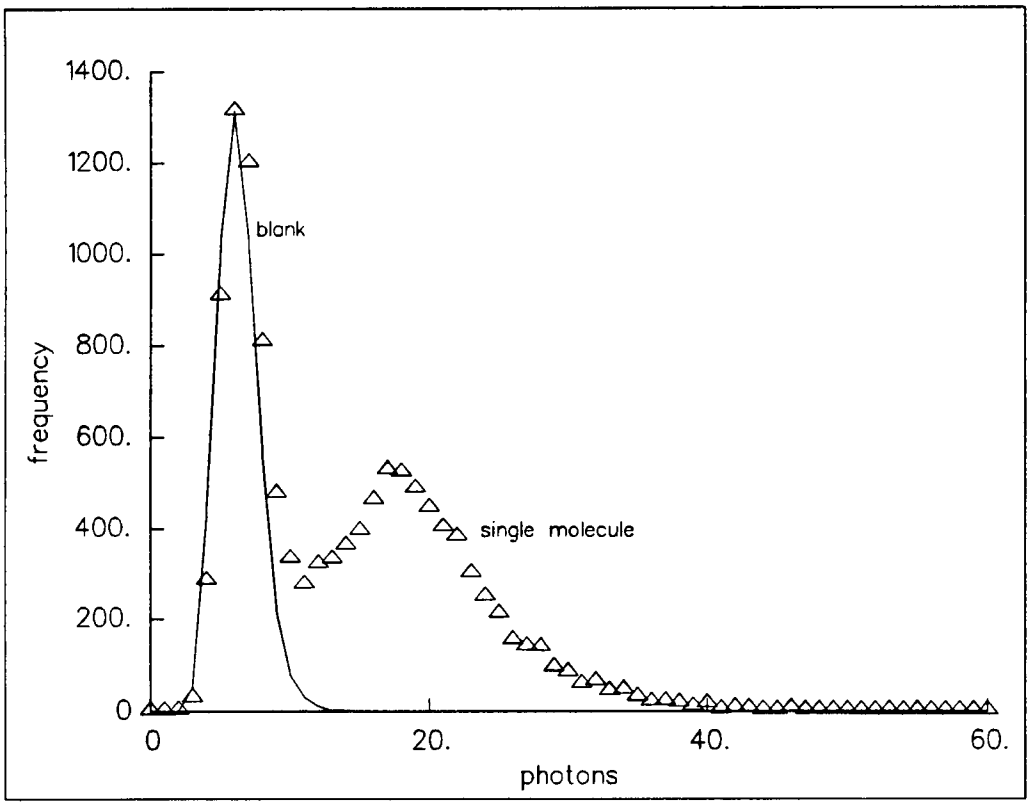



Figure 8.9: Histograms of burst amplitudes from S-101 and blank in the new flow cell

Furthermore, note that the experimental results agree well with the Monte Carlo simulation results, shown in Figs. 7.3 and 7.4. The simulation also predicts that $\sim$ 10 % of the molecules give photon bursts of amplitude less than the threshold due largely to photodegradation. In §7.4, the fraction of molecules that photodegrade while passing through the probe volume is calculated to be $\sim$ 20 %, although some molecules that photodegrade are still expected to give enough photons for detection. These values are in approximate agreement with the estimation that 80 % of the molecules are detected in the experiment. The differences between the experiment and simulation can be attributed to imperfect optical alignment, laser power fluctuations, non-uniform flow conditions near the capillary tip and uncertainty in experimental parameters.

There are substantial differences between the histograms of the photon burst amplitudes in Figs. 8.9 and 8.6. In the simple flow cell and in the previously published experiments[10] only a small fraction of the molecules pass through the laser beam and many molecules pass through the edges of the beam, resulting in a monotonically decreasing distribution of burst amplitudes. However, the new flow cell permits a distribution with a clear peak at a finite number of photons to be obtained, illustrating that most molecules pass through the center of the beam. Although the total detection efficiency is not 100 %, mainly because of photodegradation, the new flow cell gives a significant improvement in overall detection efficiency compared to previous experiments.

## 8.5 Conclusions

This dissertation presents the first experiment demonstrating that single chromophore molecules in aqueous solution can be detected with a high overall efficiency of $\sim$ 80 %. Furthermore, individual molecules are detected using the highest flow velocities to date, $\sim$ 40 times faster than in the prior experiments of Ref. [10]. High flow velocities enable large numbers of molecules to be rapidly counted. Experi-

mental results are presented demonstrating single molecules detection at a rate of $\sim 150\ s^{-1}$, considerably higher than in all previous experiments including Ref. [10] for which the detection rate is only $\sim 1\ s^{-1}$.

In order to achieve the high overall detection efficiency at fast flow velocities, several major advancements are made.

Firstly, about an order of magnitude increase in the light detection efficiency is obtained by collecting light with a microscope objective with a numerical aperture of 0.85 and by implementing a new solid-state single photon detector. The detector, which is a custom-modified single photon avalanche diode, is fully characterized in this work. It is found to give a quantum detection efficiency of $\sim 55$ % and a timing jitter as short as 168 $ps$ when used at the appropriate bias voltage and with light focused tightly onto the active area. The high light detection efficiency allows a higher yield of photons per molecule to be obtained, or alternatively, it allows a detectable photon burst to be obtained in a shorter time.

Secondly, an even higher fluorescence photon count rate is obtained by implementing a custom electronic circuit to discriminate most of the background photons that are caused by Raman scattering from solvent molecules. The higher photon detection efficiency leads to a Raman count rate of $\sim 3 \times 10^5\ s^{-1}$ and causes substantial pile-up and loss of signal throughput at the time-to-amplitude converter, which has a dead time of $\sim 5\ \mu s$. The new custom circuit consists of a two stage coincidence-anticoincidence scheme that is used as a subnanosecond temporal prefilter with a dead time $< 10\ ns$. With the new temporal prefilter circuit, detectable photon bursts from single chromophore molecules can be obtained with transit times as short as $\sim 1\ ms$.

Thirdly, a new sample delivery system is implemented to cause almost all sample molecules to pass near the center of the laser beam for detection. In all prior experiments, only a very small fraction of molecules in the flow cell pass through the laser beam. Quoted detection efficiencies do not refer to the overall efficiency but

only to those molecules that happen to pass through the beam. Furthermore, the signal amplitudes from individual molecules vary widely because many molecules pass through the edges of the Gaussian laser profile. In the new sample delivery system, molecules are introduced through a fine capillary with a sub-micron hole placed immediately above the laser beam. Fast flow velocities are used to reduce diffusional spreading so that molecules pass in a narrow stream through the probe volume. Since most molecules experience approximately the same laser excitation, the histogram of photon burst amplitudes has a distinct peak located clearly above the background noise fluctuations. With the new sample delivery system, individual molecules of sulforhodamine 101, which have a fluorescence quantum efficiency of only 0.35, are detected with an overall efficiency of $\sim$ 80 %. Calculations of the signal strength and the predicted histogram of burst amplitudes obtained from Monte Carlo simulations agree well with the experimental results.

In summary, the goals of this dissertation, to achieve single molecule detection with a faster solution flow rate and high overall detection efficiency, have been realized with a custom user-friendly instrument design.

## 8.6 Suggestions for future work

As discussed in § 8.5, many improvements have been made on different aspects of SMD in solution. Nevertheless, there are several remaining issues that need to be addressed in future work.

For some applications such as single molecule spectroscopy, a larger signal strength is required. The signal strength could be increased by using larger laser powers, since the excitation intensity is presently well below saturation. However, higher photon detection rates cannot presently be achieved due to the SPAD dead time, which is now limited by the passive quenching mechanism employed. Although the dead time of the following TCSPC electronics has been increased by stretching the CFD blocking width to $\sim$ 800 $ns$, if the SPAD dead time were reduced by using an active

quenching bias circuit, the dead time of the electronics could also be reduced. The Monte Carlo simulation predicts that if the SPAD dead time is reduced to $\sim 40\ ns$, higher laser powers can be used to improve the signal strength by at least a factor of 5. Future experiments should investigate the use of actively quenched SPADs for SMD.

Under current experimental conditions, $\sim 20\%$ of the molecules undergo photobleaching. Although some of the photobleached molecules may still emit enough photons for detection, photobleaching is the major cause for molecules to escape detection. Thus, a molecule with better photostability would give a larger overall detection efficiency. Also, molecule species with higher fluorescence quantum yields would further improve the detectability. Experiments to investigate new molecules for experiments in ultrasensitive detection are needed.

Some applications of ultrasensitive detection in solution require SMD of fluorophores that respond to shorter excitation wavelengths. In such cases, one has to pay particular attention to the impurity fluorescence from the solvent. For this reason, further experiments should be conducted to understand and solve the impurity fluorescence that was observed in the attempts to detect R6G using higher laser power excitation at 532 $nm$.

Presently, the exact flow velocity profile near the opening of the tip in the new diffusion-limited sample delivery system is unknown. Calculations on the flow velocity profile near the tip by computational fluid dynamics should be performed to provide a better understanding of the trajectory of molecules. If the flow velocity profile were known, it could be used with the Monte Carlo simulation to provide a better understanding of the experimental set-up.

Also, the sample delivery rate is unknown. It should be determined for typical values of applied pressure by measuring the time interval required to empty a predetermined volume of solution from the capillary. Knowledge of the sample delivery rate would enable a more accurate measure of the overall detection efficiency in the

experiment.

In setting up each experiment, since the new diffusion-limited sample cell is open at the top, it is necessary to carefully balance the input rate of liquid with the output rate to maintain a constant level at the top of the cell. This is a somewhat difficult process. Methods for making the set up of the flow system more user-friendly should be investigated for future experiments. Also, it is uncertain whether the flow rate changes during the course of an experiment. Fluctuations and drifts in the flow rate will change the molecule transit time and hence the histogram of burst heights will be broadened. Thus, experiments should be conducted to test the stability of the flow over the range of flow rates used in SMD experiments.

Even though the capillary tip is a distance of several times the beam waist away from the laser, constant illumination may cause some molecules in the tip to photodegrade before they even leave the tip. Experiments should be performed to test this possibility by observing if the molecule detection rate decreases over an extended period of time. If photodegradation does occur within the tip, methods for constructing a tip from opaque material such as dark glass or for coating the tip with aluminum should be investigated.

Once the problems outlined in this section are successfully addressed, steps should be taken to produce a commercial instrument to make single molecule detection more readily available for ultrasensitive analytical applications.

# Bibliography

# Bibliography


[1] G.S. Hurst, M.H. Nayfeh, and J.P. Young. *Appl. Phys. Lett.*, 30:229–231, 1977.

[2] *Rev. of Modern Physics*, 59(3):615, 1983.

[3] R. Kopelman and W.H. Tan. Near-field optics: Imaging single molecules. *Science*, 262(26):1382, 1993.

[4] E. Betzig and R.J. Chichester. Single molecules observed by near-field scanning optical microscopy. *Science*, 262(26):1422, 1993.

[5] W.E. Moerner and L. Kador. *Phys. Rev. Lett.*, 62:2535–2538, 1989.

[6] W.E. Moerner and T. Basché. Optical spectroscopy of single impurity molecules in solids. *Angew. Chem. Int. Ed. Engl.*, 32:457–476, 1993.

[7] T. Hirschfeld. *Appl. Opt.*, 15:2965–2966, 1976.

[8] D.C. Nguyen and R.A. Keller. *Anal. Chem.*, 59:2158, 1987.

[9] K. Peck and L. Stryer *et al. Proc. Natl. Acad. Sci.*, 86:4087, 1989.

[10] E.B. Shera, N.K. Seitzinger, L.M. Davis, R.A. Keller, and S.A. Soper. *Chem. Phys. Lett.*, 174:553–557, 1990.

[11] S.A. Soper, L.M. Davis, and E.B. Shera. *J. Opt. Soc. Am. B*, 9(10):1761, 1992.

[12] L.M. Davis, F.R. Fairfield, and M.L. Hammond *et. al.* Rapid sequencing of dna based on single molecule detection. *Los Alamos Science*, (20):280, 1992.

[13] S.A. Soper, L.M. Davis, and F. R. Fairfield *et. al.* Rapid sequencing of dna based on single molecule detection. *SPIE: Optical Methods for Ultrasensitive Detection and Analysis*, 1435:168, 1991.

[14] J.H. Jett, R.A. Keller, and J.C. Martin *et. al.* High-speed dna sequencing: An approach based upon fluorescence detection of single molecules. *J. of Bio. Struct. & Dynamics*, 7(2):301, 1989.

[15] L.Q. Li and Lloyd M. Davis. Single photon avalanche diode for single molecule detection. *Rev. Sci. Instrum.*, 64(6):1524–1529, 1993.

[16] J. Weber. *Optics Communications*, 7:420, 1973.

[17] G.R. Haugen, B.W. Wallin, and F.E. Lytle. *Rev. Sci. Instrum.*, 50:64, 1979.

[18] J. Brendel, E. Mohler, and W. Martienssen. *Physical Review Letters*, 66:1142, 1991.

[19] T.S. Lundgren, E.M. Sparrow, and T.B. Starr. *J. Basic Eng.*, 86:620–626, 1964.

[20] F. Zarrin and N.J. Dovichi. *Anal. Chem.*, 57:2690, 1985.

[21] Y.F. Cheng, S.L. Wu, D.Y. Chen, and N.J. Dovichi. *Anal. Chem.*, 62:496, 1990.

[22] T.K. Sherwood, R.L. Pigford, and C.R. Wilke. *Mass Transfer.* McGraw-Hill, 1952.

[23] R.E. Walker and A.A. Westergerg. *J. Chem. Phys.*, 29:1139, 1958.

[24] H.A. Wilson. *Proc. Cambridge Phil. Soc.*, 12:406, 1904.

[25] S.A. Soper, H.L. Nutter, R. A. Keller, L.M. Davis, and E.B. Shera. *Photochemistry and Photobiology*, 57:972, 1993.

[26] R. P. Haugland. *Handbook of Fluorescent Probes and Research Chemicals.* 1992-1994.

[27] *Laser Physik Laser Dye Handbook.*

[28] W.H. Press, B.P. Flannery, S. A. Teukolsky , and W.T. Vetterling. *Numerical Recipes.* Cambridge University Press, 1986.

[29] G.E. Walrafen. *J. of Chem. Phys.*, 40(11):3249, 1964.

[30] B.J. Berne and R. Pecora. *Dynamic light scattering.* Wiley, 1976.

[31] L.M. Davis and L.Q. Li. Monte carlo model of a single molecule counting experiment. *Laser Applications to Chemical Analysis–Topical Meeting*, page 206, 1994.

[32] R.A. Mathies, K. Peck, and L. Stryer. *Anal. Chem.*, 62:1786, 1990.

[33] R.M. Measures. *Laser Remote Sensing – Fundamentals and Applications.* Academic Press, 1984.

# Appendices

## Appendix A

# Fluorescence and competing processes

Fluorescence can be understood as the following. First, a molecule in its ground state absorbs a photon and is promoted to the first electronic excited state. Through fast vibrational and rotational motion, the molecule undergoes a quick relaxation to a lower energy levels within the excited electronic state. Then fluorescence occurs when the molecule decays to the ground state by giving out a photon. Thus, fluorescence is defined simply as the electric dipole transition from an excited electronic state to a lower state of the same multiplicity, usually the ground state.

From Fermi's golden rule, the fluorescence lifetime of the molecule, $\tau_F$, is given by

$$\frac{1}{\tau_F} = \frac{2\pi}{\hbar} \sum_f R_{if}^2 \, \delta(\omega_i - \omega_f), \tag{A.1}$$

where

$$R_{if} = \langle \Psi_{ef} | \widehat{D} | \Psi_{ei} \rangle \langle \Psi_{nf} | \Psi_{ni} \rangle \tag{A.2}$$

is the transition moment integral between initial state $i$ and final state $f$. In Eqn. A.2, $|\Psi_e\rangle$ represents the electronic wave function, $|\Psi_n\rangle$ is the vibrational wave function and $\widehat{D}$ is the electric dipole moment operator. Also in Eqn. A.2, the electronic and the vibrational parts of the wave functions have been separated using

the Born-Oppenheimer approximation. The first integral involves only the electronic wave functions of the system, and the second term, when squared, is the Frank-Condon factor.

For electric dipole transitions, the electronic integral is zero unless the states $i$ and $f$ are of the same spin. This means that electronic absorption in most organic molecules results in the formation of an excited singlet state from the ground singlet state. The Frank-Condon factor dictates the width of the absorption and emission spectra.

Unlike atomic and ionic spectra, the absorption spectra of organic dyes are much broader in width and typically cover several tens of nanometers. The reason is that a typical dye molecule is composed of fifty or more atoms, giving rise to a large number of normal vibrations of the molecular skeleton. Quantum-mechanically this means that transitions can occur from the electronic and vibrational ground state of the molecule to many vibrationally excited states in the first excited electronic state. Furthermore, every vibronic level has a ladder of rotational sublevels superimposed and collisional and electrostatic perturbations, caused by the surrounding solvent molecules, broaden all lines to form a wide continuum. Typically, the fluorescence band appears as a mirror image of the absorption band displaced towards longer wavelength by reflection at the wavelength of the purely electronic transitions. This displacement is the Stokes shift.

The intensity of fluorescence from an excited molecule depends upon the rate constant for spontaneous radiative decay of the excited state. $K_R$. However, competing processes other than fluorescence also affect the fluorescence intensity. Competing processes include internal conversion, energy transfer to other molecules and intersystem crossing to the triplet manifold, which sometimes results in photobleaching or phosphorescence.

Photobleaching occurs when a molecule undergoes an irreversible chemical change that renders it unable to fluoresce. When a molecule photodegrades, it will no longer

emit fluorescence photons even under laser excitation. The photodegradation quantum efficiency is the reciprocal of the mean number of cycles of excitations a molecule can go through before photodegradation. Since photodegradation only occur for an excited molecule, it generally occurs from the metastable triplet manifold.

The various processes that follow absorption can be described by a set of rate equations. The quantum yield of fluorescence, $\Phi_F$, and the fluorescence decay time, $\tau_F$, are defined in terms of the first order rate constants of the different processes.

$$\begin{array}{ll} & RateConstant \\ M + h\nu \longrightarrow {}^1M^* & Ia(absorption) \\ {}^1M^* \longrightarrow M + h\nu & K_R(fluorescence) \\ {}^1M^* \longrightarrow {}^3M^* & K_{ISC}(intersystem\ crossing) \\ {}^1M^* \longrightarrow M & K_{IC}(internal\ conversion) \\ {}^1M^* + Q \longrightarrow products & K_Q(impurity\ quenching) \\ {}^3M^* \longrightarrow products & K_D(dissociation) \end{array} \tag{A.3}$$

The quantum yield is given by

$$\Phi_F = \frac{K_R}{K_R + K_{ISC} + K_{IC} + K_Q} \tag{A.4}$$

and the fluorescence decay time is given by

$$\tau_F = \{K_R + K_{ISC} + K_{IC} + K_Q\}^{-1}. \tag{A.5}$$

The fluorescence lifetime is a measure of the average time delay between the excitation and the emission. Typical organic dye molecules have fluorescence lifetimes of a few nanoseconds.

## A.1 Polarization influence

Both the fluorescence emission from single dye molecules and scattering from solvent molecules depend on the polarization of the excitation. The physical origin of this polarization dependence comes from the electric dipole moment operator in Eqn. A.2. The transition dipole moment operator $\widehat{D}$ is a vector quantity. Hence if a perfectly ordered system such as a fixed molecule absorbs plane-polarized light, the absorption probability is maximized if the transition dipole moment of the molecule is aligned with the polarization. Since fluorescence involves a transition between the same two electronic states as absorption, the emitted light has the same polarization properties as the absorbed light and thus is also plane polarized. Furthermore, the direction of fluorescence emission is anisotropic, since fluorescence cannot be emitted in the same direction as the polarization.

However, in fluid media the molecule is no longer fixed and the analysis is only appropriate if the average rotational relaxation time, $\tau_r$, is long compared with the fluorescence decay time, $\tau_F$. If the opposite is true, i.e., $\tau_F \gg \tau_r$, all polarization anisotropy in the direction of emission is lost through rotational relaxation prior to fluorescence emission. Consequently, fluorescence emission is also isotropic. This is the case for single dye molecules in water, since typically $\tau_r \sim 100\ ps$ while $\tau_F \sim 3\ ns$. Although fluorescence emission is isotropic, the light scattered from solvent molecules is anisotropic because it occurs essentially instantaneously.

# Appendix B

# Rayleigh and Raman scattering

When monochromatic radiation of frequency $\nu_0$ is incident on systems like dust-free, transparent gases and liquids, or optically perfect, transparent solids, most of it is transmitted without change, but, in addition, some scattering of the radiation occurs. If the frequency content of the scattered radiation is analyzed, there will be observed to be present not only the frequency $\nu_0$ associated with the incident radiation but also, in general, new frequencies of $\nu' = \nu_0 \pm \nu_M$. In molecular systems, the frequency shifts $\nu_M$ are found to lie principally in the ranges associated with transitions among vibrational and rotational levels. Such scattered radiation with frequency changes is known as Raman scattering. For liquids, the prominent features in the Raman spectra are of vibrational origin.

The elastic scattering of radiation from molecules without change of frequency is known as Rayleigh scattering. The relative intensity of Rayleigh and Raman scattering depends on factors such as the chemical composition and the direction of observation relative to the direction of illumination. Typically the intensity of Rayleigh scattering is about $10^{-3}$ of the incident radiation intensity while the intensity of strong Raman bands is about $10^{-3}$ of the intensity of Rayleigh scattering. To reduce the background due to light scattering in the SMD experiment, it is important to understand the dependence of Rayleigh and Raman scattering on direction and state of polarization.

Note first that there is an important difference between Rayleigh and Raman

scattering from an assembly of molecules. This arises because Rayleigh scattering is in phase with the incident radiation, whereas Raman scattering bears an arbitrary phase relation to the incident radiation. Thus, for Rayleigh scattering, interference between the scattering from different molecules is possible. In the idealized case of a perfect crystal at absolute zero, this interference results in the Rayleigh scattering adding to zero for all directions except the forward direction. In all other cases, the random motions of the scattering molecules prevent this. The relationship between the Rayleigh scattering from one molecule and the Rayleigh scattering from $N$ molecules thus depends on the state of the matter involved. This dependence is illustrated by the fact that Rayleigh scattering from water is only 200 times stronger than from air, even though water contains 1200 times more molecules in the same volume. For Raman scattering, the random phase factor ensures that scattering from individual centers is always additive.

## B.1 Directional properties

The directional property of light scattering is a consequence of the tensor nature of the polarizability. Here only a simple classical treatment of Rayleigh scattering is considered, although the discussion is equally applicable to Raman scattering and the results are also valid for a quantum mechanical treatment.

Consider a plane, linearly polarized monochromatic wave propagation in the $z$ direction with its plane of polarization at an angle $\phi$ to the $y$ axis, as shown in Fig. B.1. The direction of observation is taken to be in the $yz$ plane, which is also called the scattering plane. For a single molecule located at the origin, the scattered field in the $yz$ plane at angle $\theta$ with the $z$ axis is proportional to[33, page 40]

$$I_{scattered}(\theta, \phi) \propto [\cos^2 \phi \cos^2 \theta + \sin^2 \phi]. \tag{B.1}$$

Since $\frac{\partial I}{\partial \theta} = -\cos^2 \phi \sin 2\theta = 0$ and $\frac{\partial^2 I}{\partial^2 \theta} = -2\cos^2 \phi \cos 2\theta > 0$ for $\theta = \pi/2$, minimum scattering occurs in the direction $\theta = \pi/2$.

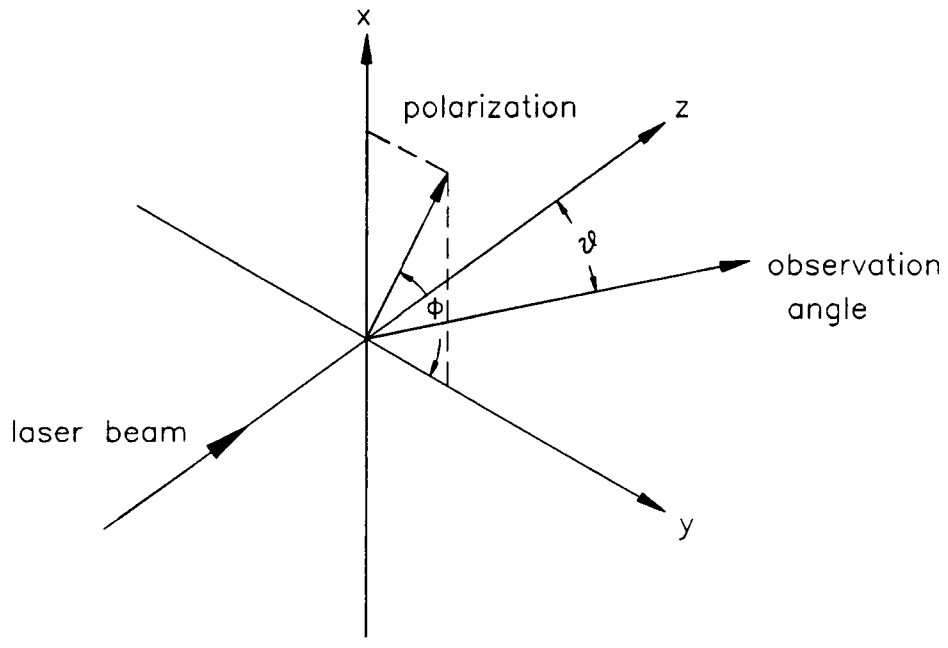



Figure B.1: Directional properties

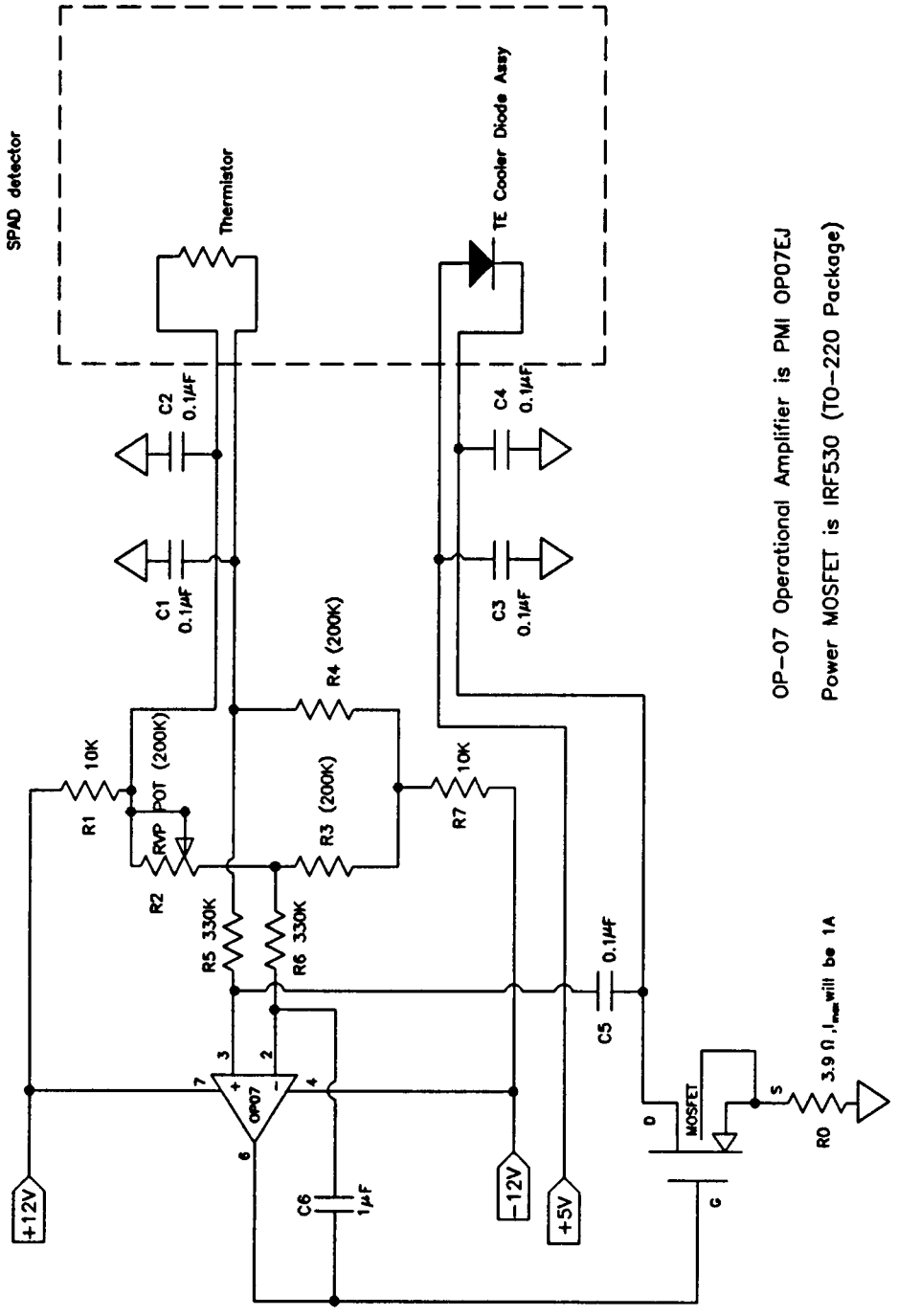
Setting $\theta = \pi/2$ in Eqn. B.1 gives

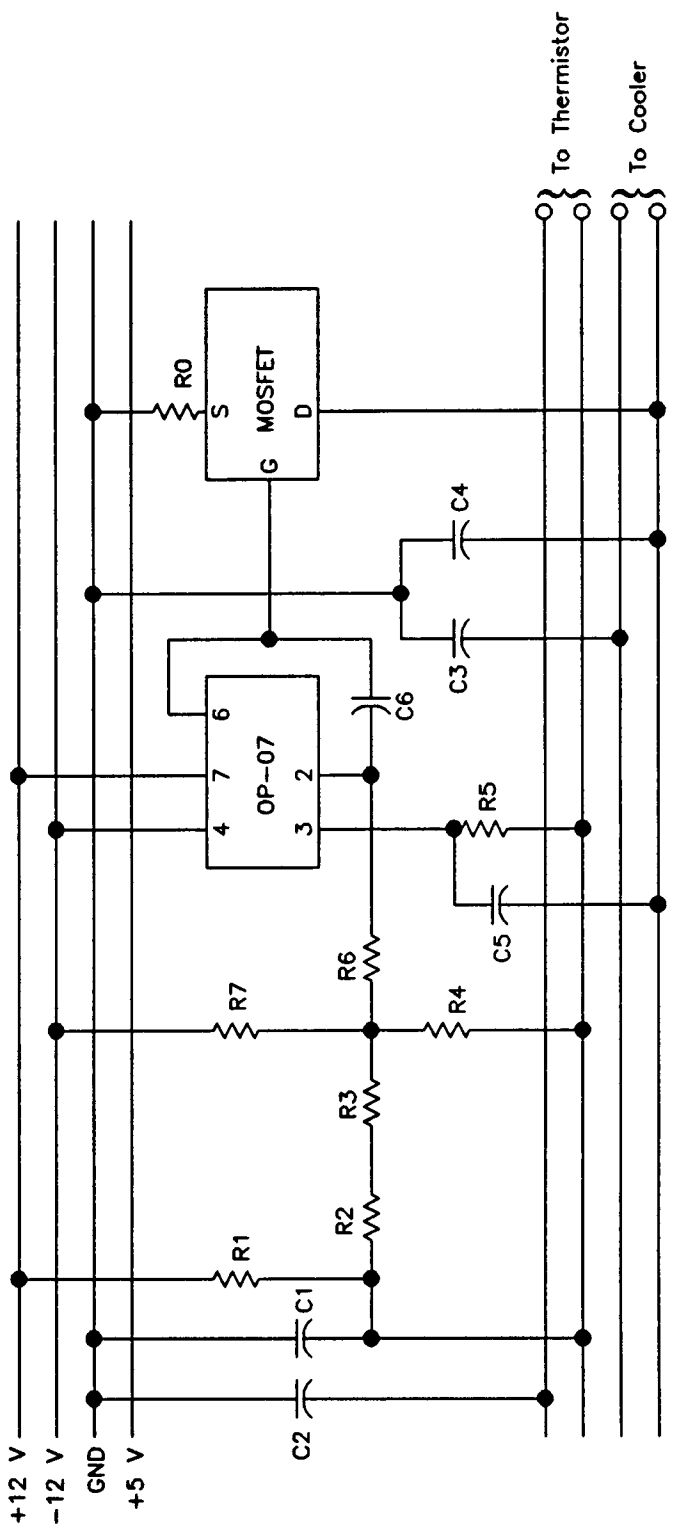
$$I_{scattered} \propto \sin^2 \phi \tag{B.2}$$

This means that for laser propagation along $z$, the scattering signal in the $y$ observation direction still depends on the incoming laser beam polarization direction, $\phi$. To minimize the scattering signal, that the polarization direction of the laser should be adjusted to be $\phi = 0$ or $\pi$. In this case, $I = 0$, but in practice zero scattering situation cannot be achieved because light collection is from a cone rather than at the single angle ($\theta = \pi/2$ and $\phi = 0$).

Appendix C

# Circuit diagram for detector cooler driver


SPAD detector
Thermistor
TE Cooler Diode Assy
C2 0.1μF
C4 0.1μF
C1 0.1μF
C3 0.1μF
R4 (200K)
R1 10K
POT (200K)
RVP
R3 (200K)
R7 10K
R2
R5 330K
R6 330K
C5 0.1μF
OP07
C6 1μF
+12V
-12V
+5V
MOSFET
D
S
G
RO 3.9 Ω, I_max will be 1A
OP-07 Operational Amplifier is PMI OP07EJ
Power MOSFET is IRF530 (TO-220 Package)


+12 V
-12 V
GND
+5 V
C2
C1
R1
R2
R3
R7
R6
R4
C5
R5
C6
C3
C4
R0
S
G
MOSFET
D
OP-07
4
7
6
2
3
To Thermistor
To Cooler

## Vita

Liqiang Li (Richard L.Q. Lee) was born in HeBei, China on January 17, 1963. He completed the degree of Bachelor of Science in mechanical engineering from HeBei Engineering Institute in 1982, specializing in *Metrology and measurement standards in length.* He received a M.S. equivalent diploma in opto-electronics from the Chengdu Electronics Technology University in 1985, specializing in *Laser technology and fiber optics.* He was employed by Guilin Institute of Optical Communications as a research engineer for three years before he came to the United States to study at Virginia Polytechnic Institute & State University in September 1988, where he received the degree of Master of Science in physics in May 1990. The following August, he entered the University of Tennessee Space Institute, and in 1994 he received the Ph.D. degree in physics.