5-2000

Analysis of ultra-sensitive fluorescence experiments

Yuxing Sun

To the Graduate Council:

I am submitting herewith a dissertation written by Yuxing Sun entitled "Analysis of ultra-sensitive fluorescence experiments." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Physics.

Lloyd Davis, Major Professor

We have read this dissertation and recommend its acceptance:

Chris Parigger, Bruce Whitehead, Horace Crater

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Yuxing Sun entitled "Analysis of ultrasensitive fluorescence experiments". I have examined the final copy of this thesis for form and content and recommend 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 Davis, Major Professor

We have read this dissertation
and recommend its acceptance:

Christie Periyp

Ben Whitehead

Hans W. Seals

Accepted for the Council:

Carolyn Minkel

Associate Vice Chancellor
and Dean of the Graduate School

Analysis of Ultra-sensitive Fluorescence Experiments

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Yuxing Sun

May 2000

Dedication

Dedicated to my parents, Guiying Qu and Xiang Sun.

Acknowledgments

I would like to thank Dr. Lloyd Davis who has been my mentor since I arrived at the University, for his patience and guidance to my academic and experimental achievements. I would like to thank Dr. Bruce Whitehead who has been very supportive throughout the course of my studies. I am also grateful to other of my committee members, Dr. Chris Parigger and Dr. Horace Crater, for their advice and assistance over the past many years.

A special thank to the staff at the Center for Laser Application, Newton Wright, Jim HornKohl and Diane Chellstrop. In addition, special thanks to my friends at UT Space Institute, and they are (in alphabetic order): Dr. Ying-Lin Chen, Ivan Dors, Dr. Meng Fan, Dr. Guoming Guan, Pavlina Jeleva, Mehul Kochar, Yongjing Li, Bin Liu, Dr. Dinesh Mehta, Dr. Wenhong Qin, Jianmin Shen, Dr. Yuanji Tang, Leson Wang, Guoping Xia, and Lin Zhu.

This work is partially supported by the National Science Foundation.

Abstract

This work primarily investigates use of the neural network(NN) method to analyze spectral data collected in single molecule detection(SMD) and identification (SMI) experiments. The 2-layer neural networks, with sigmoid as the activation function, are constructed and trained on a set of simulated data using back-propagation and the δ-learning rule. The trained networks are then used for identification of photon bursts in subsequent simulations. Results show that the NN method yields better identification of individual photon bursts than the traditional maximum likelihood estimation (MLE), particularly in cases where the fluorophores have disparate fluorescence quantum efficiencies, absorption cross-sections, or photodegradation efficiencies.

In addition, this work reports several improvements over the prior version of the Monte Carlo simulation program. The improved version considers the fluorescence probability as the convolution of the pure exponential decay function characterized by the fluorescence lifetime and the instrument impulse response function in the experiment. The setting of the time window is then implemented by monitoring the variation of signal and noise. A number of problems have been investigated by using the improved version. In particular, the effects of the number and widths of the bins within the time window on the precision of identification of molecules are studied. The results from the improved version of the simulation show that only a small number of bins (4–8) are required to achieve approximately 90% correct predictions with the NN method. Bin widths chosen in accordance with the intuitive algorithm, or equal bin widths, generally

give better predictions.

Experimental improvements are also reported in this work. In particular, the transit time of BODIBY-TR(D-6116) dye molecules in an SMD experiment was improved to less than 200 μs, and a circuit is implemented to accomplish fast and continuous data collection to be used in future single molecule identification experiments.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the past decade, the detection of individual fluorescent molecules in solution, as well as potential applications in the area of DNA sequencing, has drawn increasing attention from scientists around world, and considerable progress has been made. As the SMD experiment can be performed routinely in many labs, the main focus of the topic has been switched to identify molecules individually by using ultra-sensitive fluorescence spectroscopy techniques.

Ultra-sensitive fluorescence spectroscopy techniques for distinguishing between different fluorophores have applications for several DNA sequencing technologies, including 'rapid' and 'conventional' DNA sequencing schemes, and in gene detection and analysis by hybridization techniques that utilizes DNA micro-chip arrays or single molecule flow cytometry [1]. For example, for 'conventional' DNA sequencing by the Sanger method, the electrophoresis within a capillary or slab gel may be multiplexed by the

use of spectroscopically identifiable fluorescent labels. Lieberwirth *et al.*[2] have recently demonstrated such multiplex dye DNA sequencing in capillary gel electrophoresis by the use of four dyes with disparate fluorescence lifetimes.

In the single-molecule detection (SMD) approach to DNA sequencing currently under development at Los Alamos[3] and elsewhere [4], different fluorescently-labeled DNA bases are to be spectroscopically identified as they are enzymatically cleaved from a single strand of DNA and suspended in a flowing solution. For such applications, unequivocal identification of two different fluorophores with substantially different absorption and emission spectra has been demonstrated in SMD experiments [5]. However, the instrumentation in these experiments was complex in that it involved two laser sources, two single-photon detectors, two associated sets of electronics, and two dichroic beam splitters for combination of the laser beams and for separation of the collected fluorescence. The measurement of the fluorescence lifetimes of individual molecules was also reported in Reference [5], and alternative means for identification of different fluorophores based upon measurements of lifetimes or other spectroscopic properties, such as photo-stability and fluorescence quantum efficiency, were suggested. Such measurements can be advantageous compared to spectral discrimination because the instrumentation can be more cost effective, as it involves only a single excitation source and detector.

Spectroscopic identification of individually detected molecules based upon fluorescence lifetime measurements has now been demonstrated by several groups, as reported for example in references [6, 7, 8, 9]. In the first such experiments [6], Zander *et al.*

showed that Rhodamine 6G (R6G, with a lifetime of 3.79 ns), could be distinguished from the Rhodamine B-zwitterion form (with a lifetime of 1.79 ns). In their experiments, SMD was achieved using a probe volume of a few femtoliters with confocal epi-illumination and up to 320 photons were obtained from molecules diffusing through the probe region. A Maximum Likelihood Estimation (MLE) technique that accounted for the full decay curve and convolution with the prompt was used to estimate the fluorescence lifetime of each photon burst. Enderlein *et al.*[9] first reported the distinction of fluorophores with different lifetimes within a mixture, R6G (with an experimentally determined lifetime of 4.2 ns) and tetramethylrhodamine isothiocyanate (TRITC) (with determined lifetime of 2.5 ns). Their experiments used a much larger picoliter probe region, with the sample solution being injected from a capillary tip placed upstream from the probe region. Their experimental geometry was similar to that of the SMD experiments[10] in our lab at the University of Tennessee Space Institute (UTSI). Enderlein *et al.* used a MLE technique for categorization of the temporal decay spectra of the photon bursts, similar to that described in chapter 3.

In a general sense, all spectroscopic means for categorization of different fluorophores are alike in that the photons are eventually sorted into a number of different bins, resulting in a particular pattern for each type of fluorophore. Statistically speaking, the MLE technique yields an 'efficient' estimator in that it attains the smallest variance possible for the given statistics [13]. This fact is pointed out in much of the literature on spectroscopic identification, with the implication that it would be impossible to im-

prove on such a means for categorization. However, in practical experiments there may be other possible categories that are difficult to statistically calculate and include in the MLE analysis. Also, there may be other factors that influence the statistics in a manner that is difficult to track, but which may be helpful to include in the categorization calculations. For SMD experiments, these include the possibility of photodegradation, intersystem crossing, or multiple molecules passing simultaneously through the probe region. Hence, in the derivation of the MLE method used in Reference [9], the assumption of Poissonian photon statistics within each spectral bin may apply only as an approximation.

By contrast, the neural networks (NN) methods for categorization make no assumptions on the statistics, but develop their own rules during the learning phase. The developed rules are thus purely empirical and can more easily account for complicating factors or alternate categorizations.

Therefore, the primary goal of this work is to investigate a new method of data analysis, namely, the NN method [14], for single molecule identification. The MLE method will also be utilized to analyze the spectral data collected in the simulation, and the prediction results are to be compared with those obtained by the NN method. A procedure of how to incorporate the NN analysis into the real time SMD experiment will be proposed in this work.

In our lab, Li *et al.* first demonstrated the high efficiency single molecule detection experiment with sulforhodamine 101 (S-101) molecules in aqueous solution in 1990

[10], and the transit time of molecules in the experiment was in the order of one millisecond. In 1996, Bunfield *et al.* [11, 12] developed a Monte Carlo simulation of the single molecule detection experiment. The simulation of the individual molecule identification processes by the spectral information was also implemented as different types of molecules were present. The MLE method was used to analyze data in that work. Furthermore, the works [12, 15] showed that better accuracy for identification would be obtained if bins of unequal widths were to be used.

However, the version of the Monte Carlo Simulation of the SMD experiment developed by Bunfeld *et al.*[12], has been found to have two main problems. First, the fluorescence decay profile of the molecule is considered to be a pure exponential function, as characterized by the fluorescence lifetime. In fact, the fluorescence probability is the convolution of the exponential decay and the instrument impulse response function in the experiment. Second, the variation of the signal and noise with selection of the time window setting is not modeled. In this case, the time window was assumed to occupy a portion of the interval between successive laser pulses and the exponential fluorescence decay profiles were simply truncated to the time window. These two issues are inter-related in that realistic temporal profiles for the fluorescence decay profiles and the background noise must be considered in order to correctly model the selection of the time window. Another point to note is that the prior version of the simulation did not consider the difference of the time gate efficiency for dyes with different fluorescence lifetimes. There is a practical limitation in the selection of the time window in that the

number of background scattered photons will substantially increase if the beginning of the time window is selected to be too early. In this case, promptly scattered background photons will overwhelm the collected time spectra of each photon burst, thereby making the different types of molecules impossible to distinguish. When the convolution of the prompt and the selection of the time window are included, the simulation of SMD and SMI would give a useful direction for how to improve the selection of parameters and analysis for the experiment.

Further, the division of the time window into a number of bins, and the selection of the widths of each bin will influence the accuracy and precision of the identification of the dye molecules. Meanwhile, employment of a small number of bins within the time window will reduce the cost of the equipment and speed up the data collection in the experimental setups. Such problems need to be investigated by using the simulation. However, Bunfield *et al.* did not extensively study the effects due to the limitations of the prior version.

Hence, the next task of this work is to extend the previous simulation code to consider the fluorescence probability as the convolution of the exponential decay and the instrument impulse response function in the experiment, and then to model the selection of the time window. Having accomplished these, this work then considers various of bin selection algorithms, and discusses the influences of the accuracy of the identification of single molecules by utilizing different sets of bins. All the investigations based on the improved version of the Monte Carlo simulation will provide a realistic

evaluation of how to increase the correct predictions of individual molecules by changing the experimental setups, and to improve the efficiency of the detection.

SMD in solution offers the prospect for counting of individual molecules within small quantities of sample for ultra-sensitive chemical analysis, and for enabling new approaches for bio-technology applications. Further development of the technology and capabilities is important to facilitate the practical realization of applications. In particular, an increase in the rate at which single molecules may be efficiently detected is necessary for several of the biotechnological and analytical applications. Hence, this work also includes the experimental improvement of fast single molecule detection, and the data acquisition schemes.

In Chapter 2, the main processes in the SMD experiment and main improvements over the prior version of the simulation are explained. Implementation of such improvements are presented. A theoretical discussion of how the selection different number of bins within the time window influences the accuracy and precision of lifetime measurements by Kollner *et al*[16] is introduced, and various bin widths selection algorithms are also described. Chapter 3 gives the definition of the MLE method used in the SMD simulation, several attempts to locate the optimal bin widths are presented. In chapter 4, the NN analysis is introduced, and its categorization power is discussed. The chapter ends with the introduction of software tools used for the simulation, and of the variation of the simulation code for the NN analysis. Chapter 5 presents the results of the simulations and experimental advancements. Finally, Chapter 6 is a summary of this

work, and offers suggestions for the future work.

Chapter 2

Simulation of Single Molecule Detection and Identification

2.1 Overview of the single molecule detection experiment

In the single molecule detection (SMD) experimental setup at UT Space Institute, a synchronously pumped dye laser provides picosecond laser pulses at a nominal wavelength of 585 nm, and at a repetition rate of 76 MHz. The beam is focused to a waist of 3.5 μm in the sample flow cell. Ultra-dilute aqueous dye solution ($\sim$nM) is introduced into a sheath flow in the flow cell through the sub-micrometer opening of an injection capillary, which is placed immediately up-stream from the focused laser beam. Molecules in the sample solution are carried by the flow from the tip of the injection capillary to the probe region. Given the small cross sectional area of the probe region,

the flow speed must be sufficiently fast that very few molecules miss detection due to diffusion.

As each molecule passes through the focused laser beam, it experiences a laser intensity that varies with its position. The laser may cause the molecule to become excited, and this leads to various possibilities, including fluorescence, decay without emission by internal conversion, inter-system crossing to the triplet manifold, and photo-degradation [12]. If the molecule decays to the ground state by fluorescence, a photon is emitted. The emission time is random, with an exponential probability density function and a mean equal to the fluorescence lifetime. The dye molecule may then be excited again, and it usually cycled many times and emits many fluorescence photons during its passage across the probe region.

The experimental apparatus is designed to efficiently detect the fluorescence photons from dye molecules and to discriminate the background from scattered light. A high numerical aperture (N.A. 0.85, 60X) microscope objective collects the photons, and a pinhole (radius 300μm), placed at the image plane of the objective, acts as a spatial filter to discriminate against laser light scattered from the walls of the sample cell. Spectral filters are then used to block Rayleigh scattered light and most of the Raman scattered light from the solvent. Raman light that overlaps the fluorescence band passes through the spectral filters, but is eliminated by an anti-coincidence time gate in the electronics. This is possible because the Raman light is scattered promptly, i.e., with no time delay, whereas most fluorescence photons are emitted with a delay characterized

by the fluorescence lifetime of the dye. Only photons falling within the time window are collected and analyzed, and a burst of such photons is the signature of a molecule passing through the probe region.

Detected molecules may be identified by a variety of spectroscopic measurements. In general, all such identification methods require the collected photons to be sorted into a number of bins, which depend upon the particular spectroscopic measurement. Here, the main focus is on the identification of molecules with different fluorescence lifetimes. In this case, the nanosecond time interval between the laser excitation pulse and the fluorescence photon is used to sort each photon into the appropriate bin. The particular pattern of counts over the bins is dependent on the fluorescence lifetime of the detected molecule. For example, if the molecule has a short fluorescence lifetime, most photons will fall into bins corresponding to short time intervals. Therefore, each photon burst can be categorized by analysis of the pattern. Maximum Likelihood Estimation (MLE) and a Neural Network (NN) method will be used in this work.

2.2 Overview of the simulation

The key physical processes of the SMD experiment, as described in section 2.1, have been incorporated into a Monte Carlo simulation. These have been discussed in detail in references [10, 12], and are listed below.

- The delivery of new molecules from the injection capillary
- The transport of the molecule by the solution flow and by Einstein-Stokes diffusion

- The removal of the molecule due to transport out of the simulation region
- The possibility of excitation of the molecule, calculated with inclusion of polarization and saturation effects
- The possibility of inter-system crossing of an excited fluorophore to the metastable triplet manifold
- The possibility that a molecule in a triplet manifold photodegrades
- The exponentially distributed random time until a molecule in the triplet state relaxes
- The possibility that an excited fluorophore will otherwise decay to the ground state with the emission and subsequent detection of a photon
- The possibility that a photon will be detected due to background or dark noise.
- If a photon is detected, the paralysis of the detector for a number of iterations corresponding to its dead time.
- The possibility that the photon falls within the time window, and if so, the subsequent paralysis of the timing electronics.

Section 2.3 discusses how these processes are implemented into the simulation. The general method for determining which spectroscopic bin each photon falls within, and for recognizing photon bursts, is also discussed. Section 2.4 explains deficiencies in the prior implementation of these aspects of the simulation. In particular, for fluorescence

lifetime measurements, the temporal profile is not a simple exponential decay but must account for the time response of the detector and electronics. Section 2.4 elucidates the corresponding changes to the implementation, including the method of selection of the time gate in order to reduce the noise level. The number of bins to use, methods for selection of the time intervals for each bin, and the choice of bins for a realistic temporal profile are discussed in section 2.5. Section 2.6 examines the improvement in the simulation for determining the begining and end of each photon burst, and for recognizing which photons belong to the burst, particularly in the case of overlapping bursts. It concludes with an explanation of the method for determining which molecule has produced the burst.

2.3 Implementation of the simulation

The simulation deals with the situation where one or two types of dye molecules could be present in the probe region. Photons could be generated by the fluorescence of molecule type A only, molecule type B only, promptly scattered photons from the laser, or dark noise. Eventually, each photon burst is to be categorized as due to the passage of molecule A alone, molecule B alone, both A and B, or background noise due to either promptly scattered photons from laser or dark noise. The algorithms and equations used by the simulation have been reported in detail elsewhere [11, 10] and are explained only briefly below.

The simulation begins with the calculation of frequently used quantities, such as the

probability per time step for detection of a dark noise photon or a promptly scattered photon, the net fluorescence detection efficiency, which accounts for the throughput of the optics and the time gate, the quantum efficiency of the single photon detector, and the fluorescence quantum efficiency of the dyes. The simulation then considers the sequence of events that can occur with each step of time. With each iteration, a new molecule might be introduced from the injection capillary. All molecules present are advanced due to the flow and random diffusion. Then, for each molecule that is not in the triplet manifold, the laser intensity at the current location and the probability of excitation are evaluated and used to stochastically determine if the molecule is excited and subsequently intersystem crosses to the triplet manifold. If so, another uniform random number is generated to determine if the molecule photodegrades. If the molecule does not cross to the triplet manifold, and if the detector is not paralyzed by a previously detected photon, the relative optical collection efficiency from the present location is evaluated. This is used to accumulate the net probabilities for detection of a photon from the particular type of molecule.

After all molecules present are considered, a single uniform random number is generated and compared with the above net probabilties to determine whether a photon is detected, and if so, its origin. In the previous version of the simulation, the net probabilities account for the efficiency of the time gate. In the new version of the simulation, explained in section 2.4.1, the time gate is considered separately. If a photon is detected, both the detector and the electronics will be paralyzed for a number of time

steps according to their corresponding dead times. However, even if either the detector or the electronics are not active, the simulation still considers intersystem crossing and photodestruction.

The type of each photon will not be available in the experiment, but is used in the simulation to determine into which bin the photon will fall. This is done by assigning the photon to the first bin with cummulative probability larger than a uniform random number.

The passage of a molecule will produce a burst of photons. In order to distinguish such bursts, the stream of detected photons, $s(t_n)$, is processed with a digital filter. For optimal filtering, the photon stream is convolved with a weighting function $w(\delta k)$, which has the same profile as that of the pattern to be recognized. In the simulation, the weights are taken to be proportional to relation:

$$W(\delta k) = \sqrt{2}\exp(-2\delta k^2/\delta_t^2), \delta k = -q, ..., q \tag{2.1}$$

where δ_t is the number of iteration steps that correspond to the half-transit time of molecules[11], and q is the integer closest to δ_t. The digitally filtered photon stream, $S(t_n)$, is calculated by:

$$S(t_n) = \sum_{\delta k=-q}^{q} s(t_n + \delta k)W(\delta k) \tag{2.2}$$

Whenever $S(n)$ peaks at a value that exceeds a preset threshold, the simulation must determine which photons contribute to the burst. The method for doing so is

discussed in section 2.4.1. The numbers of photons in each of the bins during the burst are accumulated, and are stored for further data analysis, together with information such as the amplitude and duration of the burst. The burst is later categorized as one of the four types described in the beginning of this section using the MLE and/or the NN methods.

2.4 Problems in the prior version of the simulation

As introduced in chapter 1, the prior version of the simulation has two main problems, i.e., the fluorescence decay profile of the molecule is considered as mono-exponential function, and the selection of the time gate setting is not modelled. The corrections of the problems and corresponding implmentations are described in section 2.4.1. Section 2.4.2 introduces how to process the experimentally collected prompt data, which are to be used as the instrumental response function.

2.4.1 Correction and Implementation

The first approach to improve the simulation is to replace the mono-exponential decay profiles for each molecule by profiles that account for the convolution with the prompt. The curves for the decay profiles are evaluated over the entire time interval T between laser pulses. The resolution used for the evaluation is 2091 channels, which is the same as the number of channels of resolution in the experimentally measured prompt curve, as explained in section 2.4.2.

Let p^{Prompt}[i] denote the normalized probability for detection of a promptly scattered photon within each channel, where $i = 0, \ldots, M-1$, with $M = 2091$. Then, the probability that a fluorescence photon from a molecule of type γ will be detected within channel i, $p^{\gamma}[i]$, is expressed by:

$$p^{\gamma}[i] = \alpha_{\gamma}(\sum_{k=0}^{k<i} p^{Prompt}[k] \exp[-(i-k)T/(M\tau_{\gamma})] + \sum_{k=i}^{M-1} p^{Prompt}[k] \exp[-(i-k+M)T/(M\tau_{\gamma})]), \tag{2.3}$$

where α_{γ} is a normalization factor, defined as:

$$\alpha_{\gamma} = T/[M\tau_{\gamma}(1 - \exp(-T/\tau_{\gamma}))], \tag{2.4}$$

and τ_{γ} is the fluorescence lifetime of molecule type γ. In equation 2.3, the first term describes the usual digital convolution function and the second term accounts for the contribution from the fluorescence photons generated by the prior laser pulse.

To simulate the selection of the time window in the SMD experiment, the probability distributions of all types of photons must be considered. In this work, the time window was at first selected by choosing the region over which both types of fluorescence photons outnumber promptly scattered photons. To this end, the normalized probability curves, $p^{\gamma}[i]$ ($\gamma = A, B$) and $p^{Prompt}[i]$, were scaled to account for factors such as differences in the excitation and fluorescence yields of the dyes, and in the throughput of the optics, etc. The approach used was to scale the fluorescence probabilities with the expected number of fluorescence photons per iteration time step from the corresponding dye

molecules located at the origin (referred to as the brightness, B_γ), and to scale the prompt with the expected number of dark counts per iteration time step, B_p.

In the simulation, the brightness of molecule type j, B_j, is evaluated by:

$$B_\gamma = M_0[1 - exp(-(2P_0\sigma_\gamma)/(3\pi\ R_0 h\nu\omega_{0y}\omega_{0z}))]Q_\gamma \tag{2.5}$$

where M_0 is the number of pulses in each time step, P_0 is the laser power, σ_γ is absorption cross section of molecule type γ, R_0 is the laser repetition rate, ν is the frequency of the laser, and ω_{0y} and ω_{0z} are the laser beam waists. The quantity Q_γ is the net efficiency, defined as

$$Q_\gamma = Q_{E\gamma} L_f L_c Q_s \Phi_C, \tag{2.6}$$

where $Q_{E\gamma}$ is the fluorescent quantum efficiency for molecule type γ, L_f is the throughput of the interference filter, L_c accounts for the transmission loss of the remaining optics, Q_s is the SPAD quantum efficiency, and Φ_C denotes the efficiency of the collection objective, given by,

$$\Phi_C = [1 - \cos(\sin^{-1}((N.A.)/1.33))]/2, \tag{2.7}$$

where $N.A.$ is the numerical aperture of the collection objective and the factor of 1.33 is the refractive index of the aqueous sample.

The scaling factor of the prompt is defined as

$$B_p = (M_0 \beta P_0 V \Phi_C)/(R_0 \omega_{0y} \omega_{0z}), \tag{2.8}$$

where β is an experimentally determined parameter with the units of counts s^{-1} W^{-1} m^{-1}, and V is the effective sample volume evaluated in the initialization routine by Monte Carlo integration. The source of these equations is documented in Chapter 4 of reference [12].

Figure 2.1 shows typical scaled probability distributions of fluorescence from two types of dye molecules A and B (τ_A =4.2 ns, τ_B =2.5 ns) and of background from laser prompt. The points at which the scaled prompt curve crosses the scaled curves for molecule A or B could be used to choose the time window. However, if the left boundary of the time window were to be extended to below the crossing points, the promptly scattered photons would overwhelm the fluorescence photons, and the histogram collected from an individual photon burst would not appear to be similar to the expected distribution pattern from either molecule type. Note that even when the time window is selected such that the fluorescence signal is expected to be marginally stronger than that of the prompt, the photon burst would still generally contain a considerable number of prompt photons superimposed with fluorescence photons, and consequently the collected distribution would differ from that expected from fluorescence alone. For this reason, in practice, it was found to be beneficial to further tighten the time window to reduce the number of background photons in the collected distribution. Figure 2.1

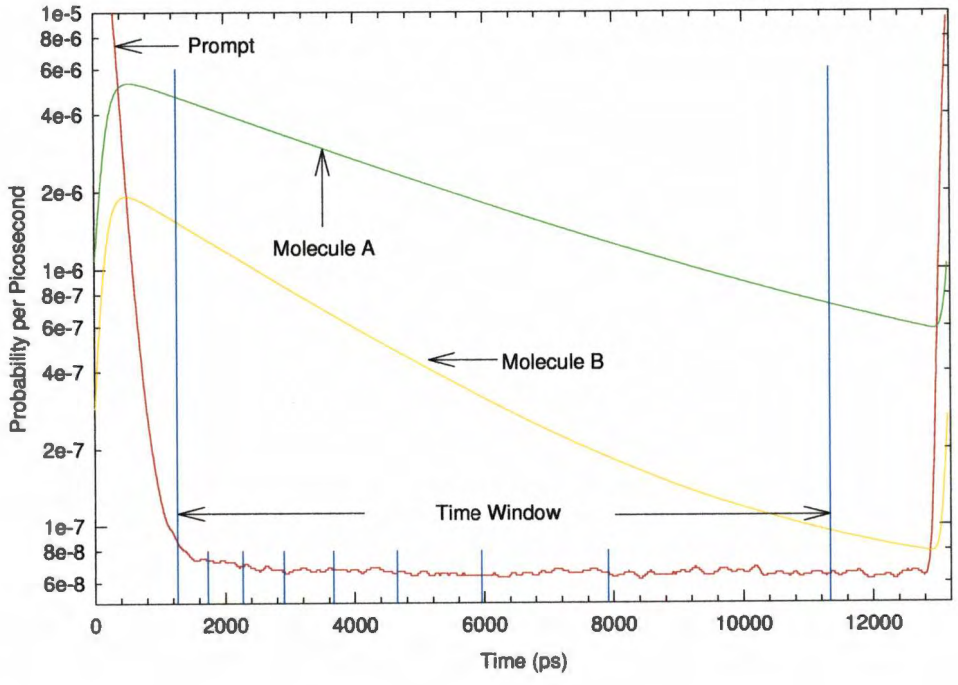


Figure 2.1: Scaled probabilities for photons from prompt, molecule A, and molecule B

shows the time window selected for the simulation results presented in Chapter 5. The use of the tighter time window is consistant with the approach taken in the experimental work reported in reference [9]. Here, the time window was started at approximately 2.1 ns after the peak of the laser pulse and ended at about 1.0 ns before the next pulse, and the period between the consecutive laser pulses was 12.2 ns.

Once the time window is selected, it can then be divided into bins into which photons belonging to each burst are sorted, as described in section 2.1. The bins are usually assumed to be non-overlapping and contiguous although in section 5.9 overlapping bins are considered. The use of a sufficiently large number of the bins is important for acheiving a high percentage of correct identifications of photon bursts, as discussed in section 2.5.1. In addition, the widths of the bins can be chosen (according to various algorithms introduced in section 2.5.2) to improve the precision of identifying molecules.

2.4.2 Prompt Data

This section describes the methods available for obtaining the normalized probability for detection of a promptly scattered photon, $p^{Prompt}[i]$, which is used in equation 2.3. The Monte Carlo simulation may either read in an experimentally collected prompt data file, or generate an ideal Gaussian shaped prompt. The latter choice is useful for debugging and testing the code, while the former generates data that more closely model the actual experimental conditions. The experimentally collected prompt is found to be assymmetric with an exponential tail characteristic of the electron diffusion time of the SPAD detector [24].

The experimental prompt data are obtained with only pure water in the sheath flow system. The time correlated single photon counting instrumentation used for the measurement was the same as that described in [10], with the exception that the prompt data were collected by a multichannel analyzer (MCA) with 8192 channels of resolution. Further, the widths of stop pulses for the time-to-amplitude converter were increased to approximately 20 ns. This enabled the prompt profile to be collected over a duration of longer than the 13.2 ns period T between laser pulses, thereby ensuring that $p^{Prompt}[i]$ could be deduced over the entire time interval T. The MCA accumulated data for several minutes and the data were saved in instrument specific binary format (CHN format) and later converted to an ASCII file for the simulation. Figure 2.2 presents a semi-logarithmic plot of the experimental prompt collected in this way. There are approximately 50,000 counts at each of the peaks. Note that the time calibration (6.3 ps/channel) can be obtained by measuring the number of channels between the peaks in Figure 2.2.

The experimental data exhibits shot noise and a small contribution of detector dark counts, which are negligible compared to the number of the photons in the prompt. The fluctuations due to the shot noise are filtered by digitally smoothing the data. A Gaussian digital filter with a standard deviation of 10 channels is used. Then, in order to extract one period of prompt data from the 8192 channels, the starting point at the half maximum of a pulse on its rising edge and the ending point at the half maximum on the rising edge of the next pulse are located. As a result, the duration between

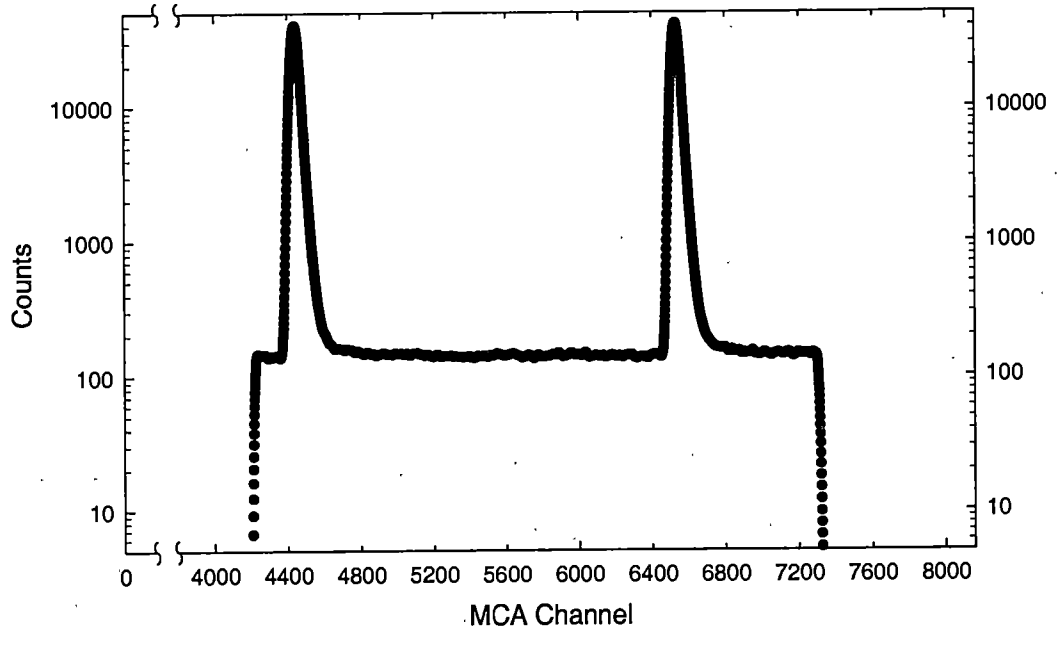


Figure 2.2: Experimentally collected prompt

the peaks is 2091 channels, and the corresponding $p^{Prompt}[i]$ are used to calculate the convolution in section 2.3. A FORTRAN program was written to perform the data analysis, namely to read in an experimentally collected prompt data stored in instrument specific binary format, smooth the data, truncate over one period, normalize, and write the resultant $p^{Prompt}[i]$ as an ASCII file for use by the simulation. In addition, a routine was written for determining the width of the peak at any specified fraction of height, as this information provides a helpful comparison between the fluorescence decay profile of dye and prompt profile. The full width at half maximum of the prompt peaks in Figure 2.2 is 309 ps, which is the same as that reported in reference [24] for an actively-quenched SPAD.

2.5 Bin selection algorithms

Section 2.4 discussed modifications to the simulation to account for the convolution with the instrument response function and to appropriately select the time window. This section discusses the next step needed for spectropic identification of detected dye molecules, namely the selection of bins within the time window. The questions that need to be answered are:

- how many bins are needed to correctly distinguish different molecule types with a certain degree of accuracy?
- how should the width of each bin be selected to give best accuracy?

These questions are investigated by using Monte Carlo simulations, with the results presented in Chapter 5. In an experiment, the number of bins to be used would be determined by the complexity of the hardware and the processing time. Although the number of bins inside the time window could be chosen to be large, the implemention of hardware and data analysis for this would be more costly, and computationally slower. Section 5.7 discusses hardware that would result in a small number of bins of arbitrary width.

A related problem of how the mean variance in the determination of the fluorescence lifetime of a single detected molecule depends on the number of bins has been addressed using an analytical theoretical approach by Kollner *et al.*. [16]. Section 2.5.1 reviews this analysis, which predicts that there is little improvement in the precision of determined lifetimes as the number of bins is increased above about 4. Accordingly, little improvement in the accuracy of identification of molecules with different lifetimes is expected if the number of bins is increased above about 4. The analysis of Kollner *et al.*. pertains to bins of equal width, but when the number of bins is small, experimental hardware for bins of arbitrary width can be implemented. The question of how to best select the bin widths is addressed in section 2.5.2. Several procedures for choosing bin widths are also introduced, and details of the implementation are given in section 2.5.3.

2.5.1 Bin number theory

A typical data file from a SMD experiment consists of many sets, each of which corresponds to a photon burst. Each data set contains the total number of photons within

each bin for the burst. Kollner *et al.*. considers the analysis of the data set to estimate the fluorescence lifetime of the dye molecule by a general statistical estimator. How the estimator analyzes these data and determines the fluorescence lifetime, and precisely how it works, are some of the important questions discussed in Kollner *et al.*'s paper [16].

As stated in the Rao-Cramer theorem, the standard deviation or variance of an estimate cannot be smaller than a well-defined limit. Estimators reaching this limit are said to be efficient. In addition, for sufficiently large signals, the distribution of the parametric estimates for efficient estimators is normal with a covariance matrix equal to the inverse of the Fisher-information matrix, F. For the problem of fluorescence lifetime estimations, F is defined by

$$F_{lm} = \sum_{j=1}^{N} (1/y_j)(\partial y_j/\partial \alpha_l)(\partial y_j/\partial \alpha_m), \tag{2.9}$$

where y_j is the expected number of photons in bin j, $j = 1, ..., N$, as predicted by the model function y. The function y depends on the parameters α_l, $l = 1, ..., L$, where in this case, one of the α_l refers to the lifetime.

As single photon counting involves distributing a total of K photons, $K = \sum n_j$, over N bins, the for a particular distribution $\{n_j\} = n_1, n_2, ..., n_N$ is the multinomial distribution, as discussed in reference [12]

$$P(\{n_j\}, \alpha_1, \alpha_2, ...) = (K!/n_1!n_2!...n_N!)p_1^{n_1}\ p_2^{n_2}...\ p_N^{n_N}, \tag{2.10}$$

where n_j is the number of photons observed in bin j, and p_j is the probability that a photon will fall into bin j. As $y_j = Kp_j$, equation 2.9 becomes

$$F_{lm} = K \sum_{j=1}^{N} (1/p_j)(\partial y_j/\partial \alpha_l)(\partial y_j/\partial \alpha_m). \tag{2.11}$$

Note that equation 2.11 is linear in K and because the covariance matrix is equal to the inverse of the matrix F, the variance of τ given that K photons are observed equals to $1/K$ times the variance given that 1 photon is observed, i.e.,

$$var_K(\tau) = (F)^{-1}_{\tau\tau} = (1/K)[F(K=1)]^{-1}_{\tau\tau}. \tag{2.12}$$

If $var_1(\tau)$ denotes the variance of τ when the signal contains only one photon, the required number of photons for a desired variance $var_D(\tau)$ is

$$K \geq var_1(\tau)/var_D(\tau). \tag{2.13}$$

Consider a simple case where only a constant background and a pure mono-exponential fluorescence decay are present. The decay is characterized by the fluorescent lifetime τ_γ of the particular molecule, γ. If a detected photon originates from dark noise, then the arrival time with respect to a laser excitation pulse is random and the probability

density for a photon to fall in the j-th bin is

$$P_j^d = \omega_j / (\sum_{j\prime=1}^{N} \omega_{j\prime}), \tag{2.14}$$

where ω_j (i=1,2,..., N) is the width of the j-th bin. If a detected photon originates from fluorescence, the probability will be

$$P_j^\gamma = \int_{t_{j-1}}^{t_j} \exp(-t/\tau_\gamma)dt / \int_0^T \exp(-t/\tau_\gamma)dt, \tag{2.15}$$

where, for simplicity, the time window is assumed to start at 0 and end at T, the temporal bin j starts and ends at t_{j-1} and t_j, respectively, and j= 1,2,...,N. Substituting

$$\sum_{j\prime=1}^{N} \omega_{j\prime} = t_N - t_0 = T \tag{2.16}$$

into equation 2.15 gives

$$P_j^\gamma = \frac{\exp[-(t_j - t_0)/\tau_\gamma](\exp[(t_j - t_{j-1})/\tau_\gamma] - 1)}{1 - \exp[-T/\tau_\gamma]}. \tag{2.17}$$

Substituting P_j^γ into equations 2.11, 2.12, and 2.13 yields

$$var_K(\tau, T, N) = \frac{\tau^2 N^2}{K} (\frac{\tau}{T})^2 \frac{1 - e^{-T/\tau}}{e^{T/(N\tau)}(1 - e^{-T/\tau})/(e^{T/(N\tau)} - 1)^2 - N^2/(e^{T/(N\tau)} - 1)} \tag{2.18}$$

Figure 2.3 plots $var_1(\tau, T, N)/\tau^2$ calculated using equation 2.18 for selected values of

N.

Equation 2.18 can also be used to determine the number of photons required for a given precision. For example, suppose the desired precision, $var_D(\tau)$, is 10%, the lifetime, τ, is 2.5 ns, the time window, T, is 10 ns, and the number of bins, N, is 64. Substitute these numbers and $K = 1$ into equation 2.18, and the result for $var_1(\tau)$ into equation 2.13 yields $K \geq 90$, i.e., at least 90 photons are required to achieve the desired precision.

It is impossible to choose a time window larger than the period between laser pulses, and the expected range of T/τ is around 2 to 5. If T/τ is optimally selected, the number of photons required to achieve 10% precision is found to be about 48. The example demonstrates that the experimental setup needs to be configured so that the number of photons within each burst is as large as possible.

The equations can also be used to optimize experimental conditions such as the width of the time window and the number of bins by finding the minimum in the plot of $var_1(\tau)/\tau^2$. From figure 2.3, if a large time window is selected, say T/τ is greater than 100, and the number of bins is less than 16, most of the photons will fall into the first bin and the precision of the measurement of τ will be poor. Conversely, if T/τ is properly selected, little improvement in precision is gained beyond about 4–8 bins. Similarly, if molecules with different known lifetimes are to be identified from the distribution of photons, little improvement in the accuracy in the identification is expected if the number of bins is greater than about 4–8.

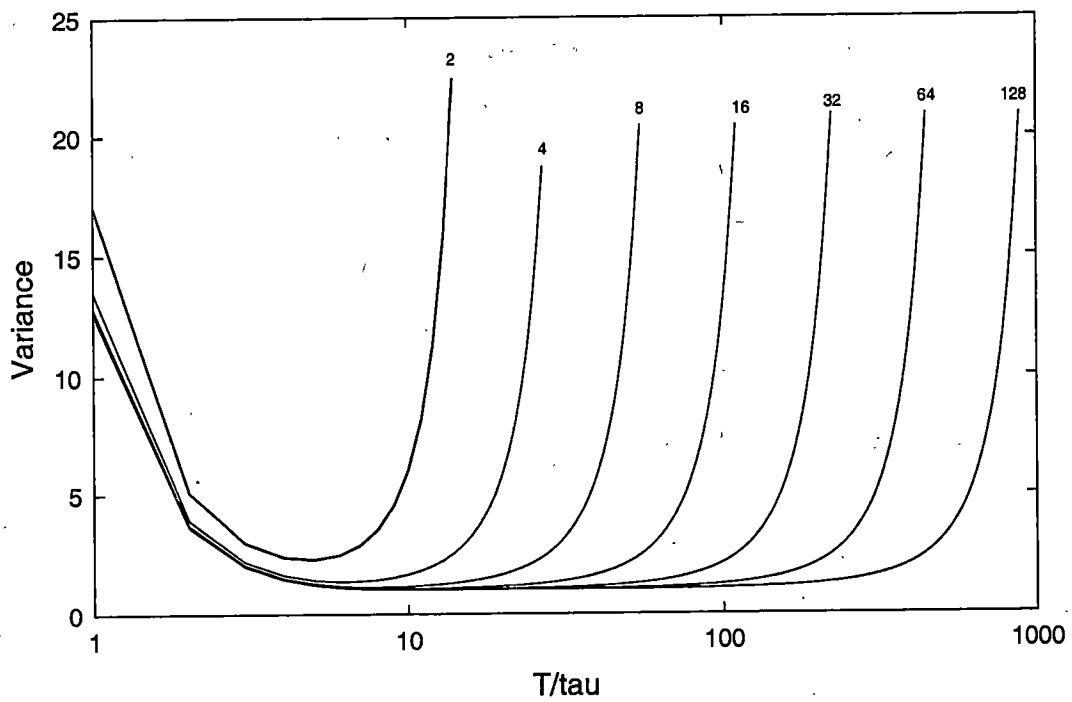


Figure 2.3: The normalized variance τ, var_1/τ^2, for mono-exponential decays as a function of T/τ for different number of bins.

2.5.2 Intuitive algorithm for bin width selection

This section considers how to select the widths of the bins in order to give better identification of molecules with different fluorescence lifetimes. In particular, an intuitive method of bin width selection, which was introduced in [12], is explained. Section 2.5.3 describes how the intuitive method is implemented. Section 2.5.4 explains an alternate bin selection method, and results in chapter 5 indicate that this method gives improved molecule identification.

For bins of equal widths, molecules with different lifetimes tend to distribute photons similarly, namely, with a monotonically decaying distribution with most photons falling into the first few bins. Intuitively, one would expect molecules with different lifetimes to be more easily distinguishable if they yield dissimilar distributions. The distribution for a molecule with a long lifetime may be made to increase if the widths of the earlier bins are made narrower than those of the later bins. The increase of the widths of the bins must be chosen such that the distribution for molecules with the shorter lifetime is still decreasing. Ideally, the widths of the bins should be chosen so as to make the expected distributions for the different types of molecules as dissimilar as possible.

To this end, in reference [12], the widths of the bins were selected so as to yield a flat distribution if the lifetime of a molecule were equal to the algebraic mean of the two known lifetimes that were to be distinguished. That is, if the molecule were to have a fluorescence lifetime equal to

$$\overline{\tau} = (\tau_A + \tau_B)/2, \tag{2.19}$$

where τ_A and τ_B are the known lifetimes of molecules type A and B, then the expected distribution should be flat. Figure 2.4 illustrates the selection of 8 bins for the case of distinguishing between molecules with lifetimes of τ_A =4.2 ns and τ_B = 2.5 ns.

The equations for derivation of the widths of bins neccesary to yield a flat distribution for a hypothetical molecule with fluorescence lifetime τ are given incorrectly in reference [12], and hence are rederived below. Consider the case of N bins within the time window. If background and convolution with the prompt are ignored so that the probability density decays mono-exponentially, then, from equation 2.17, the probability for bin j should be

$$P_j^{\overline{\gamma}} = \frac{\exp[-(t_j - t_0)/\overline{\tau}](\exp[(t_j - t_{j-1})/\overline{\tau}] - 1)}{1 - \exp(-T/\overline{\tau})} = 1/N. \tag{2.20}$$

Substituing

$$\omega_j = t_j - t_{j-1} \tag{2.21}$$

and

$$\zeta = [1 - \exp(-T/\overline{\tau})]/N, \tag{2.22}$$

into equation 2.20 and rearranging equation 2.20 yields

$$\exp(\omega_j/\overline{\tau}) = 1 + \frac{\exp[(t_j - t_0)/\overline{\tau}]}{N}(1 - \exp(-T/\overline{\tau})) = 1 + \zeta \ \exp[(t_j - t_0)/\overline{\tau}]. \tag{2.23}$$

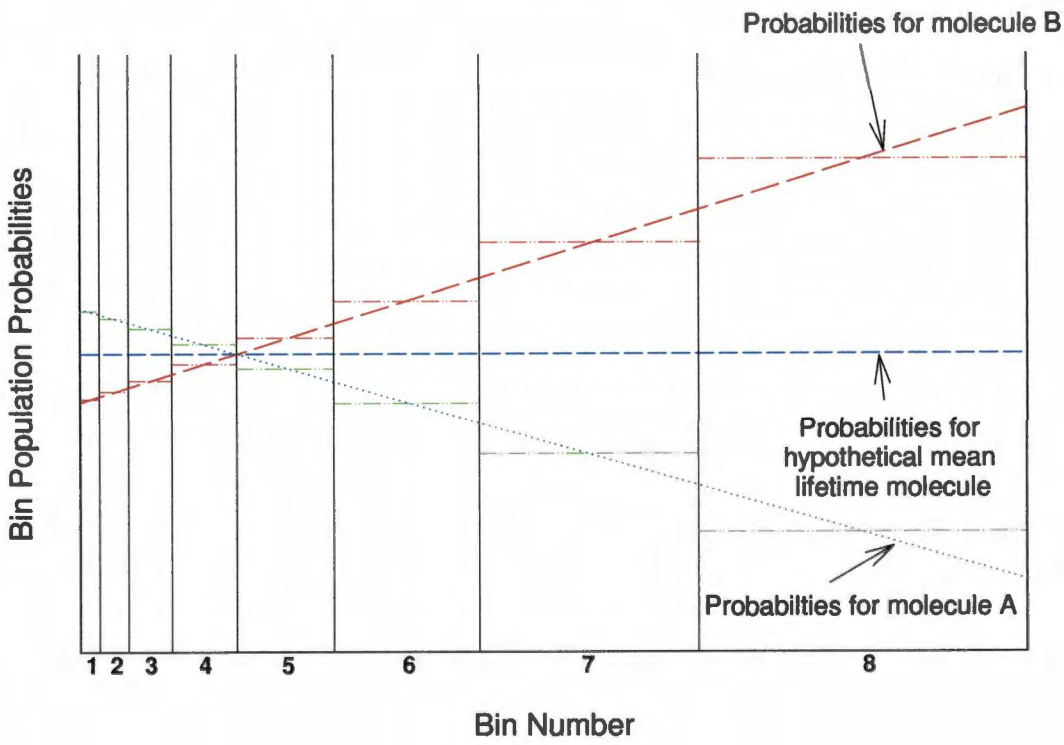


Figure 2.4: Representation of bin widths which produce the most dissimilar distributions for molecule types A and B

This may be solved for each ω_j as follows:

$$\begin{aligned}
&j=1, \quad \exp(\omega_1/\overline{\tau})=\tfrac{1}{1-\zeta}, \quad \omega_1=-\overline{\tau}\,ln(1-\zeta)=\overline{\tau}[ln(1)-ln(1-\zeta)],\\
&j=2, \quad \exp(\omega_2/\overline{\tau})=\tfrac{1-\zeta}{1-2\zeta}, \quad \omega_2=\overline{\tau}[ln(1-\zeta)-ln(1-2\zeta)],\\
&j=3, \quad \exp(\omega_3/\overline{\tau})=\tfrac{1-2\zeta}{1-3\zeta}, \quad \omega_3=\overline{\tau}[ln(1-2\zeta)-ln(1-3\zeta)],\\
&\quad\vdots\\
&j=N, \quad \exp(\omega_N/\overline{\tau})=\tfrac{1-(N-1)\zeta}{1-N\zeta}, \quad \omega_N=\overline{\tau}[ln(1-(N-1)\zeta)-ln(1-N\zeta)].
\end{aligned} \tag{2.24}$$

The summation of all the widths in equation 2.24 should be T, i.e.,

$$\sum_{i=j}^{N}\omega_j=-\overline{\tau}\ln(1-N\zeta)=T. \tag{2.25}$$

Section 2.5.3 discusses the extension of the above described bin selection algorithm to the case where the decay profile is not a pure exponential but considers convolution with the prompt.

2.5.3 Implementation of the intuitive algorithm

To account for the convolution with the prompt, equation 2.3 is used to evaluate the probability that a fluorescence photon will be detected within channel i from a hypothetical molecule with fluorescence lifetime $\overline{\tau}$ as given by equation 2.19. Figure 2.5 illustrates the probability densities if molecules A and B have fluorescence lifetimes of

4.2 ns and 2.5 ns, and the hypothetical molecule has a fluorescence lifetime of 3.35 ns. The figure shows that the probability density of the fluorescence is no longer a pure exponential decay, and hence the bin widths given in equation 2.24 are not applicable. In order to extend the intuitive algorithm to the case that considers the convolution with the prompt, the probabilities for each bin for the hypothetical molecule with mean fluorescence lifetime must be set to be equal, that is,

$$P_j^{\overline{\gamma}} = \int_{t_{j-1}}^{t_j} p^{\overline{\gamma}}(t)dt = 1/N, \quad j = 1, 2, ..., N, \tag{2.26}$$

where $p^{\overline{\gamma}}$ in equation 2.26 may be determined from equation 2.3, and t_{j-1} and t_j denote the start and end time of the j-th bin. These locations are to be adjusted for each bin in order to satisfy equation 2.26.

As indicated by equation 2.3, the probability densities of the prompt and molecules are known at M discrete channels. Therefore, the discrete form of equation 2.26 becomes

$$P_j^{\overline{\gamma}} = \sum_{i=i_{j-1}}^{i_j} p^{\overline{\gamma}}[i] = 1/N, \quad j = 1, 2, ..., N, \tag{2.27}$$

where i_{j-1} and i_j correspond to the start and end channels of the j-th bin.

If the total number of channels is much larger than the number of bins, then the discrete nature of the start and end channels will not pose a problem. However, when only 1600 channels are to be grouped to 64 bins, the resultant probabilities per bin for molecules of type A or B are bumpy, as shown in figure 2.6 (a). In order to solve this

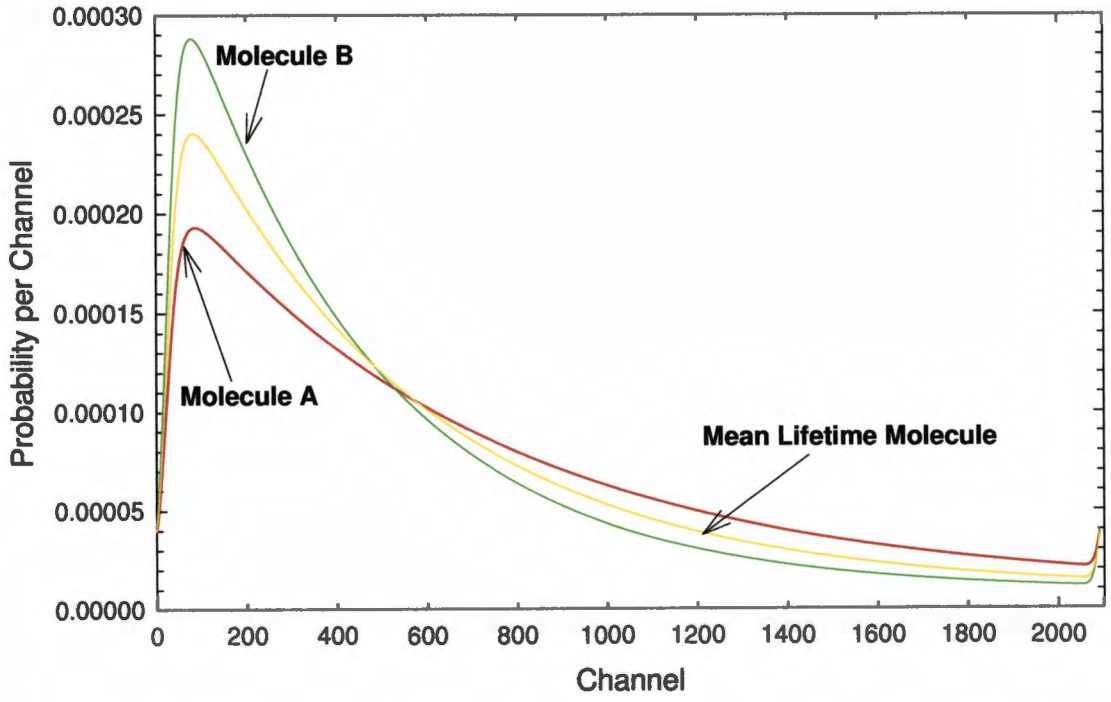


Figure 2.5: Probability densities for molecules A (4.2 ns), B (2.5 ns), and a hypothetical molecule $\overline{\gamma}$ (3.35 ns)

problem, linear interpolatation (smoothing) is used, and the start and end of the j-th bin are defined as

$$t_{j-1} = i_{j-1} + \delta t_j^s, \tag{2.28}$$

$$t_j = i_j + \delta t_j^e, \tag{2.29}$$

where, for $j = 1, 2, ..., N$,

$$0 < \delta t_j^s < 1 \quad , \quad 0 < \delta t_j^e < 1 \quad \text{and} \quad t_j^s = t_{j-1}^e. \tag{2.30}$$

Equation 2.27 then becomes

$$P_j^{\overline{\gamma}} = \sum_{i=i_{j-1}+1}^{i_j-1} p^{\overline{\gamma}}[i] + (1 - \delta t_j^s) p^{\overline{\gamma}}[i_{j-1}] + \delta_j^e \, p^{\overline{\gamma}}[i_j] = 1/N. \tag{2.31}$$

By applying the linear interpolation algorithm, the bumpy probability densities in figure 2.6 (a) are smoothed, and results are shown in figure 2.6 (b).

An important point that needs to be made here is that the time window is now shorter than the interval between two consecutive laser pulses, and hence photons may fall outside of the time window. However, the present MLE method only uses the photons inside the time window, and the probabilities, such as $P_j^{\overline{\gamma}}$ in equation 2.31, P_j^{Prompt}, P_j^A, and P_j^B, need to be re-normalized among all the bins that belong to the time window. All the probabilities are re-normalized inside the time window by

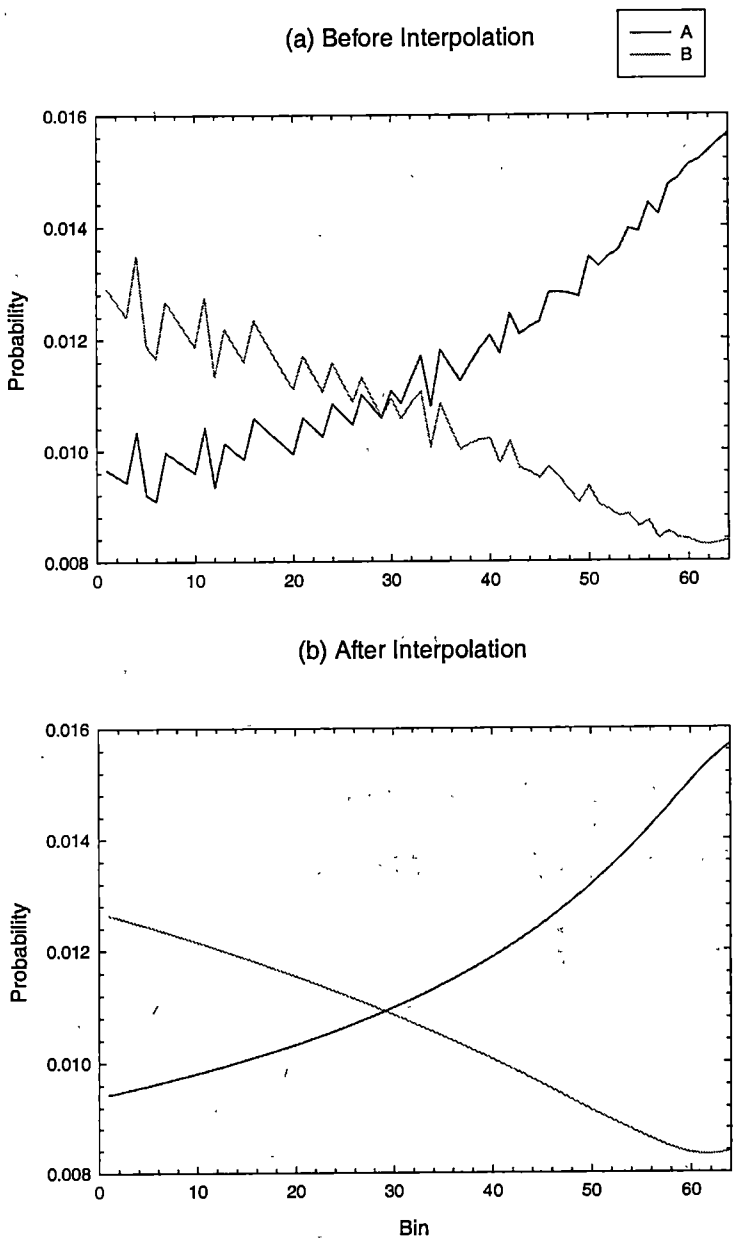


Figure 2.6: Accumulation of probabilities of 64 bins among the 1600 MCS channels

multiplying a normalization factor, defined as:

$$L_\mu = 1/[\sum_{i=i_0}^{i_N} p^\mu[i]], \quad \mu = \overline{\gamma}, Prompt, A, B. \tag{2.32}$$

Thus, equation 2.31 becomes

$$\begin{aligned} \overline{P_j^{\overline{\gamma}}} &= P_j^{\overline{\gamma}} L_{\overline{\gamma}} = L_{\overline{\gamma}} \sum_{i=i_{j-1}+1}^{i_j-1} p^{\overline{\gamma}}[i] + (1-\delta t_j^s) p^{\overline{\gamma}}[i_{j-1}] + \delta_j^e \, p^{\overline{\gamma}}[i_j] & (2.33) \\ &= 1/N. & (2.34) \end{aligned}$$

2.5.4 Bin widths changing by a constant factor

Another bin-width selection procedure is also examined in the coming chapter, i.e., the factorized bin-width selection, which uses a constant factor C_0 to obtain the consecutive bin widths recursively,

$$\omega_{j+1} = C_0 \omega_j, \tag{2.35}$$

where the first bin width is given by

$$\omega_1 = (C_0 - 1)T/(C_0^N - 1), \tag{2.36}$$

and the widths are defined in equation 2.21. The selection gives a set of decreasing bin widths when the constant, C_0, is less than one, or a set of increasing bin widths when it is larger than one.

As introduced in the above sections, several bin width selections are proposed. In the

simulation, the initialization stage gives the option to choose one of three methods for bin width selection, viz., bin widths generated by the intuitive algorithm given by equation 2.31, equal bin widths, and arbitrarily defined bin widths (without interpolation). The number of bins can be also specified at this stage. Typically, the intuitive algorithm is used and the number of bins is chosen as 64. Then, in order to save time, data for the case of 32 bin s are obtained by adding the photons in each of the consecutive pairs of bins. The process is repeated to obtain data for the case of 16, 8, 4, 2, and 1 bins. In this way, the probability density of the hypothetical mean lifetime molecule is still flat among all newly combined bins, and the approach of the intuitive algorithm is still valid.

2.6 Peak analysis

2.6.1 Determination of the start and end points of a burst

As introduced in section 2.3, the passage of molecules through the probe region produces bursts of photons, and the photons collected from each burst above a pre-set threshold are used to distinguish the identity of the molecule. The simulation must determine the start and end of each burst and hence which photons are to be included in the analysis of each burst. In the previous version of the simulation, the start and end of each burst is determined by locating the turning points of the weighted sliding sum. That is, the simulation records the times at which the weighted sliding sum varies from going down to up. The burst is discarded if the intervening maxima is below the pre-set threshold,

nut otherwise the local minima in the weighted sliding sum are used to define the start and end of each burst.

However, this method of determining the extent of a burst brings some problems when background noise is present. As illustrated in figure 2.7, when a burst is superposed on a broad pedistal of noise photons, the local minima are widely separated. Consequently, there are many background photons contained between these two points. The background photons degrade the performance of the data analysis, and result in incorrect predictions of the identity of the bursts. To circumvent this problem, the burst finding algorithm was modified to consider only those photons contained within a constant time interval, centered about the maximum of the burst. A time interval equal to the mean transit time of a molecule across the laser beam is used. This results in the new choice of the start and end of the burst, as shown in figure 2.7. Only those photons contained between the start and end of the burst are used in the subsequent analysis. Furthermore, the duration of the burst is taken to be the time interval between the first and last of these photons. If a molecule photodegrades during its transit through the laser beam, the duration of the burst will be shorter than the mean transit time, unless a background photon happens to fall just before the end point.

A problem with the new algorithm can arise if two photon bursts are closer together than the mean transit time. Figure 2.8 illustrates two bursts that are close together but not overlapping, together with the starts and ends of each of the bursts, as determined by the new algorithm. If the start point of the second burst $s2$ were to occur earlier than

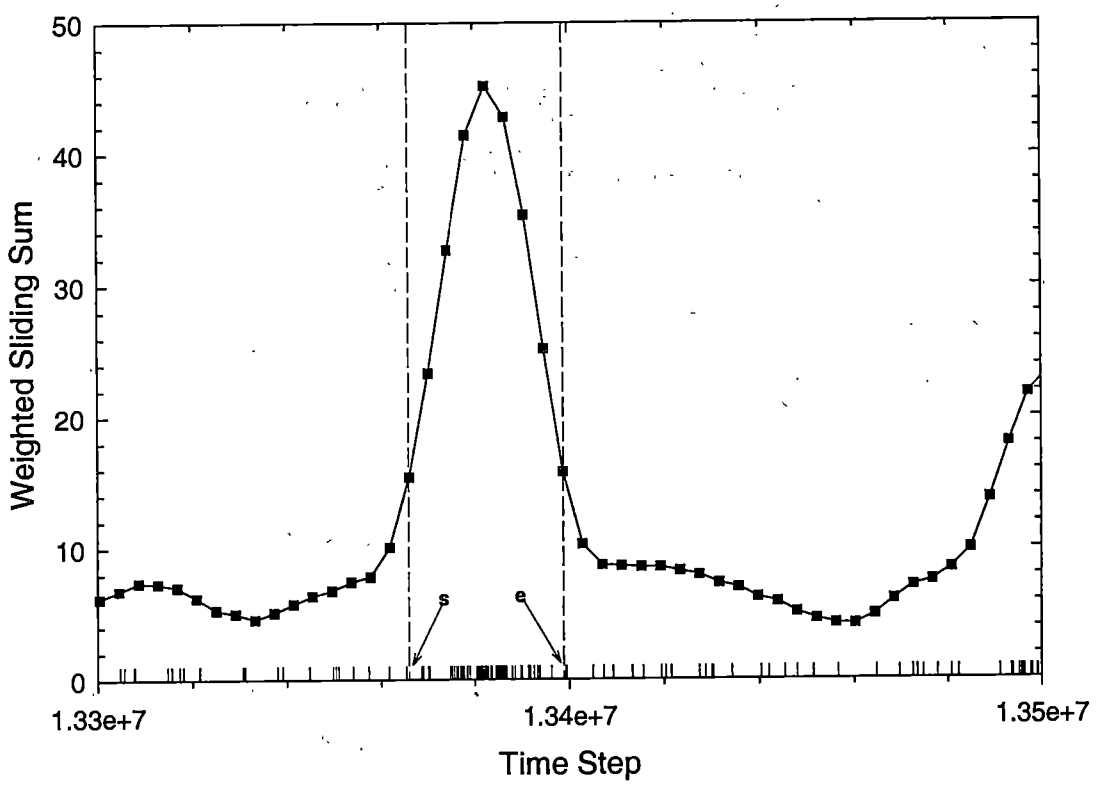


Figure 2.7: Weighted sliding sum of a photon burst superimposed on a pedistal of background photons

the end point of the first burst $e1$, then some of the photons belonging to the second burst would be counted into the first burst, and vice versa. To minimize this effect, the new algorithm is further modified to set the locations of $e1$ and $s2$ as the intervening minimum point in this event.

2.6.2 Burst type determination

The simulation must determine the actual cause of each burst, i.e., whether the burst originated from the transit of one or more molecules of type A, or of type B, or the transit of both types of molecules simultaneously, or from background noise alone. The type of each burst is output into a file for training the neural network analysis. It is also used to evaluate the accuracies of the MLE and the NN predictions.

The process of determining burst type is as follows: As discussed in section 2.3, whenever a photon is detected, the simulation records its origin, i.e., whether a photon originated from fluorescence from molecule type A, type B, from prompt, or from background. This information is used to determine the cause of each of the bursts. Having determined the start and end photons of each burst above the pre-set threshold, the cause of the burst is then initialized as being produced by only background. The simulation traverses all the photons in the burst, and checks if there are fluorescence photons other than background ones. If it finds one, it continues to look for whether, among the rest of the photons, there is a photon produced by the other type of molecule. If so, the burst will belong to the simultaneous passages of both types of molecules. Otherwise, the burst is due to the passage of only one type of molecule. If no fluorescence photon

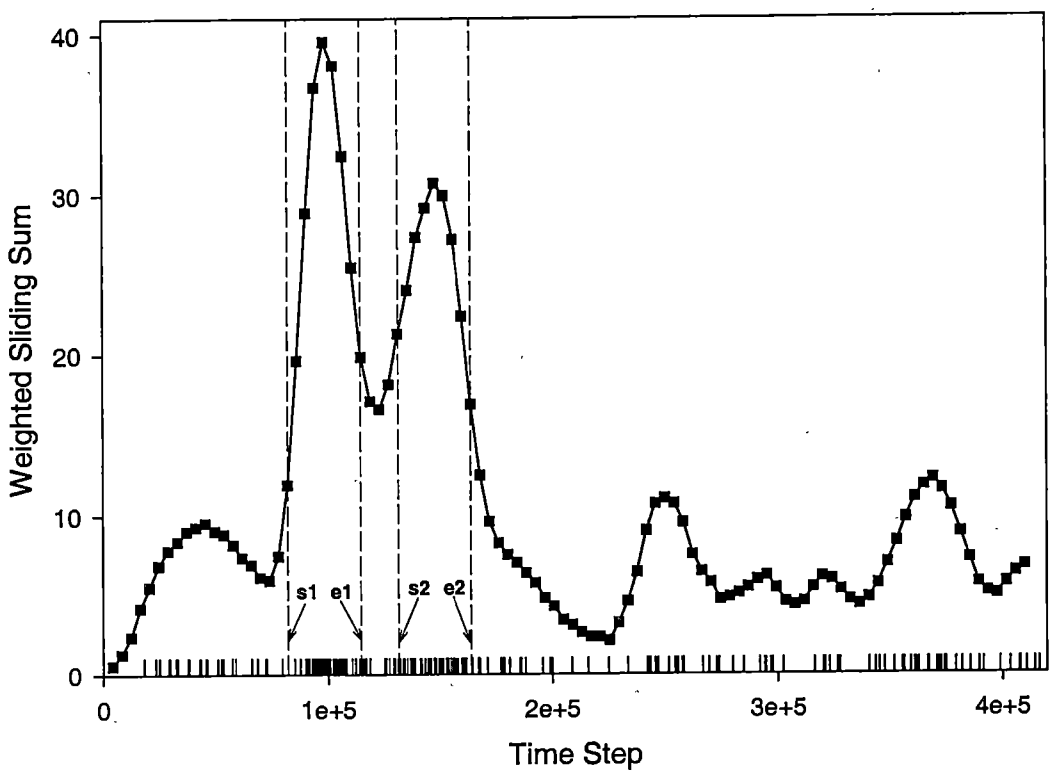


Figure 2.8: Consecutive photon bursts close to each other

is found in the burst, the burst will keep its initial type, i.e., from background.

Chapter 3

Maximum Likelihood Estimation

3.1 Overview

In the Monte Carlo simulation of single molecule detection, a maximum likelihood estimation (MLE) method is used to categorize molecules according to their fluorescence lifetimes, which differ for the distinct types of dyes. Other photophysical properties of such dyes may also be different, although these are not used when molecules are distinguished by the MLE method. In this work, only two types of dye molecules are to be distinguished as they are detected in the sheath flow system. The number of photon counts in each temporal bin are directly used in the categorization analysis.

In section 3.2, the principle of the MLE method for distinguishing two types of molecules is introduced. The implementation of the MLE method when background from promptly scattered photons is present is explained in section 3.3. A log-likelihood representation of the MLE analysis is introduced in section 3.4 for the case when counts

are Poisson distributed, and equations for determining the mean error in predicting molecules are given. In section However, there is no clear solution in the error analysis for considering a multinomial distribution and therefore, a direct approach to accumulate the mean error is proposed in section 3.6. The computational complexity of the direct approach is increased as the number of bins used becomes large or if background is included in the analysis. Some results are shown for the selection of 3 bins when background is omitted.

3.2 MLE method for distinguishing two types of molecules

The MLE method uses the distribution of photon counts $\{n_j\} = n_1, n_2, ..., n_N$ over each of the temporal bins for categorization of molecules with different fluorescence lifetimes. The conditional probabilities for obtaining the particular distribution of $\{n_j\}$, given that molecule A, or molecule B, is present in the probe region are evaluated and compared. According to equation 2.10, the multinomial probability for obtaining the distribution $\{n_j\}$ under the assumption that all the photons originate from one or more molecules of type A is

$$P(\{n_j\}|A) = K! \prod_{j=1}^{N} (\overline{P}_j^A)^{n_j}/n_j!, \tag{3.1}$$

and for molecule B, the probability is

$$P(\{n_j\}|B) = K! \prod_{j=1}^{N} (\overline{P}_j^B)^{n_j}/n_j!, \tag{3.2}$$

where $\overline{P}_j^A$ and $\overline{P}_j^B$ are given by equation 2.33, in which $p^{\overline{\gamma}}[i]$ is replaced by $p^A[i]$ and $p^B[i]$, as defined in equation 2.3. The ratio of the probabilities is

$$R = \frac{P(\{n_j\}|A)}{P(\{n_j\}|B)} = \prod_{j=1}^{N} (\frac{\overline{P}_j^A}{\overline{P}_j^B})^{n_j}. \tag{3.3}$$

The events are then categorized by

$$\begin{aligned} &\text{if } R > 1,\ P(\{n_j\}|A) > P(\{n_j\}|B), \quad \text{predict molecule } A, \\ &\text{if } R < 1,\ P(\{n_j\}|A) < P(\{n_j\}|B), \quad \text{predict molecule } B. \end{aligned} \tag{3.4}$$

3.3 Inclusion of background in the MLE analysis

The MLE method described in section 3.2 does not consider the case in which some of the photons in the burst originate from background. In order to account for background photons in general, all the possible distributions for all possible combinations of fluorescence and background photons that can give rise to the observed distribution need to be considered. For example, if photons could be generated by molecule A or background with equal likelihood, the probability for obtaining the observed distribution $\{n_j\}$ would

be given by

$$P(\{n_j\}|\text{molecule } A) = \sum_{b_1=0}^{n_1}\sum_{b_2=0}^{n_2}...\sum_{b_N=0}^{n_N} P(n_1-b_1, n_2-b_2,..., n_N-b_N|A)P(b_1,b_2,...,b_N|bg) \quad (3.5)$$

where $P(n_1-b_1, n_2-b_2,..., n_N-b_N|A)$ is determined by equation 3.1, and $P(\{b_j\}|bg)$ is the multinomial probability for obtaining the distribution $\{b_j\} = b_1, b_2, ..., b_N$ if all photons originate from background, i.e.,

$$P(\{b_j\}|bg) = (\sum_{j=1}^{N} b_j)! \prod_{j=1}^{N} (\overline{P}_j^{bg})^{b_j}/b_j!. \quad (3.6)$$

However, in most cases the majority of photons are from fluorescence and hence background and fluorescence photons are not equally likely. Equation 3.5 would need to be modified to account for the expected ratio of background and fluorescence photons. However, due to the computational complexity, it would be impractical to obtain results from such equations in a reasonable amount of time. Therefore, another approach is taken to account for the presence of background. If K photons are detected and if a number of these are expected to be due to background, then the probability that any given photon will fall within bin j is assumed to be the weighted average of $\overline{P}_j^{bg}$ and either $\overline{P}_j^A$ or $\overline{P}_j^B$. The expected number of photons within a burst of duration of Δt is $b\Delta t$, where b is the mean rate of background photons. Hence, if a burst contains a total

of K photons, the probabilities for each bin are

$$\widetilde{P}_j^A = [(K - b\Delta t)\overline{P}_j^A + b\Delta t\overline{P}_j^{bg}]/K,$$
$$\widetilde{P}_j^B = [(K - b\Delta t)\overline{P}_j^B + b\Delta t\overline{P}_j^{bg}]/K. \tag{3.7}$$

The background is composed of dark noise and promptly scattered light, but the probability of promptly scattered light is much larger than that of dark noise, and hence $\overline{P}_j^{bg}$ in equation 3.7 is replaced by $\overline{P}_j^{Prompt}$, which is obtained by replacing $\overline{\gamma}$ in equation 2.33 by *Prompt*.

Equation 3.3 then becomes

$$R = \frac{P(\{n_j\}|\text{molecule } A)}{P(\{n_j\}|\text{molecule } B)} = \prod_j \left[\frac{(K - b\Delta t)\overline{P}_j^A + b\Delta t\overline{P}_j^{Prompt}}{(K - b\Delta t)\overline{P}_j^B + b\Delta t\overline{P}_j^{Prompt}}\right]^{n_j} \tag{3.8}$$

Furthermore, if the photon burst is assumed to originate from either molecule A or molecule B (i.e., there are no other possibilities), then the confidence that a burst is due to A, or B, is given by

$$c_A = \frac{P(\{n_j\}|\text{molecule } A)}{P(\{n_j\}|\text{molecule } A) + P(\{n_j\}|\text{molecule } B)} = \frac{R}{1+R}, \tag{3.9}$$

$$c_B = \frac{P(\{n_j\}|\text{molecule } B)}{P(\{n_j\}|\text{molecule } A) + P(\{n_j\}|\text{molecule } B)} = \frac{1}{1+R}. \tag{3.10}$$

Note that $c_A + c_B = 1$, and thus only c_B is calculated in the simulation. If c_B is greater than 0.5, the passing molecule is predicted as B, otherwise, the molecule is predicted

as A.

As introduced in section 2.5, various numbers of bins are chosen in the simulation to study the influence of changing the number of temporal bins on the precision of the identification of the molecules. In the simulation, the data from the maximum number of bins (64) generated in a single run are used in subsequent analyses for the cases of 32, 16, 8, 4, 2 and 1 bins. In order to do so, some of the intermediate variables of the simulation code must be recorded for each burst, in particular, the quantities $b\Delta t\overline{P}_j^{Prompt}$, which are used in equation 3.8 must be saved and regrouped accordingly.

3.4 Log-Likelihood function and error calculation

Another approach of the MLE method, which is discussed for the case of negligible background, is to evaluate the difference of the logarithms of $P(\{n_j\}|A)$ and $P(\{n_j\}|B)$, instead of the ratio of the two probabilities, as given in equation 3.3. The advantage of this approach is that the computations are reduced from powers to multiplications, and the execution time of the analysis is shortened. An overview of the Log-Likelihood approach [9] will be briefly introduced in this section as some of the approaches to evaluate the bin width selections are based on this information.

Recall that the $\overline{P}_j^A$ and $\overline{P}_j^B$ in section 3.2 satisfy the normalization condition, i.e.,

$$\sum_{j=1}^{N}\overline{P}_j^A = 1, \quad \sum_{j=1}^{N}\overline{P}_j^B = 1. \tag{3.11}$$

Reference [9] considers the case in which the probability of obtaining a certain photon distribution in the time window, e.g., n_j photons in the j-th bin, is the product of Poissonian probabilities, instead of multinomial distribution as formulated in equation 3.1. A clarification of the difference between the multinomial and Poisson description of the counts n_j is given in reference [23]. Each n_j is Poisson distributed only if the total number of photons, $K = \sum_j n_j$, is Poisson distributed. In the case of SMD, photodegradation, triplet crossing, and other processes result in K actually not being Poisson distributed. On the other hand, once K is measured, n_j rigorously follows a multinomial distribution. Thus the Poisson approach does not make use of all the information, it introduces an insecurity in K that is actually not present.

For Poisson distributions within each bin, given that molecule A or B is present, the probabilities are defined as

$$P(\{n_j\}|A) = e^{-K} \prod_{j=1}^{N} \frac{(K\overline{P}_j^A)^{n_j}}{n_j!},$$
$$P(\{n_j\}|B) = e^{-K} \prod_{j=1}^{N} \frac{(K\overline{P}_j^B)^{n_j}}{n_j!}. \tag{3.12}$$

The log-likelihood function (or M function) is defined as

$$M = \sum_{j=1}^{N} n_j m_j, \tag{3.13}$$

where

$$m_j = \ln(\frac{\overline{P}_j^A}{\overline{P}_j^B}). \tag{3.14}$$

To predict the event, a similar approach as in equation 3.4 is given by

$$\begin{aligned} &\text{if} \quad M > 0, \quad \text{predict molecule } A, \\ &\text{if} \quad M < 0, \quad \text{predict molecule } B. \end{aligned} \tag{3.15}$$

Further, the mean error for predicting molecule B when A is actually present is defined as

$$err_K(A) = \int_{-\infty}^{0} dM P_K^{(A)}(M), \tag{3.16}$$

and the mean error of predicting A while B is present is defined as

$$err_K(B) = \int_{0}^{+\infty} dM P_K^{(B)}(M), \tag{3.17}$$

where $P_K^{(A)}(M)$ and $P_K^{(B)}(M)$ are given by summing equation 3.12 over all possible sets of $\{n_j\}$ that give a particular value of M. Reference [19] gives the form of $P_K^{(\gamma)}(M)$ as

$$P_K^{(\gamma)}(M) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \{\prod_{j=1}^{N} \frac{(K\overline{P}_j^{\gamma})^{n_j}}{n_j!} \exp[-K\overline{P}_j^{\gamma}]\} \delta[M - \sum_{j=1}^{N} n_j m_j], \tag{3.18}$$

where $\gamma = A$, or B. In order to evaluate the mean errors in equations 3.16 and 3.17, the

Dirac δ function in equation 3.18 is replaced by

$$\delta(z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikz}. \tag{3.19}$$

Substituting this into equation 3.18 yields

$$P_K^{(\gamma)}(M) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} Q_K^{(\gamma)}(k) e^{ikM}, \tag{3.20}$$

where

$$\begin{aligned} Q_K^{(\gamma)}(k) &= \prod_{j=1}^{N} \sum_{n_j=0}^{\infty} \frac{(K\overline{P_j}^{\gamma})^{n_j}}{n_j!} \exp(-K\overline{P_j}^{\gamma} - ikn_j m_j) && (3.21) \\ &= \exp[K < \exp(-ikm) - 1 >_{\gamma}]. && (3.22) \end{aligned}$$

The abbreviation in equation 3.22 denotes

$$< f >_{\gamma} = \sum_{j=1}^{N} f_j \overline{P_j}^{\gamma}. \tag{3.23}$$

Hence, the mean error of identification of molecule A in equation 3.16 or that of B in equation 3.17 can be expressed as

$$err_K(A) = \int_{-\infty}^{0} dM \int_{-\infty}^{\infty} \frac{dk}{2\pi} Q_K^{(A)}(k) e^{ikM}, \tag{3.24}$$

or

$$err_K(B) = \int_0^{+\infty} dM \int_{-\infty}^{\infty} \frac{dk}{2\pi} Q_K^{(B)}(k) e^{ikM}. \tag{3.25}$$

The mean error for a mixture of molecules A and B is defined by

$$err = \frac{\rho_A err(A) + \rho_B err(B)}{\rho_A + \rho_B}, \tag{3.26}$$

where ρ_A and ρ_B are the actual concentrations for molecule A and B.

3.5 Errors for bin selection with widths changing by a constant factor

The bin selection algorithms introduced in section 2.5 can be evaluated by numerically solving equations 3.24, and 3.25 by calling a standard math library for performing Gauss-Kronrod integration, and comparing the predicted errors for different bin selections. This is done here for the case of bin widths changing by a constant factor, and the result is shown in figure 3.1.

In the calculation of figure 3.1, the total number of bins, N, is 4, the fluorescence lifetimes of molecule A (τ_A) and B (τ_B) are 2.0 ns and 4.0 ns, the total number of photons K is 100, the width of the time window T is set to be 8.0 ns, and the concentrations of molecule A and B are equal. Figure 3.1 gives the results corresponding to the bin widths changing by a series of constant factors, C_0. For the purpose of comparison, a straight line with mean error of 1.3%, which is obtained from the intuitive bin width

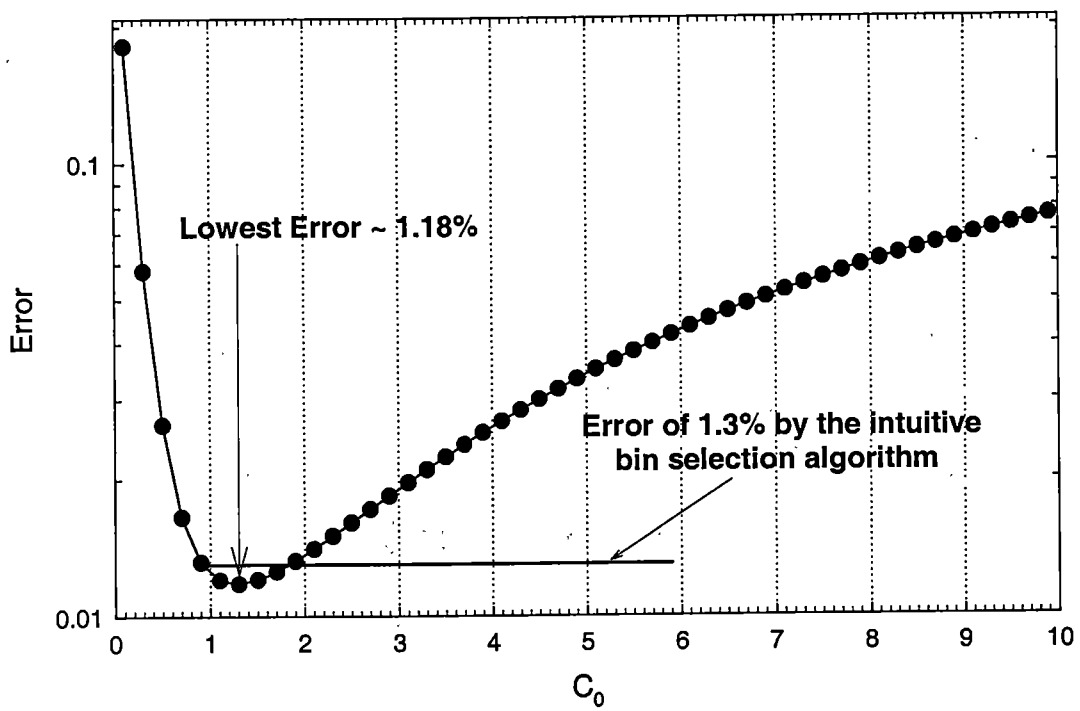


Figure 3.1: Mean identification error of a mixture of molecule A and B using a set of bin widths changing by a constant factor C_0.

algorithm introduced in section 2.5.2, is also illustrated.

In figure 3.1, the part of the curve with C_0 <1 represents the selected bins with a set of decreasing widths, and the corresponding error of identification is expected to be large as discussed in section 2.5.2. Similarly, the identification error would also be significantly increased if C_0 is large, say C_0 >2. Hence, the part of curve within $1 < C_0 < 2$ is expected to give least error. For the assumed parameters, the lowest error of 1.18% is achieved when C_0 is 1.3. Under the same circumstances, the intuitive algorithm gives 1.3% of identification error, which is higher than that of the bin widths selection changing by a constant factor of C_0=1.3.

In summary, using the intuitive algorithm introduced in section 2.5.2 to select bins may not be the best way to minimize the error. However, the bin width selection with widths changing by a constant factor may not be the best solution either. There may be another algorithm that yields an even lower error. Also, with another set of parameters, bin widths determined by the intuitive algorithm may give a lower error than any of the bin width selections changing by a constant factor. Note that a Poissonian rather than multinomial distribution has been assumed in the analysis in section 3.4. An extension of the above analysis to the multinomial case would not be straight forward because it would no longer be possible to obtain the simplified form in equation 3.22, which is used in the numerical evaluation of equations 3.24 and 3.25. In addition, the contribution from background photons is not considered, and the addition of background into the analysis would further complicate the problem. Therefore, a direct approach to calculate

the error is pursued, and details are given in the next section.

3.6 Direct method for error calculation

Recall that, in equation 3.16 and 3.17, the error of predicting molecule A is accumulated for those cases of $\{n_j\}$ that give rise to a negative M value, and the error of predicting molecule B is accumulated for those cases of $\{n_j\}$ that give rise to a positive M value. Instead of evaluating the distribution of the Log-likelihood function, as well as the following equations from 3.22 to 3.25, the error for predicting either type of molecule can be obtained by systematically considering all possible distributions $\{n_j\}$ and accumulating the probabilities for obtaining each such distribution for which M is negative when A is present or for which M is positive when B is present. For simplicity, background photons are not considered in the following discussions.

Consider the case in which there are totally K photons to be sorted into N bins, and hence there is a total of (N+K)!/(N!K!) ways to sort them into the bins. Each way of allocation of the photons into the bins has a certain probability, i.e., the multinomial probability, given by

$$P(\{n_j\}|\gamma) = K! \prod_{j=1}^{N} (\overline{P}_j^{\gamma})^{n_j}/n_j!, \tag{3.27}$$

where $\gamma = A, B$. Therefore, the error of predicting molecule A (i.e., predicting B when

A is actually present) is

$$err_K(A) = \sum_{\forall\{n_j\}\ni M<0} P(\{n_j\}|A), \tag{3.28}$$

and error of predicting molecule B (i.e., predicting A when B is actually present) is

$$err_K(B) = \sum_{\forall\{n_j\}\ni M>0} P(\{n_j\}|B). \tag{3.29}$$

Once the time window is selected, detected photons will always fall within the time window and the probabilities $\overline{P}_j^{\gamma}$ have accordingly been normalized in equation 3.27. However, the problem of interest is how to simultaneously choose both the time window and the bins within the time window to minimize the error. In order to address the problem, consider the case in which the entire temporal decay profile is divided into N bins, with probabilities P_j^{γ}, j=1, 2, ..., N. Only the distribution of photons within the bins j=2, ..., $N-1$ is used for prediction, and hence M is evaluated using

$$M = \sum_{j=2}^{N-1} n_j \ln(\frac{\overline{P}_j^A}{\overline{P}_j^B}), \tag{3.30}$$

rather than equations 3.13 and 3.14, but the multinomial probabilities in equations 3.28 and 3.29 are given by

$$P(\{n_j\}_{j=1}^N|\gamma) = (\sum_{j=1}^N n_j)! \prod_{j=1}^N (P_j^{\gamma})^{n_j}/n_j!. \tag{3.31}$$

Selection of the time window among 5 bins is illustrated in figure 3.2. All the bins are numbered from 1 to 5, and the time window occupies bins 2, 3, and 4. Therefore, the photons falling into bins 1 and 5 will not be considered when evaluating the Log-likelihood function M.

A number of bin width selections are evaluated using the 3 bins case, illustrated in figure 3.2. The widths of all the 5 bins are adjusted according to the selecting methods below. The M function is evaluated by using equations 3.13 and 3.14 while the error is accumulated by using equation 3.27, 3.28, and 3.29. The following bin width selections are considered

- Case 1: Increasing bin widths by a factor of $C_0 = 10$,
- Case 2: Decreasing bin widths by a factor of $C_0 = 0.5$,
- Case 3: Intuitive bin width selection introduced in section 2.5.2,
- Case 4: Equal bin width, C_0=1.0,
- Case 5: Increasing bin widths by a factor of $C_0 = 1.3$, with which the lowest error is achieved in section 3.5.

The comparison among the above bin width selections is shown in figure 3.3. Here, the mean errors are evaluated for different total numbers of photons, K. Other parameters used in the calculation are kept the same as in section 3.5.

The results show that the mean identification errors are large for all 5 cases when K is small. This is consistent with the discussion in section 2.5.1, where the total

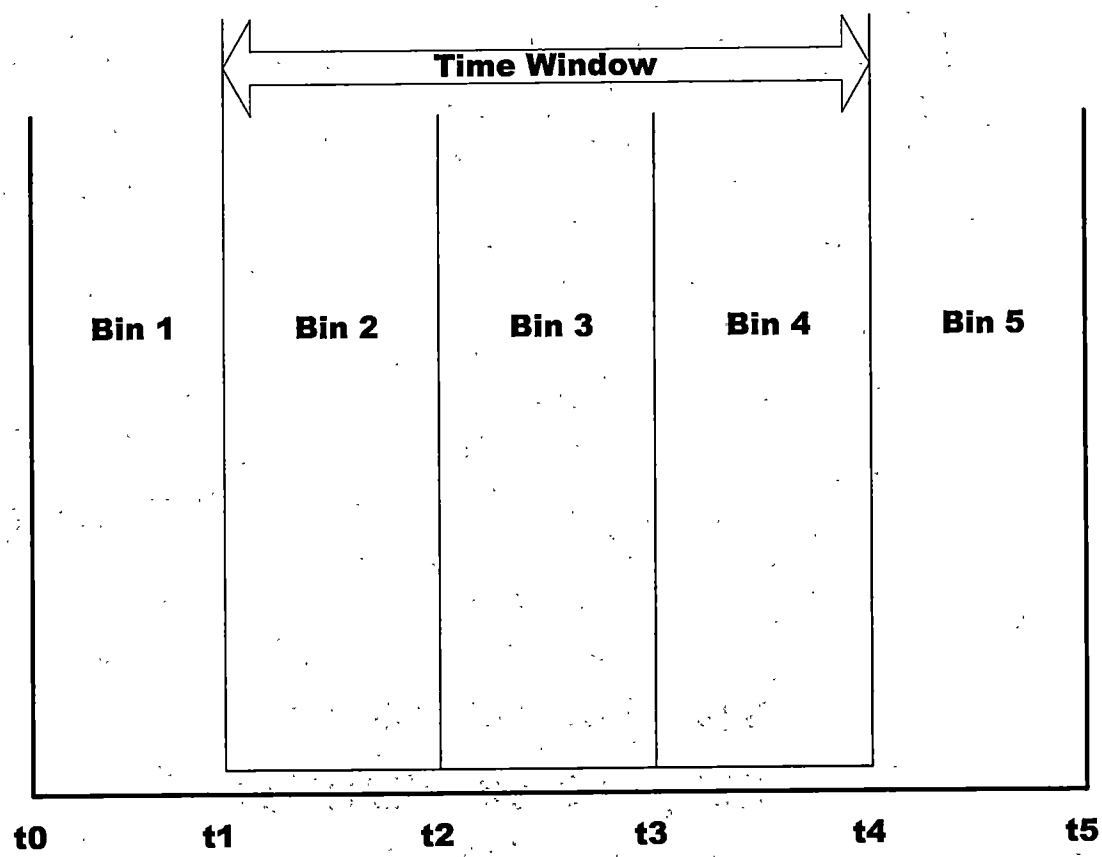


Figure 3.2: 3 bins out of 5 form the time window

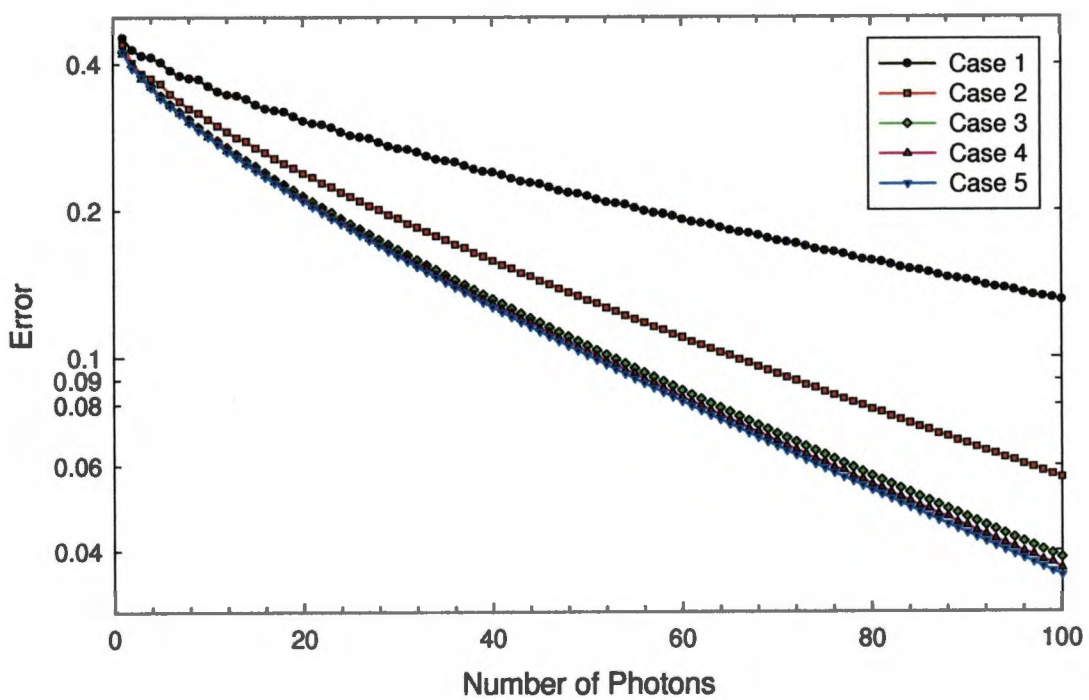


Figure 3.3: Comparison of identification errors among several bin selection methods

number of collected photons was shown to influence the precision of the estimation of the lifetime of a molecule. An important point to note here is that the total number of photon counts K within the burst does not directly influence the decision, but it does impact the reliability of the categorization. Furthermore, the cases for a selection of decreasing bin widths and of increasing bin widths with C_0=10 exhibit larger errors than other cases. This is consistent with the results shown in figure 3.1.

The bin selections with equal widths, with a set of increasing widths by a factor, C_0, of 1.3, and with the widths chosen by using the intuitive algorithm have nearly the same low error. There is no clear evidence showing which of these selections is the best. Minor differences in the errors obtained by each method are attributable to differences in the widths of the time window, which causes the mean number of photons within the time window to differ. The width of the time window formed by bins 2, 3, and 4 is 1.0, 4.5, 3.8, 6.0, and 5.7 ns for cases 1–5 respectively. The result from the set of bin widths determined by the intuitive algorithm is expected to be poorer than that from the equal bin widths because the time window is narrower. In practice, the time window selection will be determined largely by the temporal profile of the background.

3.7 Possibility of systematically searching for the optimal bin width selection

The motivation of the discussions in this section is to look for an approach for finding an optimal selection of time window and bin widths within the time window, so as to

achieve the minimum identification error of molecules. One possible way to accomplish this would be to perform a systematic search over all possible bin width selections. In specific, all the start and end points of the time window and of the bins within the time window can be varied within the allowed range such that

$$t_0 < t_1 < t_2 < ... < t_{N-1} < t_N. \tag{3.32}$$

The time window starts at t_1 and ends at t_{N-1}. Thus, all possible bin widths could be obtained, and, by recording the mean error for each set of widths, the minimum error could be found. The identification error for a selection of bin widths is evaluated by the direct method, introduced in section 3.6.

A simple illustration of the systematic searching approach uses the 5-bin case in figure 3.2, and the points t_1 and t_4 are fixed at 2.0 ns and 8.0 ns, i.e., the start and end points of the time window are fixed. The inclusion of the variable start and end points of the time window would require more computational memory and time. In addition, the background is ignored to simplify the implementation of the analysis. The points t_2 and t_3 in figure 3.2 are varied to give all possible widths for bins 2 , 3, and 4. The total number of photons is set to be 100, and the resulting 2 dimensional surface is drawn in figure 3.4. The data in figure 3.4 was obtained with approximately 90 minutes execution time on a Pentium 200 MHz PC (64 MB memory).

The result shows that the mean error fluctuates as the bin widths within the time window are varied, resulting in a bumpy two dimensional surface with many local min-

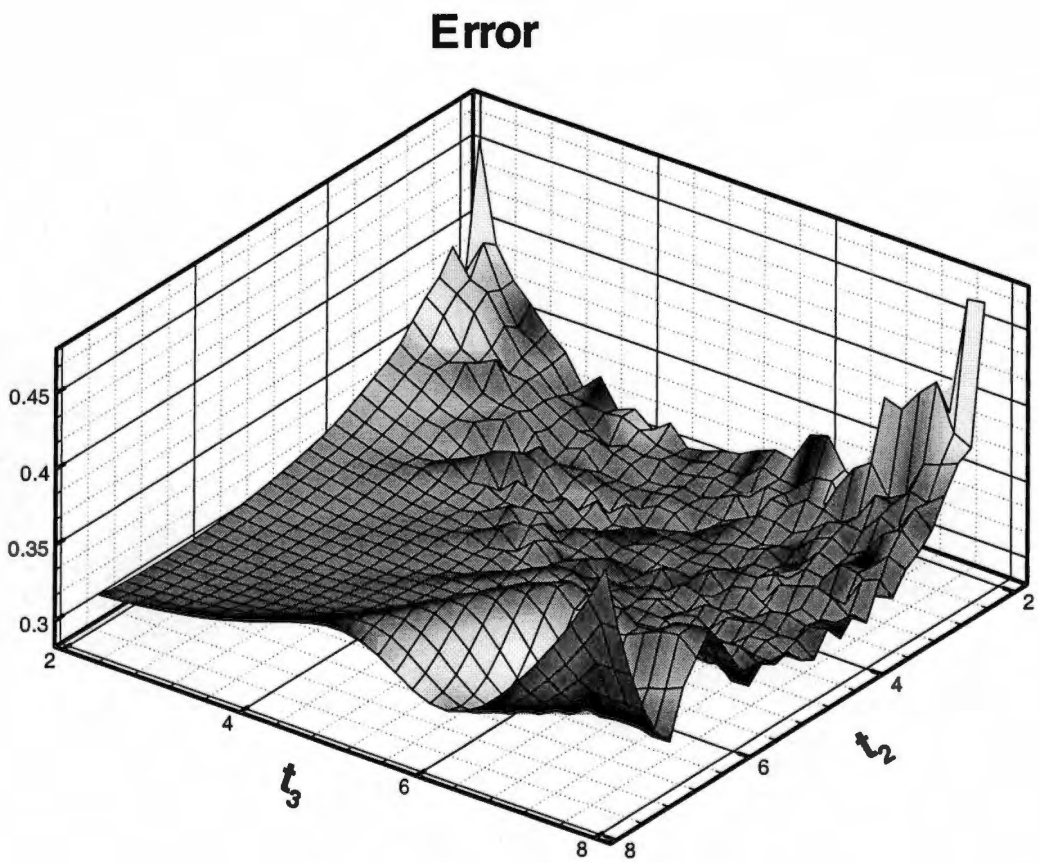


Figure 3.4: Error rate vs. two dimensional searching surface

ima. This would make it difficult for a searching algorithm to locate the global minimum without evaluation of the error at every point of the domain. In the calculation presented in figure 3.4, the searching step is around 0.1 ns, and the lowest error is achieved when t_2 is 4.0 ns, and t_3 is 6.1 ns. This is approximately equivalent to selecting equal widths for bins 2, 3, and 4.

Note that the total number of photons may differ over a wide range from one burst to the next, whereas the calculation of figure 3.4 assumes a constant burst size of K=100 photons. A rigorous approach to finding the bin selection would require calculation of the weighted average of the identification error $err_K(\gamma)$ for all possible burst sizes K, i.e.,

$$err(\gamma) = \sum_{K=1}^{\infty} \Omega_K err_K(\gamma), \tag{3.33}$$

where Ω_K is the probability that a detected burst contains K photons. However, such an approach would be computationally intensive, and was not attempted here. Further, it would be considerably more computationally intensive to implement such a searching algorithm to obtain a set of optimal bin widths when there is a large number of bins within the time window. Even in the case of 3 bins within the time window, shown in figure 3.2, the computational complexity would be substantially increased if one were to allow to consider the selection of the start and end points of the time window, i.e., t_1 and t_4 to also be varied.

The above illustration is simplified in that, from the point of view of the SMD experiment, one would need to consider the situation described in section 2.4, where

photons may originate from either background or fluorescence of molecules. The direct method could be extended to consider the contribution from the background given that all the collected bursts are generated by the fluorescence of molecules. This could be achieved by setting a higher threshold to block background bursts, which normally have smaller peak amplitudes than those from molecules. To account for the influence from the background, the $\overline{P}_j^{\gamma}$ in equation 3.27 would be replaced by $\widetilde{P}_j^{\gamma}$, which is expressed explicitly in equation 3.7 for $\gamma = A, B$. Thus, the multinomial probability of distribution $\{n_j\}$, considering the background, would be given by

$$P(\{n_j\}|\text{molecule } \gamma) = K! \prod_{j=1}^{N} [\frac{K - b\Delta t}{K} \overline{P}_j^{\gamma} + \frac{b\Delta t}{K} \overline{P}_j^{Prompt}]^{n_j} / n_j!, \tag{3.34}$$

and the error of predicting molecule A or molecule B would be defined as

$$err_K(A) = \sum_{\forall\{n_j\} \ni M<0} P(\{n_j\}|\text{molecule } A), \tag{3.35}$$

or

$$err_K(B) = \sum_{\forall\{n_j\} \ni M>0} P(\{n_j\}|\text{molecule } B). \tag{3.36}$$

The log-likelihood function would be evaluated by

$$\begin{aligned} M &= \sum_{j=2}^{N+1} n_j \ln(\frac{\widetilde{P}_j^A}{\widetilde{P}_j^B}) \\ &= \sum_{j=2}^{N+1} n_j \ln[\frac{(K - b\Delta t)\overline{P}_j^A + b\Delta t \overline{P}_j^{Prompt}}{(K - b\Delta t)\overline{P}_j^B + b\Delta t \overline{P}_j^{Prompt}}]. \end{aligned} \tag{3.37}$$

The above equations could be used to incorporate the presence of background in the calculation of the weighted errors for identification of bursts, given by equation 3.33. In principle, the multidimensional parameter space representing the boundaries of the time bins, t_1, t_2, ..., t_{N-1}, could be searched to find the minimum error. However, in practice, a very large amount of computational power would be required for the calculations and search. The selection of the time window by the method described in section 2.4.1, so that the expected signal in the time window exceeds the background, would reduce the dimensionality of the parameter space by 2, but the computations required would still be very large. The Monte Carlo simulations, as presented in Chapter 5, offer a viable alternative approach for comparing the identification errors obtained with different bin selections. For example, in section 5.5, a comparison of the errors in prediction are made between equal bin widths and bin widths selected by the intuitive algorithm.

Chapter 4

Neural Networks Analysis

4.1 Overview

Neural networks analysis provide a computational methodology different from that of traditional hard computing or artificial intelligence. Hard computing methods normally refer to scientific calculations, in which a complete set of inputs to a program is provided and the exact results are deduced by executing the program. On the other end, an artificial intelligence system, such as the weather forecast system, does not strictly require comprehensive inputs, and thus there may exist many possible outcomes for the provided incomplete information. Neural networks model processes information in a manner that imitates the functionality of human brains, where training is required prior to testing or use. In the training stage, the network is provided information(training data) that contains inputs to the networks and the corresponding desired outputs. In the testing stage, the trained network, with its current state of knowledge, provides the

answers to new problems(testing data), and such testing data contain only inputs to the network in this phase.

In section 4.2, the basic concepts of neural networks are introduced, and power of the neural networks analysis is discussed in section 4.2.6. Section 4.3 discusses how to implement the neural networks analysis for categorization of photon bursts from simulated SMD experiments. A procedure of how to incorporate neural networks analysis into the SMD experiment is proposed in section 5.8 of chapter 5 .

4.2 Basic concepts in NN analysis

The neural networks model was first introduced to imitate processes in human brain, where information is processed in the form of electrical pulses, transmitting from one neuron to others. Each neuron, as the basic processing unit in the brain, collects all the incoming signals from its neighboring neurons, from which it determines whether to activate its output to the next set of neurons. During training, the extent to which it needs to activate each of the next set of neurons is determined, and the connection between neurons is strengthened or weakened accordingly. In general, human brains contain up to 10^{10} neurons, each one of which has a response time of the order of a milli-second. However, due to the massive connections among the neurons($\sim 10^4$ per neuron), human brains are able to process information in a fast and accurate fashion, and are more powerful than current computer implementations.

In a similar way, neural networks are constructed to have many processing units that

connect to each other with certain weights. The weights denote the importance of each connection in the network. Each unit first evaluates all its inputs in a certain way, e.g., by summing up all the weighted incoming values. It then uses the resultant value to calculate an activation function and output from this is passed to the next neuron. In this way, input values to the network are propagated through a network of neurons. In the training stage, the output neurons compare the final results with the desired results provided by the training data. The calculated error from such a comparison will provide information for how much the weights of connections should be adjusted, and the goal is to minimize this error. Having obtained the lowest error, the trained networks would save the information of the training data in the form of its current state of the networks, i.e., the set of fixed weights for each connection.

4.2.1 Architecture of neural networks

All the neurons in the neural networks are grouped into layers. The first layer of the neural networks is the input layer, to which users provide the inputs of a particular task. The last layer, the output layer, is composed of neurons for outputting results. Users could choose to have some intermediate layers or hidden layers to increase the analyzing power as discussed in section 4.2.4. When counting the number of layers in a network, only the hidden layers and the output layer are counted, e.g., 2-layer networks have one input layer, one output layer, and one hidden layer. According to the location of layers, processing units can be divided into three categories, i.e., input units that receive data from outside of the network, output units which send data out of the network, and

hidden units, whose input and output signals remain within the neural networks.

There are two types of networks which differ from each other by the direction of their connections. In feed-forward networks, the signals propagate only in the forward direction, and all units receive values only from neurons in preceding layers and send values only to following layers. Units may not connect to neurons in the same or previous layers. In recurrent networks, feedback connections may exist. In contrast to feed forward networks, the dynamical properties of recurrent networks are important. In some cases, the activation values of the units undergo a relaxation process such that the network evolves to a stable state in which the value do not change any more. In this work, only the feed-forward type of network is considered.

4.2.2 Weighted summation and activation rules

In a feed-forward network, each processing unit receives inputs from preceding neurons or, in the case of input neurons, from external sources, and calculates the weighted summation of all inputs and corresponding activation function. The output value from the unit is propagated to the next set of neurons, and is re-calculated through each of the units along the path. In the stage of learning or training, the units in hidden layers and output layer adjust the weights of those input connections according to their individual importance leading to the desired output. The diagram of a typical processing unit is shown in figure 4.1.

In figure 4.1, the unit k has a number of preceding connections, among which the importance of the connection from unit j is denoted by a weight factor w_{jk}. In most

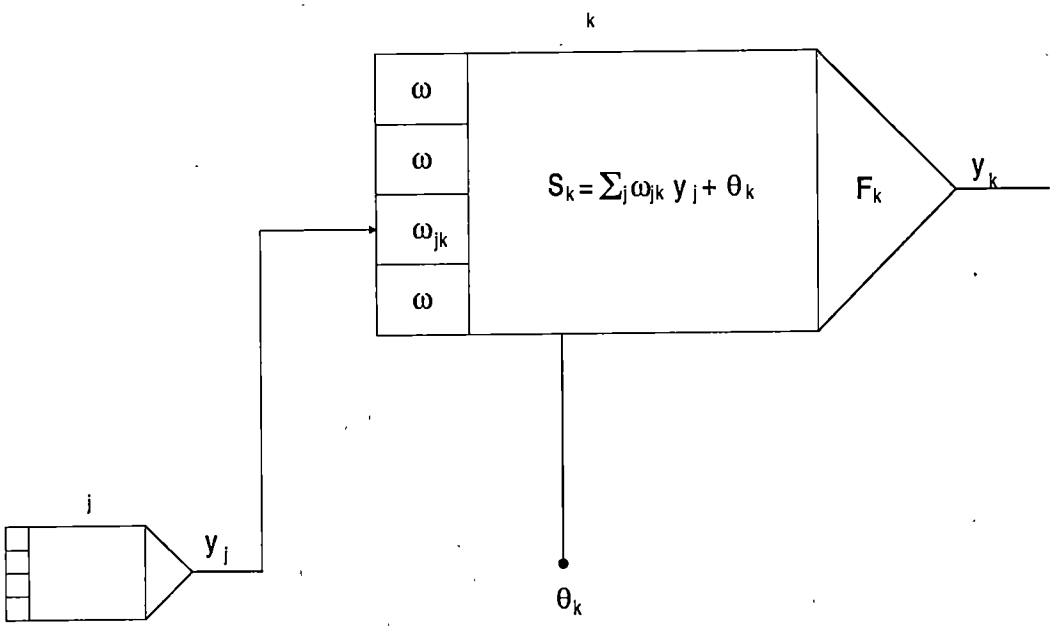


Figure 4.1: The basic components of an artificial neural networks

cases, each processing unit performs a summation over all its weighted inputs, and a bias or offset term θ_k for unit k. In figure 4.1, the summation is expressed as

$$s_k(t) = \sum_j w_{jk}(t) y_j(t) + \theta_k(t), \tag{4.1}$$

where t denotes the current execution time, and $y_j(t)$ is the input value from unit j.

As described in the beginning of this section, whether the current inputting connection has the significant contribution to its output is determined by a pre-defined threshold function (or activation function). Generally, the neural networks would provide a rule that gives the effect of the total inputs on the activation of the unit. The activation function, F_k, takes the total inputs $s_k(t)$, and produces a new value of the activation of the unit k at the next time step, i.e.,

$$y_k(t+1) = F_k[s_k(t)] = F_k[\sum_j \omega_{jk}(t) y_j(t) + \theta_k(t)]. \tag{4.2}$$

There are three types of threshold functions normally used. The sign function, as shown in figure 4.2 (a), is defined as

$$\begin{aligned} s_k < T_0,\ F_k(s_k) = 0, \\ s_k > T_0,\ F_k(s_k) = 1, \end{aligned} \tag{4.3}$$

where T_0 is a pre-set constant. The linear or semi-linear function in figure 4.2 (b), is

defined as

$$s_k < T_0,\ F_k(s_k) = 0,$$

$$s_k > T_1,\ F_k(s_k) = 1,$$

$$s_k \in [T_0, T_1],\ F_k(s_k) = \frac{s_k - T_0}{T_1 - T_0}, \tag{4.4}$$

where T_0 and T_1 are constants. The sigmoid function, shown in figure 4.2 (c), is defined as

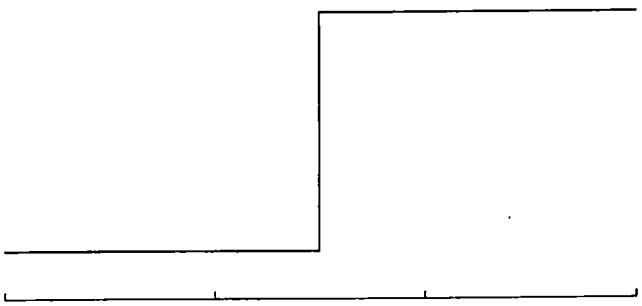
$$y_k = F_k(s_k) = \frac{1}{1 + e^{-s_k}}. \tag{4.5}$$

4.2.3 Learning of the neural networks

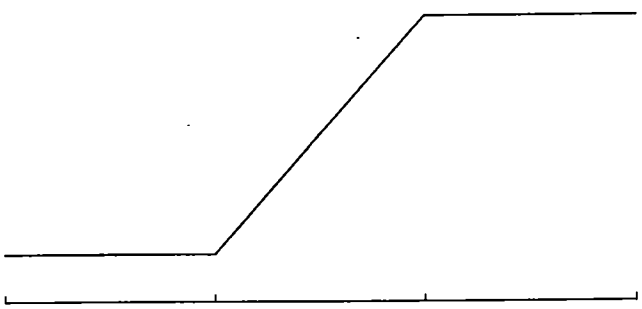
There are two types of learnings for neural networks. One is supervised learning where the networks is trained by provided inputs and corresponding desired outputs. The other is the unsupervised learning where each unit is trained to recognize the pattern of the input sets. In the latter learning process, the networks are supposed to discover statistically salient features of the input, and no priori knowledge of categorization of the input pattern is required. In this work, only the supervised training is considered.

As described in section 4.1, in the training stage, the weights of each connection get updated upon receiving the input signals. The basic idea is that if the errors of the desired activation of unit k are due to changing of the weight factor, w_{jk}, between the

(a) Sign function



(b) Linear/ Semi-linear Function



(c) Sigmoid function

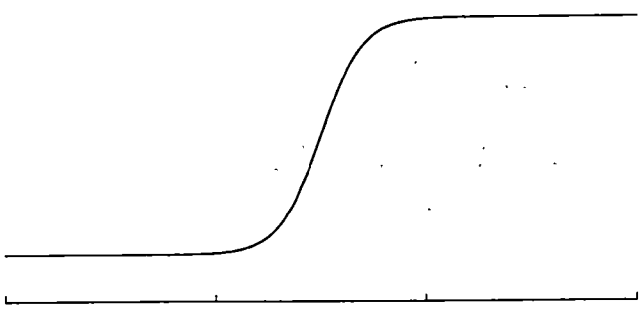


Figure 4.2: Activation functions

two unit j and k, their inter-connection need to be corrected by a factor of

$$\delta w_{jk} = \gamma\, y_j (d_k - y_k), \tag{4.6}$$

where γ is a constant representing the learning rate, and d_k is the desired activation of the unit k provided by the training set. This is called δ learning rule or the steepest gradient descent. In the training stage, the weights coming to the output layer are adjusted by comparing the actual processing results and the desired results, and the resultant desired activations of the hidden layer prior to the output layer are obtained. Then, the hidden layer performs the same procedure until all the weights coming to each hidden layer are corrected. In this fashion, the error between the processing results and the desired results is propagated backwards in the network, and the weights through the paths are adjusted. This type of neural networks are also called the back propagation networks.

The networks change the weights after each set of the training data is inputted. Having trained by all the data in the training set, the networks may not reach the minimum error, and multiple cycles of trainings are required until the asymptotic limit of the error is reached.

4.2.4 Single layer and multi-layer networks

For a single layer neural networks, there are no hidden layers, and the input units go directly to the output units. Thus, during the learning process, the weights of the

connections between each input and output pair get updated directly. From equation 4.6, it is obvious that the changes of weights are linearly proportional to the activation function in this case. If the total error function E_k of the output unit k prior to each update is defined by

$$E_k = \sum_p E_k^p = \frac{1}{2}\sum_p [d_k^p - y_k^p]^2, \tag{4.7}$$

where E_k^p represents the error on pattern p, p ranges over the set of input patterns. The procedure finds that the minimum error has achieved when all the values of the weights are adjusted by

$$\delta\delta w_{jk} = -\gamma(\partial E^p)/(\partial w_{jk}), \tag{4.8}$$

which is proportional to the negative of the derivative of the error as measured on the current pattern with respect to each weights. Note that y_k^p is a linear function of its input x_j, i.e.,

$$\frac{\partial y^p}{\partial w_{jk}} = x_j, \tag{4.9}$$

and note that

$$(\partial E^p)/(\partial w_{jk}) = \frac{\partial E^p}{\partial y^p}\frac{\partial y^p}{\partial w_{jk}}, \tag{4.10}$$

thus,

$$\frac{\partial E^p}{\partial w_{jk}} = -(d_k^p - y^p), \tag{4.11}$$

is hold, therefore,

$$\delta\delta w_{jk} = \gamma\delta^p x_j, \tag{4.12}$$

where $\delta^p = d^p - y^p$. This means only a linear system could be analyzed. However, the goal of the NN analysis is to study non-linear systems. The introduction of hidden layer will solve this problem.

Under the generalized δ learning rule, the multi-layer back-propagation networks is suitable to solve the problem of nonlinear system. The basic idea is that when a learning set is inputted, the activation values are propagated to the output units, and the actual outputs of the networks are compared with the desired output values. This will result in an error for each of the output units, say e_o for output unit o. Therefore, all the preceding weights of all the output units and hidden units have to be changed accordingly in order to minimize the error. The general form is given by

$$\delta w_{jk} = -\gamma \partial e_o / \partial w_{jk}, \tag{4.13}$$

and the partial derivative in equation 4.13 can be evaluated by the chain rule. The adjustment of the weights preceding the output layer is no longer linearly proportional to the inputs.

4.2.5 Addition of dimensions of input data

In the case where the existing networks is not able to distinguish the data patterns and data belonging to different group are merged, addition of extra dimensions of information may be helpful. However, these extra information needs not to be any linear combination of the existing inputs.

4.2.6 Discussion of the classification power of neural networks

In data categorization, there are some possible categories that are difficult to statistically calculate and to be included by any formulated analysis. Also, there may be other factors that influence the statistics in a manner that is difficult to track, but which may be helpful to include in the categorization. In a practical SMD experiment, such information includes the brightness of each photon burst, the possibility of photon-degradation, inter-system crossing, or multiple molecules passing simultaneously through the probe region. Furthermore, the assumption of independent photon statistics within each spectral bin used by MLE analysis may not be accurate in all occasions. By contrast, neural networks methods make no assumption on the statistics, but develop their own rules during the learning phase. The developed rules are thus purely empirical and can more easily account for complicating factors or alternate categorizations.

There are two main issues to be discussed in regard to whether the neural networks method would give adequate performance. The first is whether the weights and bias at each unit have been optimally set for the classification task at hand. This is the issue of learning, and there are several different convergent algorithms by which the weights may be adjusted during the training of the networks. To obtain the optimal weights and bias, one must ensure that the training set adequately spans the range of possibilities that neural networks will encounter, and make sure that the training set has enough distinguishable properties for different groups. The information content of the learning data is crucial for the classification ability of a networks.

The second is the issue of representation, i.e., assuming that the weights have somehow been optimally chosen, does the neural networks have enough flexibility or degree of freedom to give correct results for all possible inputs? In this regard, it has been proven[18] that a single layer (no hidden layer) neural networks has severe limitations in its representation power. However, when one or more hidden layer are added into the networks, the representation power will be considerably enhanced. In addition, the universal approximation theorem states that just a single hidden layer is sufficient to approximate any function to arbitrary precision provided that the activation function in the hidden layer is nonlinear. The currently used architecture in the SMD data analysis provides sufficient representational power for most of the classification problems, and also enables moderately fast speed for training.

4.3 Implementation of the NN analysis into the SMD simulation

4.3.1 Construction of the training and testing data

In order to implement the neural networks analysis into the SMD simulation, some changes have to be made in the simulation code so as to output the data for the training and testing by the neural networks. Recall that, in the SMD simulation, having located the starting and ending points of a detected burst, all the photons in a burst are then sorted into temporal bins within the time window, as described in section 2.4, and hence

the photon counts in each bin can be directly used as the inputs to the neural networks. In addition, the peak amplitude and duration of each burst are also recorded for the NN analysis as the extra dimensions of input, which might be benificial to the NN analysis, as discussed in section 4.2.5. As described in section 2.2, the type of each photon burst is known in the simulation, thus the actual type of the burst can be used to construct the training data for the NN analysis. Furthermore, it can be used to estimate the effectiveness of the data analyses, and this is done by comparing the predictions by either the NN or the MLE method with the desired types.

Note that training data and testing data are require to be generated under the same experimental condition. Also, the testing data could not be statistically related to the training data. A general approach is to generate statistically significant number of bursts data under certain experimental condition, and then to randomly take a portion of the data as the training sets and the remaining of the data as the testing sets.

4.3.2 Tool for the NN analysis

The "NeuralWorks Professional II/Plus", by NeuralWare, Inc, is utilized to perform the NN analysis. The software runs under 32-bit Microsoft Windows environment, and has the ability to construct up to 28 neural networks models plus their variations.

The software provides a user-friendly interface for constructing neural networks, and parameters, such as, the number of layers, units in each layer, type of activation function and learning rule, etc., are specified in this stage. In addition, the software is able to train the networks, and to test the trained networks by user provided data.

Another important feature of the software is that it is able to output the trained networks in the form of an ANSI C subroutine, which can be called in a real time experiment, as being discussed in section 5.8. A typical pop-up window of constructing a networks is shown in figure 4.3.

Figure 4.4 is produced by the neural networks software during the set up and learning phase of the networks. The current value of weights is shown in the top right inset of the figure. The weights are initially set as random and during the training stage their values fluctuate while the weights are adjusted. The goal of learning is to minimize the root mean square (rms) error or the difference between the actual and desired output, which is monitored in top left inset of figure 4.4. The fluctuation of weights diminishes as the rms error approaches its asymptotic limit.

To implement the NN analysis into the experiment, preliminary calibrations need to be performed prior to the testing, i.e., in such calibrations, the experiment collects burst data of individual molecule or background for training the neural networks. A proposed procedure with calibrations will be introduced in chapter 5, and the results are shown in the section 5.8.

4.3.3 Neural networks architectures used in the SMD simulation

The 2-layer back-propagation δ learning rule networks are chosen to perform the data analysis for the SMD simulation, and the sigmoid function is used as the activation function for both the hidden layer and the output layer.

In this work, two types of input formats are used for the neural networks. One that

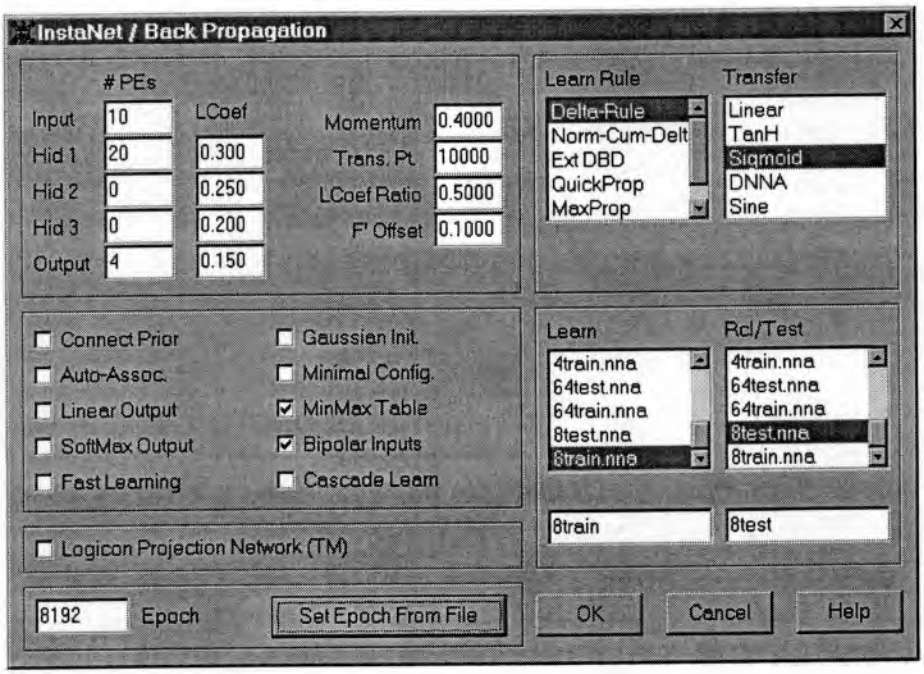


Figure 4.3: Graphic output of neural networks software package to construct a new networks

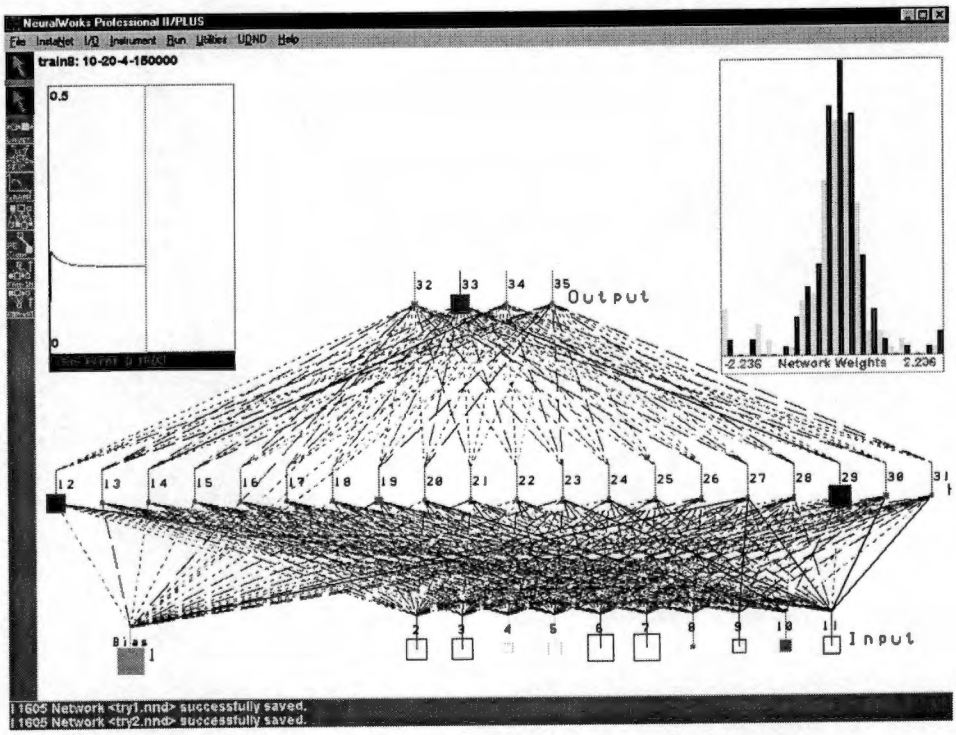


Figure 4.4: Graphic output of neural networks software package during the learning phase

is used in section 5.2 contains only the photon counts in each bin as the inputs; the other format that is used in sections from 5.4 to 5.9 contains also the peak amplitude and the burst duration as the extra inputs besides the photon counts. The peak amplitude, which denotes the brightness of each burst, and the duration of each burst are not linearly dependent on the photon counts in each bin.

For the improved version of the SMD simulation, the format of the output is set up to distinguish 4 types of events. The events and the corresponding output formats are as below:

- Molecule A: OUT_1=1, OUT_2=0,OUT_3=0,OUT_4=0
- Molecule B: OUT_1=0, OUT_2=1,OUT_3=0,OUT_4=0
- Backgrond: OUT_1=0, OUT_2=0,OUT_3=1,OUT_4=0
- Both molecule A and B present: OUT_1=0, OUT_2=0,OUT_3=0,OUT_4=1

For the prior version of the SMD simulation, there are only three groups being distinguished, i.e.,

- Molecule A: OUT_1=1, OUT_2=0,OUT_3=0,
- Molecule B: OUT_1=0, OUT_2=1,OUT_3=0,
- Both molecule A and B present: OUT_1=0, OUT_2=0,OUT_3=1.

As the sigmoid function is a continous function, and it provides float types of results at the output layer, a step function is usually applied to the testing resulting of the

neural networks. The processing result will be changed to 1 if it is larger than 0.5, or 0 if it is less than or equal to 0.5.

A rule of thumb for choosing the number of units in the hidden layer is that, for a single hidden layer networks, the number of units in the hidden layer are twice the number of units as in the input layer. For a two hidden layers networks, the first hidden layer has three times of units as in the input layer, and the second layer has one half number of units as in the first hidden layer.

In the example of the 2-layer neural networks shown in figure 4.4, the improved version of the SMD simulation has total of 10 inputs, which are composed of 8 inputs from the photon counts in each bin and 2 inputs from the peak amplitude and burst duration. According to the rule of thumb, 20 units are required in the hidden layer.

In the SMD data analysis, a further investigation of making direct connections between input and output layer is not promising, and the asymptotic limit of the rms error is not substantially decreased. Extensive numbers of other type of networks and their variations have been tried, and no obvious advantages have been observed. This indicates that the given architecture provides an adequate representation for the training set.

Chapter 5

Results and Discussions

5.1 Overview

As introduced in chapter 2, a burst of photons is detected when a molecule passes through the probe region, and these photons are sorted into a number of bins according to their arrival times with respect to the laser pulses. To identify each molecule, one analyzes the distribution of the photons among all the bins within the time window. Two methods are be used to analyze the data for each burst, the MLE method introduced in chapter 3 and the NN method introduced in chapter 4.

The development and implementation of the analysis software was undertaken before the Monte Carlo simulation was extended to consider the convolution of the fluorescence decay profiles with the instrument impulse response function, the temporal profile of the background, and the improved method for determining the start and end of each photon burst. Therefore, section 5.2 presents preliminary results of photon burst categorization

by the MLE and the NN methods using data generated by the previous version of the SMD simulation, described in section 2.2. The simulations used fluorophore parameters obtained from the literature and indicated that the dyes S-101 and R640, with lifetimes of 4.6 ns and 2.2 ns, would be readily distinguishable. In preparation for an experiment using these dyes, measurements were made of their fluorescence lifetimes, as reported in section 5.3. However, under the conditions of the SMD experiment, the dye R640 was found to actually have a fluorescence lifetime of 4.3 ns, comparable to that of S-101. Lifetimes of other dyes were also measured, but two dyes with disparate lifetimes, suitable for SMD with an excitation wavelength of 585 nm, could not be found. Therefore, the research efforts were directed towards the simulation of an experiment from the literature[9] on single molecule identification of R6G and TRITC, with fluorescence lifetimes of 4.2 ns and 2.5 ns and an excitation wavelength of 514.5 nm. The simulation was performed using the improved model described in section 2.4, and results are given in section 5.4. Section 5.4 also presents a comparison of the MLE and the NN categorization methods. The NN categorization method is found to be superior and reasons for this are discussed. In order to illustrate how the selection of the bins influences the precision of the identification of molecules, section 5.5 gives the results using different numbers of bins and different bin width selections.

When preparing for the experiment for single molecule identification of S-101 and R640, several improvements in the experimental hardware were made. Improvements aimed at speeding up single molecule detection are discussed in section 5.6, and results

show that molecules can be detected with a transit time less than 200 μs. In order to achieve identification of photon bursts in such experiments, low dead time circuit was developed for temporal binning of photons, as described in section 5.7. The circuit would be usable in experiments using other excitation wavelengths if suitable optical filters were to be purchased. Because such filters were not available, only simulations of single molecule identification experiments using such low dead time circuit could be pursued.

In section 5.8, a procedure for training of neural networks using experimental calibration data is proposed and evaluated using simulated data. The results indicate that the procedure is viable. The low dead time circuit for temporal binning of photons may result in the overlapping of bins, and therefore the results of simulations with such overlapped bins are presented in section 5.9. Lastly, the issue of the time required for the NN analysis is discussed in section 5.10.

5.2 Preliminary investigation with the previous version of the simulation

5.2.1 Conditions of the simulation

In this section, both the MLE and the NN analyses are applied to data generated by the previous version of the SMD simulation. In the previous version of the SMD simulation [12], only a small constant background, and a selection of the time window with fixed

width of 13.2 ns was considered, as described in section 2.4. A total of 4 bins are used to sort photons. The intuitive algorithm introduced in section 2.5.2 is applied to select the widths of the four bins. The parameters used in this simulation are listed in Tables 5.1–5.4.

As discussed in section 4.2, the NN analysis uses a 2-layer (1 hidden-layer) architecture, and back-propagation δ learning rule. A sigmoid function is applied as the activation function to both the hidden layer and the output layer. The photon counts in the 4 temporal bins are used as the 4 inputs of the neural network. By the rule of thumb described in section 4.3.2, eight processing units are used in the hidden layer. The data for each burst are saved in the required format for the analyses by the MLE and the NN methods, as introduced in sections 3.3 and 4.3, respectively. A step function is applied to all the testing results of the NN analysis, and the values of all the outputs from the networks are changed to either 0 if they are less than or equal to 0.5, or 1 if they are larger than 0.5. Following this, the testing results have the same formats as defined in section 4.3.2, representing different types of molecule detection events.

Table 5.1: Laser related parameters

Variable Name	Description (Units)	Value
lambda	laser wavelength (μm)	0.585
p_energy	calculated by hc/λ(J)	3.4×10^{-19}
rep_rate	repetition rate of laser pulses(Hz)	7.6×10^{7}
pulses_per_time_step	laser pulses per defined time step	1
w0y/w0z	beam waists in y or z (μm)	4.5
power	laser power(W)	0.1

Table 5.2: Optical parameters in the light path

Variable Name	Description	Value
pinhole	diameter of the spatial filter (μm)	600
magnification	magnification of the collection objective	60
NA	numerical aperture of the collection objective	0.85
image_loss	optical losses due to uncoated optics	0.75
filter	throughput of interference filter	0.5
spad_eff	quantum efficiency of SPAD	0.65
tgate	time gate efficiency for passing signal	0.6
refract_index	refractive index of the solvent	1.33
tip_radius	capillary tip radius (μm)	0.5
tip_to_beam	distance from the tip to beam (μm)	15

Table 5.3: Photo physical parameters for the molecules

Variable Name	Description	Value	
		S-101(A)	R640(B)
Fluor_QE	fluorescence quantum efficiency	0.35	0.35
SIGMA	absorption cross section (μm^2)	3.7×10^{-8}	3.7×10^{-8}
PHI_D	photon-destruction quantum efficiency	5.5×10^{-5}	5.5×10^{-5}
PHI_CROSS	triplet crossing probability	0.002	0.002
lifetime	fluorescence lifetime(ns)	4.6	2.2

Table 5.4: Other simulation parameters

Variable Name	Description	Value
dark_in_window	detector dark count (counts/s)	50.0
noise_in_prompt	counts $s^{-1}W^{-1}m^{-3}$	3.5×10^5
noise_in_window	counts $s^{-1}W^{-1}m^{-3}$	3.48×10^3
diffusion	molecular diffusion coefficient ($\mu m^2 s^{-1}$)	4.5×10^2
ratio_0	fraction of molecule of type A	0.5
concentration	solution concentration(M)	1.0×10^{-11}
flow_rate	sheath flow rate (μm s^{-1})	1.0×10^4
ns_SPAD_dead_time	dead time of SPAD (ns)	65
ns_TAC_dead_time	dead time of TAC (ns)	1000
NUM_OF_BINS	number of temporal bins	4
thres	threshold introduced in section 2.3	30

5.2.2 Results and discussions

Over a simulation run time corresponding to 231.6 seconds, 8161 photon bursts were recorded. The data from the first 4081 bursts are used for training the neural network, and that from the remaining 4080 bursts are tested by both the MLE and the NN methods. In the first 4081 data, 1945 are due to the passage of only molecule type A, 1917 are due to only molecule type B, and 219 contain photons from both A and B. Among the last 4080 data, 1903 are due to the passage of only molecule type A, 1966 are due to the passage of only molecule type B, and 211 contain photons from both types of molecules and thus are specified as being due to the simultaneous passage of both types.

According to the intuitive bin width selection algorithm, the widths of each of the bins are chosen to make the two types of molecules most distinguishable, as discussed in section 2.5.2. The photons in all bursts generated by each type of molecule should be distributed within the bins in a different manner. To illustrate this, the average numbers of photons in each bin for bursts due to molecule type A, and due to molecule type B are calculated for the testing data, and the results are drawn in figure 5.1. Photons from bursts due to molecule type A (τ_A=4.2 ns) yield an increasing distribution, while those due to molecule type B (τ_B=2.6 ns) yield a decreasing distribution over the bins. This is consistent with the motivation of the intuitive bin width selection algorithm. The presence of background photons within the bursts can reduce the difference between the two distributions.

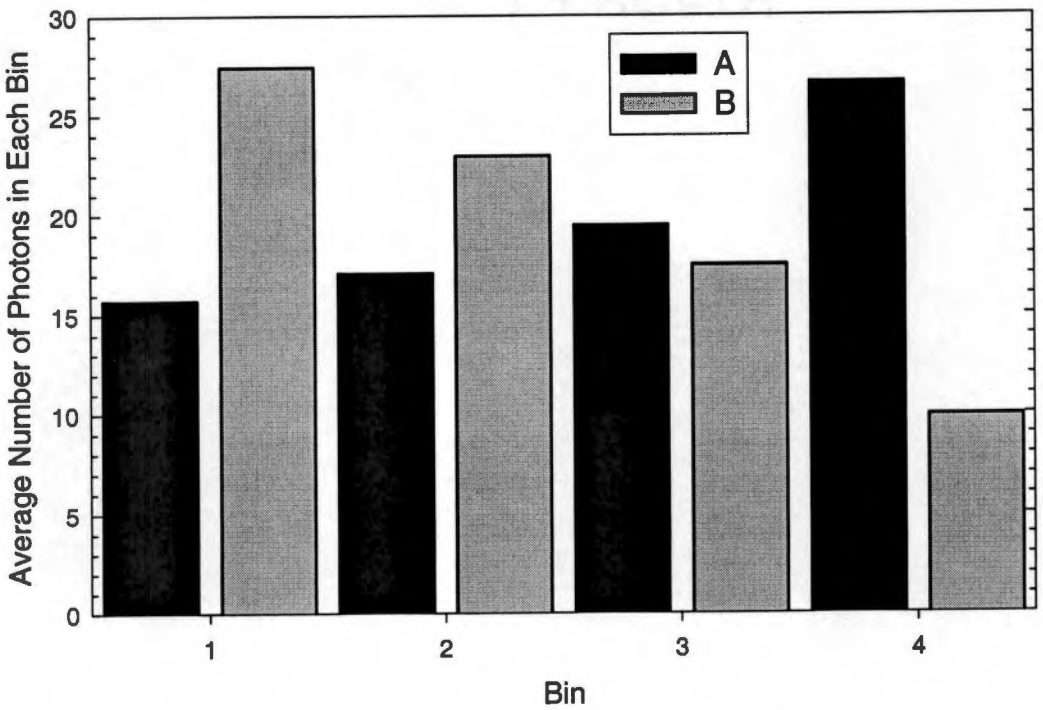


Figure 5.1: Average photon distribution in each temporal bin

The results of the predictions by the MLE and the NN analysis, listed in tables 5.5 and 5.6, show that, both methods categorize the bursts well when only one type of molecule passes through the probe region. However, bursts that have been attributed to both A and B are not predicted by either method. The MLE method is of course limited by its definition in section 3.2 so that it is not an option to predict both A and B. The NN method is more flexible in that both A and B is an option, but it fails to identify any of the 211 bursts attributed to both molecules. Actually, this is most probably a consequence of the algorithm by which the bursts were located and initially identified in the simulation, rather than a failure of the NN method. In the previous version of the simulation, the start and end points of bursts are located at the transition points of the weighted sliding sum, as described in section 2.6.1. This causes one burst to immediately follow another and hence photons from one burst to often fall within the limits of an adjacent burst. Consequently, a relatively large number of bursts (211) are attributed to both molecules, but the photons in many of these bursts would be predominantly from one species.

The expected number of bursts due to the simultaneous passage of A and B can be

Table 5.5: Identification of burst type by the MLE method

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	1903	1734(169)	91.1
molecule B	1966	1886(80)	95.9
both A and B	211	0(211)	0.0
Total	4080	3620(460)	88.7

Table 5.6: Identification of burst type by the NN method

Type of burst	Actual Occurrence	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	1903	1839(64)	96.6
molecule B	1966	1922(44)	97.8
both A and B	211	0(211)	0.0
Total	4080	3761(319)	92.2

estimated as follows. The parameters of the simulation are such that the mean transit time of a molecule across the laser beam is 0.9 ms and the mean time between molecules is 21.15 ms, and thus the probability that a second molecule enters the laser beam before the first one leaves is $1 - \exp(-0.9/21.15) = 0.04$. If the molecule detection efficiency were 100%, approximately 5256 bursts would be expected during a 115.8 s interval, including only about 110 due to both A and B.

To illustrate how the NN performs the analysis in general, differences in the information content of the data from different types of bursts are considered. The input data set from each burst is multidimensional, and in the case of 4 inputs corresponding to the number of photons within 4 temporal bins, the number of dimensions is 4. If bursts are to be successfully identified, the data sets from different types of bursts must separate into regions within the multidimensional space that are largely non-overlapping. Further, the NN must sufficiently grasp the differences between the different regions, so as to categorize the bursts. In order to demonstrate the salient differences in a 2-dimensional plot, figure 5.2 plots the fraction of photons within the first two bins versus the total number if photons. A clear separation of the data from molecule type A and

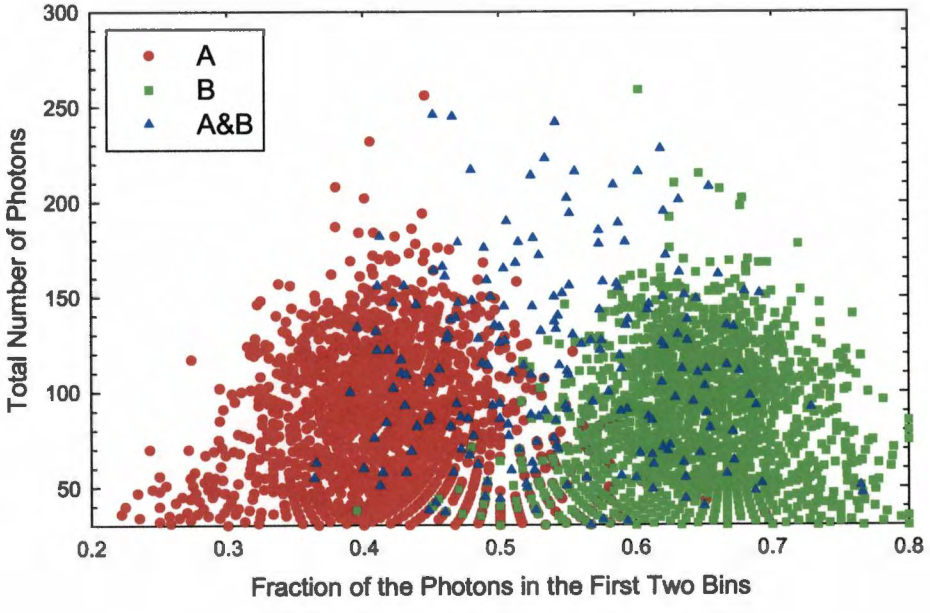


Figure 5.2: Fraction of the first two bins among the total number of photons in the burst

molecule type B is apparent, and this depends primarily on the difference of the fluorescence lifetimes of the two types of molecules. The more the lifetimes of the two types differ, the further the two groups of data are separated, and the better the chance that they can be distinguished. Therefore, the choice of dye molecules with widely different lifetimes is crucial for effective categorization.

If the lifetimes of the two types are well separated, either the MLE or the NN analysis will work. However, the MLE method, as defined in section 3.2, is only able to utilize the information pertaining to differences in the fluorescence lifetimes and is unable to classify bursts that are separated by other factors. On the other hand, the NN analysis utilizes all information in an empirical manner, and is able to discern differences that are not readily apparent, or for which a statistical analysis would be difficult to formulate. Such information could be originated from the brightness of each burst, photo-degradation, or inter-system crossing, etc. The results obtained by the NN method, shown in table 5.6, generally gives better predictions than the MLE analysis, shown in table 5.5.

5.3 Experimental measurements of lifetimes

To further investigate the possibility of an experiment for the single molecule identification, the fluorescence lifetimes of the dyes R640, and S-101, which were used in the simulation in section 5.2, are measured.

The measurement is performed by the time correlated single photon counting tech-

nique, which is explained in detail in section 5.6.1. The experimental setup is similar to that of section 5.6.1 except that the data is collected by a multi-channel analyzer (MCA). In the measurement, micro-molar concentrations of dye solutions are flushed into the sheath flow capillary. The MCA collects data for several minutes, and saves the data in CHN file format. The data are processed as described in section 2.4.2, and results are shown in figure 5.3. The figure also contains the curve for dye DCM-S.

In reference [25], the fluorescence lifetimes of S-101 and R640 were taken to be 4.6 and 2.2 ns, respectively. However, at an excitation wavelength of 585 nm, the fluorescence lifetimes of S-101 and R640 are found to be approximately 4.3 ns and 4.4 ns, which are not sufficiently different from each other for the single molecule identification experiment. Although the fluorescence lifetime of DCM-S is found to be about 2.7 ns, its absorption spectrum has a peak at 475 nm, and little absorption exists at 585 nm. It was not possible to detect single molecules of DCM-S in our system.

The dye DODCI was also considered. It has moderate absorption at 585 nm. However, its fluorescence lifetime has been measured to be 0.66 ns [26], which is rather short and indicates a low fluorescence quantum yield. Because of the short lifetime, few photons would be captured beyond the instrument prompt and within the time window. Indeed, it was found to be not possible to detect single molecules of DODCI within our instrument.

Although two dyes suitable for single molecule detection with an excitation wavelength of 585 nm and with disparate lifetimes could not be found, such dyes are known

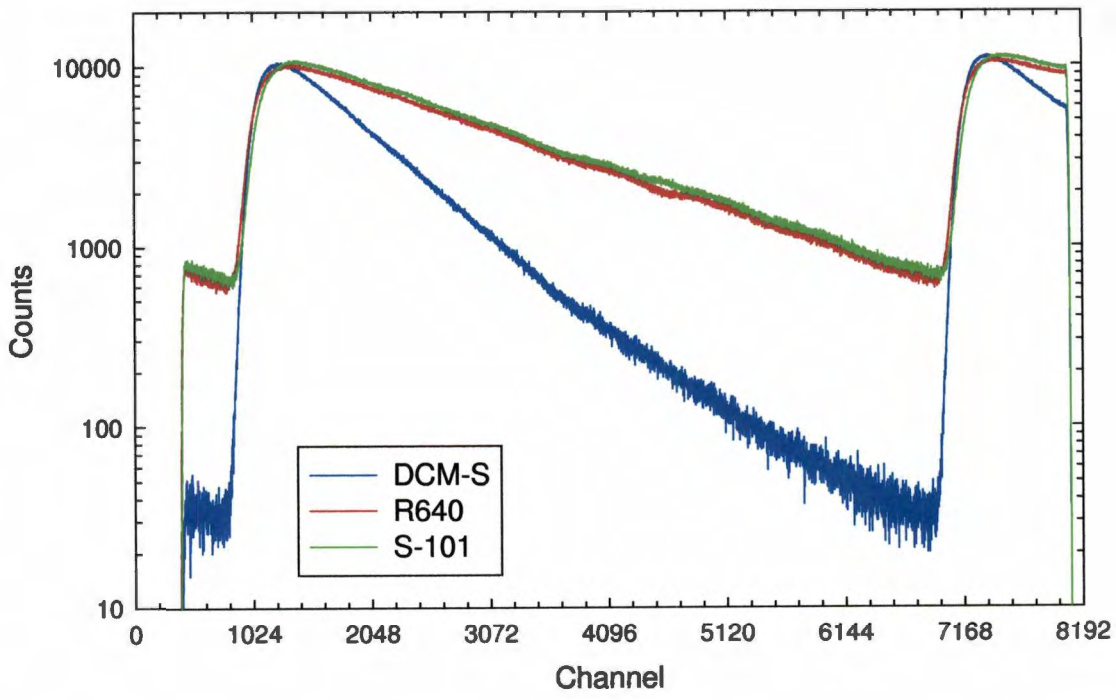


Figure 5.3: Measurements of the fluorescence lifetimes of S-101 and R640

to exist for use of other wavelengths. However, it had been decided not to change the experiment to another excitation wavelength, because of other ongoing research projects, and also the high cost of the optical filters that would be needed. Therefore, efforts at this point were directed towards the simulation of a single molecule identification experiment that had been recently reported in the literature, using an excitation wavelength of 514.5 nm.

5.4 The SMD simulation with the improved version of the simulation

5.4.1 Conditions of the simulation

By using the improved version that considers the background, as introduced in section 2.4, a similar approach as in section 5.2 is carried out in this section to simulate the experiment published in reference [9]. As the experiment in the literature was done by using a laser wavelength at 514.5 nm, which is not usable in our lab, only the simulation is presented in this section.

The parameters used in the simulation are obtained from references [9], [10] and [11], and are listed in Tables 5.7–5.10. Parameters not listed here are the same as those given in section 5.2. The photo-physical constants of the dyes, R6G (type A) and TRITC (type B), as listed in table 5.9, are obtained from reference [17]. The parameter "beta_prompt" in table 5.10, which denotes the degree of the background noise from

Table 5.7: Laser related parameters

Variable Name	Description	Value
lambda	laser wavelength (μm)	0.5145
p_energy	calculateed by hc/λ(J)	3.866×10^{-19}
rep_rate	repitition rate of laser pulses(Hz)	8.2×10^{7}
pulses_per_time_step	laser pulses per defined time step	50
w0y/w0z	beam waists in y or z (μm)	10.0
power	laser power(W)	0.03

Table 5.8: Optical parameters in the light path

Variable Name	Description	Value
pinhole	diameter of the spatial filter (μm)	500
magnification	magnification of the collection objective	40
tip	capillary tip radius (μm)	2.0
tip_to_beam	distance from the tip to beam (μm)	50.0

Table 5.9: Photo physical parameters[17] for the dye molecules

Variable Name	Description	Value	
		R6G(A)	TRITC(B)
Fluor_QE	fluorescence quantum efficiency	0.45	0.15
SIGMA	absorption cross section (μm^2)	2.2×10^{-8}	1.6×10^{-8}
PHI_D	photon-destruction quantum efficiency	1.9×10^{-5}	5.6×10^{-6}
lifetime	fluorescence lifetime(ns)	4.2	2.5

Table 5.10: Other simulation parameters

Variable Name	Description	Value
beta_prompt	prompt photons per second	1.0×10^5
total_dark	average dark counts per laser pulse	220
diffusion	molecular diffusion co efficiency ($\mu m^2 s^{-1}$)	4.5×10^2
ratio_0	fraction of molecule of type A	0.33
concentration	solution concentration(M)	5.0×10^{-13}
flow_rate	sheath flow rate (μms^{-1})	5.0×10^3
ns_SPAD_dead_time	dead time of SPAD (ns)	65
ns_TAC_dead_time	dead time of TAC (ns)	0
NUM_OF_BINS	number of temporal bins	64
convChNumber	number of bins per period of laser pulse output from the MCS	2090
startCh	start bin of time window	200
endCh	end bin of time window	1800
thres	threshold introduced in section 2.3	10

laser pulses, is set to be 1.0×10^5 counts per second so as to obtain approximately 5% of the bursts due to background. Other parameters, such as the distance between the tip and beam, and capillary tip radius in table 5.8, are obtained from the experiment current developed in our lab [10], and [11]. As the fluorescence lifetimes of R6G and TRITC are not as well separated as the two dyes in the simulation of section 5.2, poorer differentiation between the dyes is expected.

In the simulation, photons may originate from background or from fluorescence from either type of molecule. Each photon burst is then categorized as due to molecule type A, type B, both A and B simultaneously, or background only. As mentioned above, the time window is set so as to obtain 5% of bursts due to background among all the collected bursts. The setting of the time window effectively reduces the influence from

the background noise, and in this case, the time window is started at approximately 1.3 ns after the peak of the laser pulse and ended at about 1.9 ns before the next pulse, as compared to the time window setting in reference [9]. The bin widths are chosen by the intuitive algorithm introduced in the section 2.5.2. In addition, the dead time of Time-to-Amplitude converter (TAC) is set to be zero because a high-speed multi-channel scalar (MCS) is used to collect data in the experiment in replacing of the TAC with long dead time. A total of 64 bins are chosen within the time window, and the data for the case of 4 bins is obtained by joining adjacent bins, as described in section 2.5.4. The results for other numbers of bins(i.e., 64, 32, 16, 8, 2, and 1 bins), will be presented in section 5.5. In addition to the numbers of photons in each bin, the duration and the peak amplitude of each burst, are also recorded and taken as the inputs for the NN analysis.

In the NN analysis, a 2-layer (1 hidden-layer) back-propagation network is constructed. The δ learning rule and sigmoid activation function are applied to the network. The number of processing units in the network is obtained by the rule of thumb described in section 4.3, and all information pertaining to the architectures is summarized in table 5.11.

5.4.2 Results and discussions

The improved version of the SMD simulation is used to generate a total of 16384 bursts, of which the first 8192 are used as the training set for the NN analysis. The prediction of the identities of each of the bursts is obtained for the last 8192 bursts by using both

Table 5.11: The NN parameters selected for the data analysis

Number of bins in time window	Number of input units	Number of units in the hidden layer	Number of output units
64	66	132	4
32	34	68	4
16	18	36	4
8	10	20	4
4	6	12	4
2	4	8	4
1	3	6	4

Table 5.12: Identification of molecule by MLE in 4 bin case

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3662	2943(719)	80.4
molecule B	4233	2330(1903)	55.0
Prompt	131	0(131)	0.0
both A and B	166	0(166)	0.0
Total	8192	5273(2919)	64.4

the MLE and the NN methods. Among the training set for the NN analysis, there are 3725 bursts due to the passage of molecule type A, 4184 bursts due to the molecule type B, 129 bursts due to the background, and 154 due to both A and B. Among the testing data set, there are 3662 bursts due to the passage of molecule type A, 4233 bursts due to the molecule type B, 131 bursts due to the background, and 166 due to both A and B. The results of the prediction by the MLE and the NN methods are shown in tables 5.12 and 5.13, respectively.

The results show that neither the MLE nor the NN method performs as well as in section 5.2. The NN analysis out-performs the MLE method when only a single type

Table 5.13: Identification of molecule by NN in 4 bin case

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3662	3282(380)	89.6
molecule B	4233	4153(80)	98.1
Prompt	131	0(131)	0.0
both A and B	166	0(166)	0.0
Total	8192	7435(757)	90.8

of molecule is present. However, the NN analysis is not able to correctly categorize the bursts due to both molecule A and B present simultaneously, and it incorrectly categorizes such events as due to either type A or B only.

To understand why both methods predict poorer results, the information content of the data is investigated. The fraction of photons that fall within the first two bins versus the total number of photons in the burst is plotted in figure 5.4. Bursts corresponding to molecule type A with a longer fluorescence lifetime are centered at 0.4 in the x axis, i.e., on average, such bursts distribute 40% of their photons into the first two bins and the remaining 60% into the second two bins. For bursts due to molecule type B, the photons are distributed evenly between the two portions of the time window. The mean separation of the two groups is only about 0.1, which is smaller than the value of 0.5 obtained in section 5.2.2. Note that on the basis of the fluorescence lifetimes, the distance between the two groups is not expected to be this close. Also, molecule B, which has a shorter fluorescence lifetime, would be expected to distribute more than 50% of the photons into the first two bins. There are two reasons why the distributions

do not appear according to these expectations.

First, more of the bursts from molecule B are not detected, as molecule B is less bright. Recall that a threshold, which is set to be 10 in this case, is introduced to block the background bursts, and some of the bursts from molecule B will also be blocked. In the simulation, molecule B is introduced at twice the rate as that of molecule A, but the number of detected bursts due to B is only 1.15 times that of A.

Second, and more importantly, most bursts contain a large number of background photons, which distort the resultant distributions. As shown in section 2.4, if not for background, the distribution of photons into each bin would follow the probability density of the corresponding molecule type, as illustrated for molecule A, B, and background in figure 5.5(a). However, when background is included, the number of photons in each bin for bursts due to A or B becomes distorted as shown in figure 5.5 (b). The curves corresponding to A and B have a similar shape, which is similar to the probability density of the background in figure 5.5(a). Thus, if a burst contains too many background photons, the fluorescence photons will be overwhelmed by the background photons, and the timing information from the fluorescence decay profile will not be sustained.

The NN method yields more accurate predictions than the MLE method because of the same reason as discussed in section 5.2.2. In this case, the brightnesses of the fluorescence for each type of molecule differs, and is apparently used by the NN method to help classify the bursts. To illustrate how the classification of a burst is correlated to its brightness, a comparison between the MLE and NN methods is shown in figures

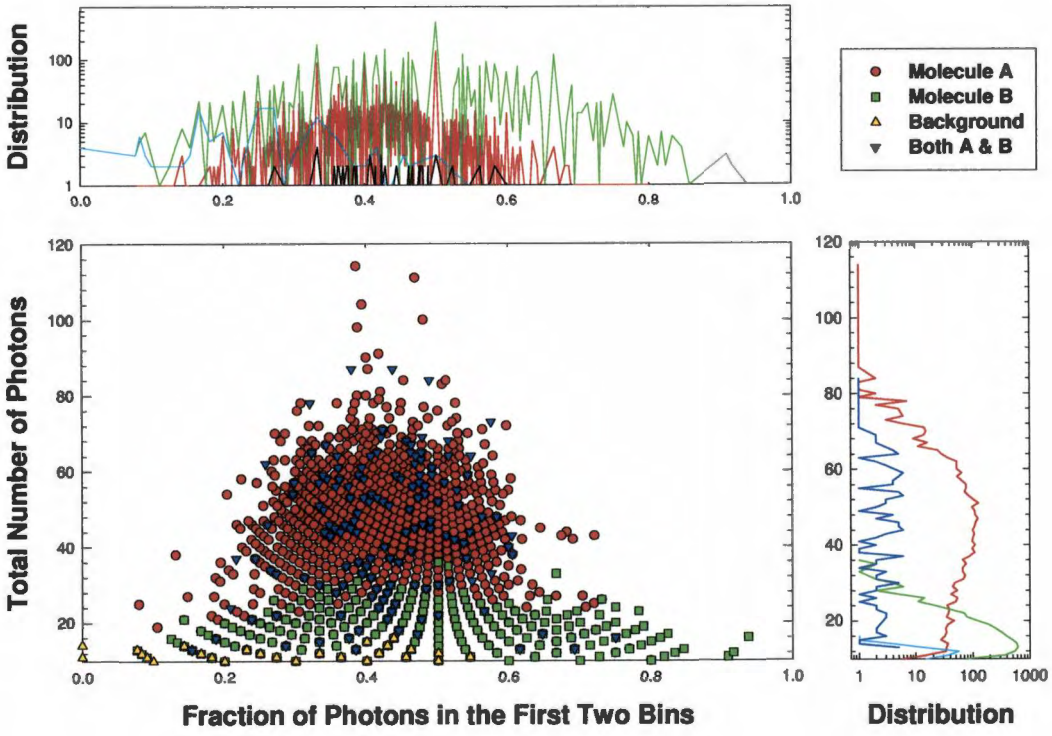


Figure 5.4: Fraction of photons in the first two bins versus the total number of photons in the burst

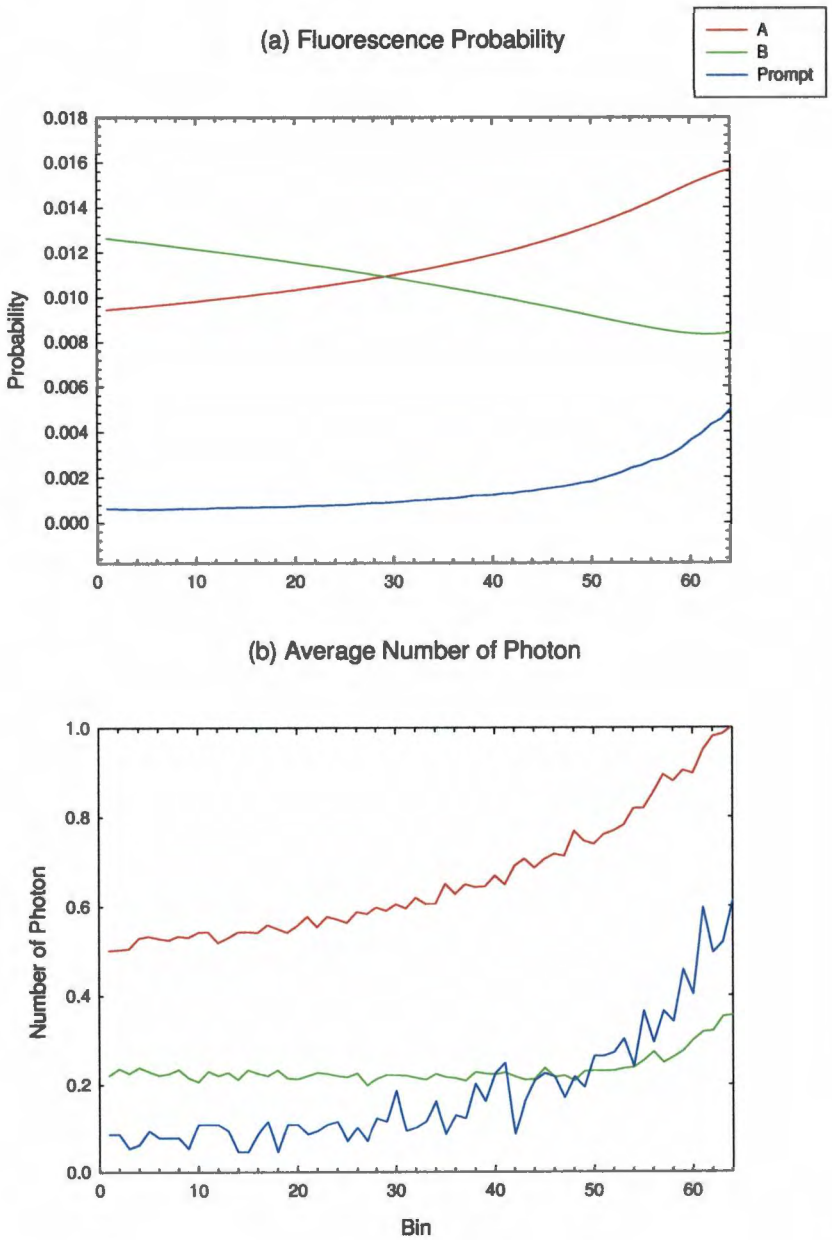


Figure 5.5: Probability distribution of photons of each burst for two types of molecules and prompt

5.6 to 5.9. Figure 5.6 shows the events predicted correctly by both methods, figure 5.7 shows predicted incorrectly by the MLE method but correctly by the NN method, figure 5.8 shows the events predicted correctly by the NN method but incorrectly by the MLE method, and figure 5.9 shows where both methods fail to correctly predict the events. The numbers of events predicted by each method are also included in each figure legend. Note that in figure 5.7, where the NN outperforms the MLE method, all the bursts from molecule A contain more photons than those from molecule B. However, in figure 5.8, where the MLE outperforms the NN method, there are smaller number of bursts and for these, molecule B is brighter than A. From this it can be inferred that the NN method yields more accurate predictions than the MLE method when the bursts from molecule A are brighter than those from molecule B. Note that molecule A has a higher quantum efficiency than molecule B, and hence its bursts are expected to be brighter.

In the cases where both methods fail, the bursts are due to background or both types of molecule, and are only 2% among all the collected events, all such bursts were categorized as single molecule events, i.e., either type A or type B. lots of bursts due to the prompt and both molecules presenting. In the simulation, the NN fails to predict the events where the burst is from either background or both types of molecules because the bursts belonging to these two cases are mixed with the cases where only single type of molecule is present, as illustrated in figure 5.9. Furthermore, the failure of classifying these cases by the NN method is expected since the majority (96%) of the bursts are

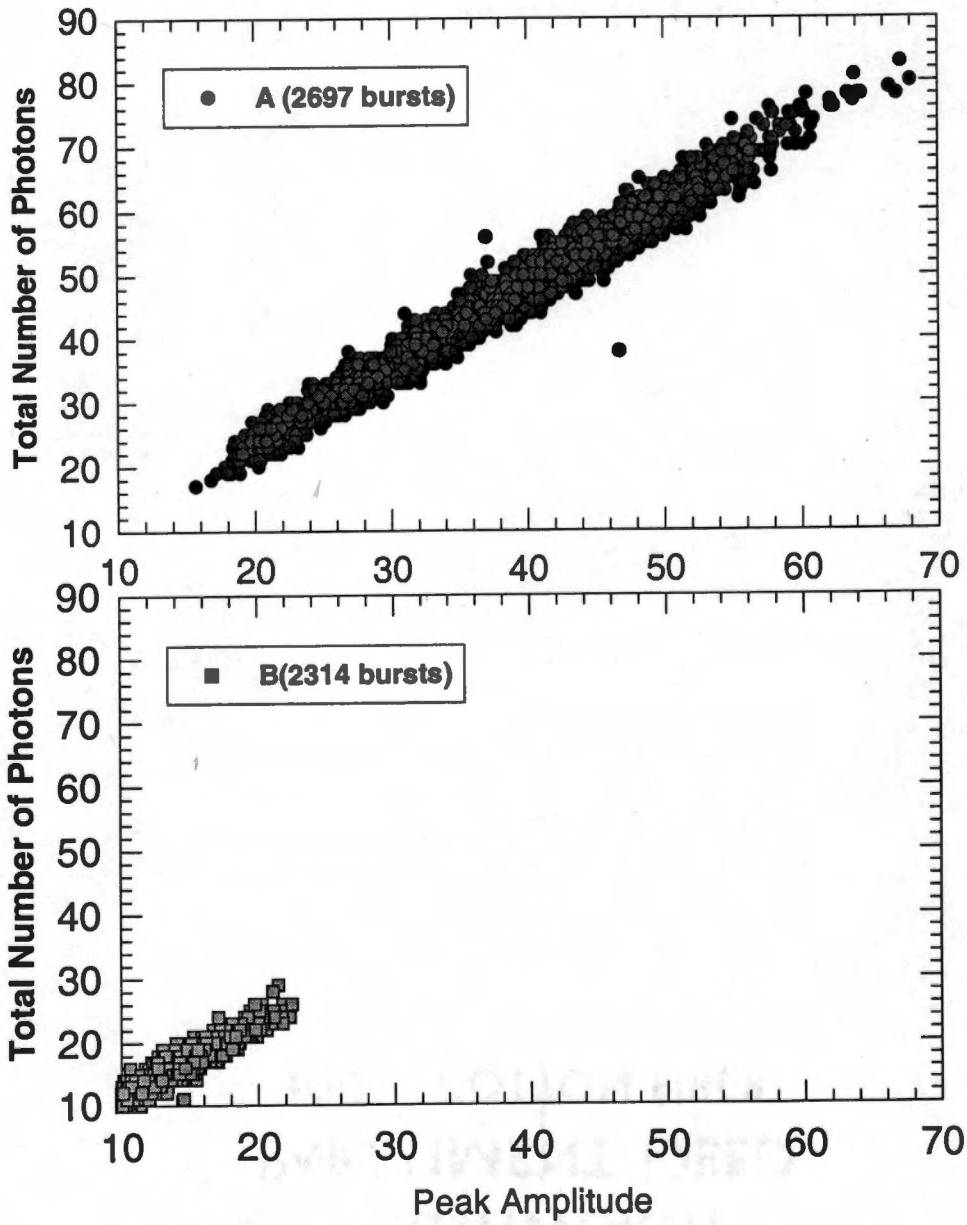


Figure 5.6: Both the MLE and the NN methods predict correctly

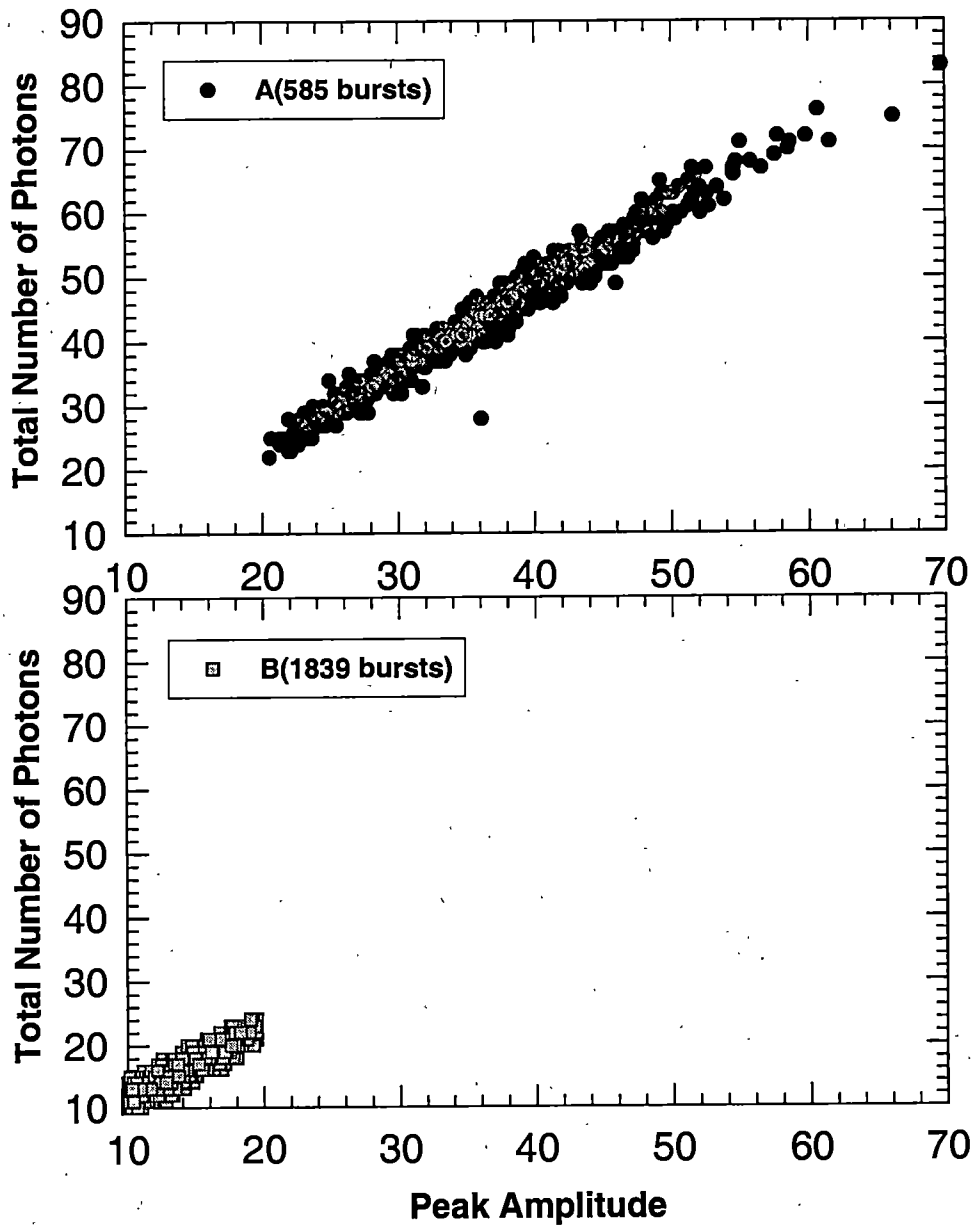


Figure 5.7: The MLE method predicts wrong while the NN method predicts correctly

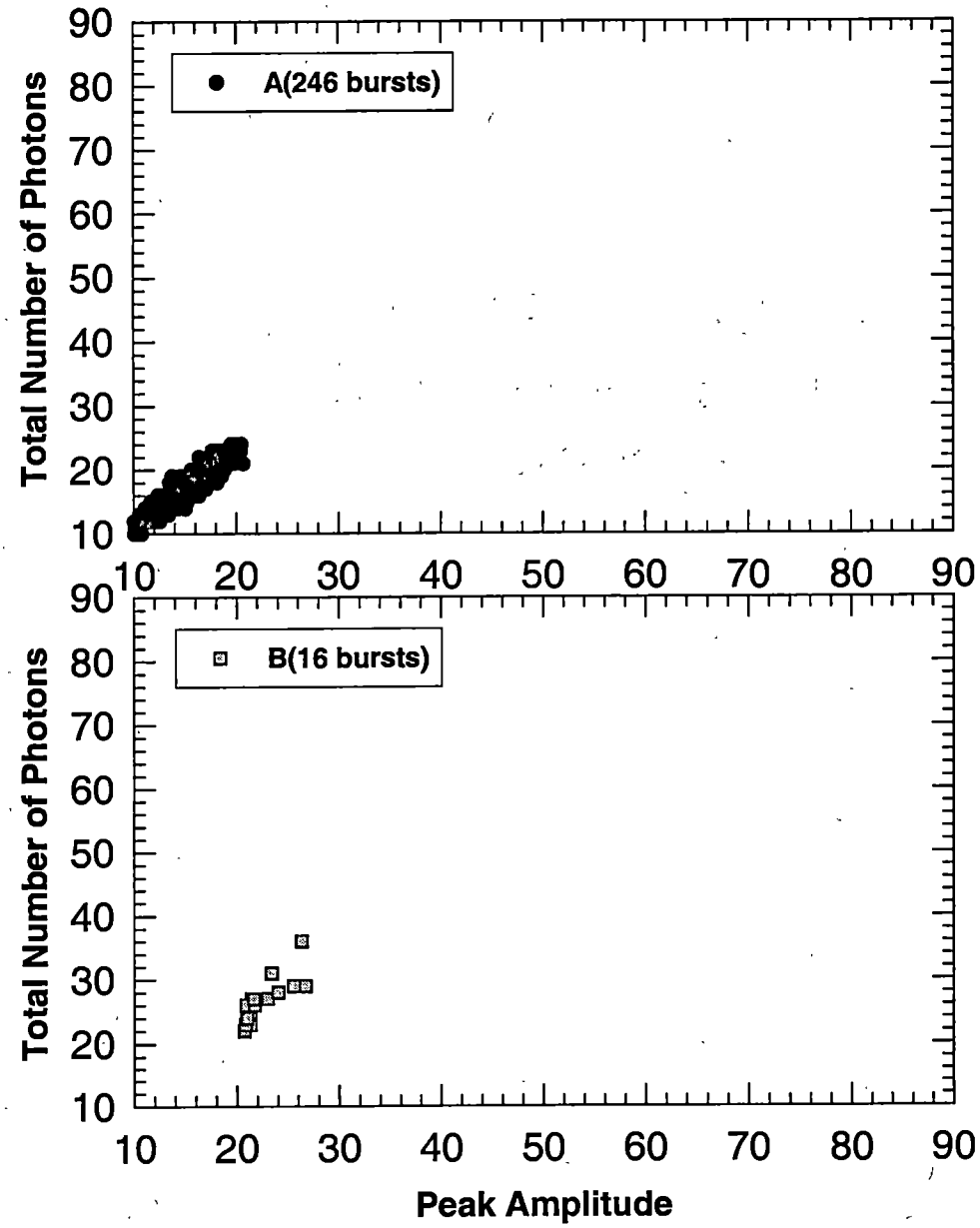


Figure 5.8: The MLE method predicts correctly while the NN method predicts wrong

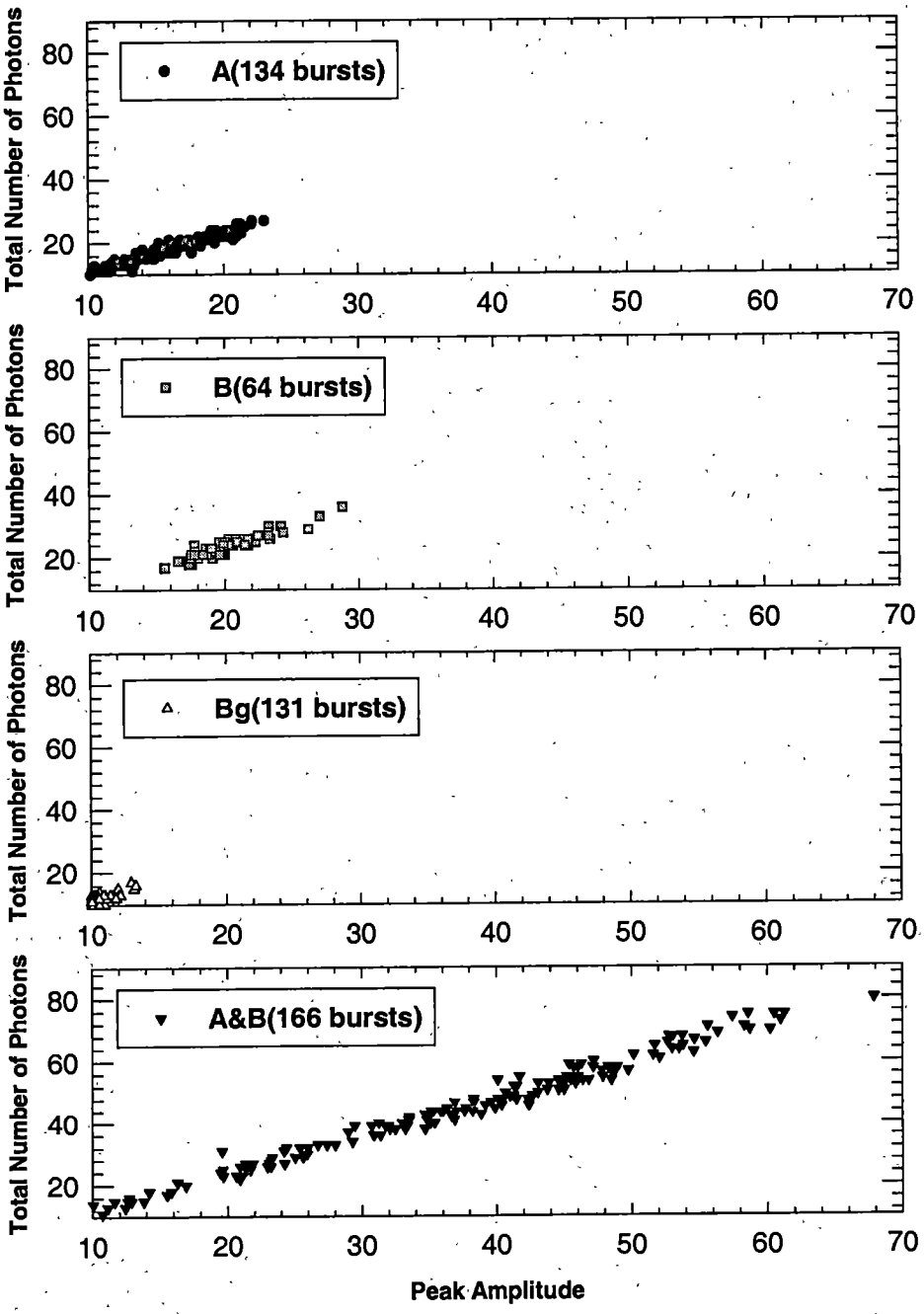


Figure 5.9: Both the MLE and the NN methods fail to predict

originated from the passage of only a single type of molecule in the learning set. The NN analysis always favors the majority of the data group so as to achieve the minimum error.

The results above are what are expected as the background was introduced into the simulation, and this also can be used as the direction of designing an efficient SMD experiment. Firstly, dye molecules with distinctive fluorescence lifetimes are preferred, as discussed in section 5.2.2. Secondly, the dye molecules need to have high quantum efficiency, and hence, under the high laser power, a burst would contain many photons from the fluorescence. Thirdly, the dye molecules with other photo-physical properties that discriminate them would help the NN analysis, for example, different internal conversion, inter-system crossing, and photo-degradation properties. Finally, the solution used in the experiment need to be diluted so as to avoid occasions where both types of molecules are present.

5.5 SMD simulation with different bin selection schemes

In section 2.5.1, the influences of choosing various number of bins within the time window on the precision of identification of molecules was discussed, and the conclusion was that only a small number of bins is needed to obtain fairly good results. Addition of extra bins within the time window is not expected to gain much benefit, and furthermore a small number of bins will speed up the process of data collection and analysis. In this section, the MLE and NN methods will be used to analyze the cases for which different

numbers of bins within the time window are chosen.

The simulations in this section use the same parameters as in section 5.4. Only the data for the case of 64 bins are generated by the simulations, and the other cases of 32, 16, 8, 4, 2, and 1 bins are obtained as described in section 2.5.4. The equal bin widths and bin widths selected by the intuitive algorithm, introduced in section 2.5.2, are used to select the bins within the time window.

There are 16384 bursts collected in the simulations, and among these bursts, the first 8192 are used for training the neural network and the last 8192 are used for testing by both MLE and NN methods. The results for the selection of equal bin widths is shown in table 5.14. The total number of correct predictions by the MLE method for each case of total number of bins within the time window is drawn in figures 5.10, where the squares correspond to the results from the equal bin widths selection. It is apparent that the reduction of temporal bins will lead to the decrement of total number of correct prediction when the resolution of the bins is less than 8. The NN analysis gives consistent results for all cases of equal bin widths, as shown in table 5.14 and figure 5.11.

Table 5.15 gives the results for the bin selection determined by the intuitive algorithm. The corresponding total number of correct predictions by the MLE method for each case of total number of bins within the time window is also illustrated in figure 5.10 (circles in the figure). In this case, the MLE method is able to give consistent prediction results for all the cases of bin selections. The NN method also gives results

Table 5.14: Results for even bin width selection case for various number of bins: A: Fluorophore A (R6G), B: Fluorophore B (TRITC), C: Background, AB: Both fluorophore

Number of temporal bins	Categories	A	B	C	AB	Total
	Actual number of cases	3661	4233	131	167	8192
64	Cases correctly predicted by NN	3249	4158	0	0	7407
	Cases correctly predicted by MLE	2965	2294	-	-	5259
32	Cases correctly predicted by NN	3242	4166	0	0	7408
	Cases correctly predicted by MLE	2966	2289	-	-	5255
16	Cases correctly predicted by NN	3251	4169	0	0	7420
	Cases correctly predicted by MLE	2966	2305	-	-	5271
8	Cases correctly predicted by NN	3246	4171	0	0	7417
	Cases correctly predicted by MLE	2955	2313	-	-	5268
4	Cases correctly predicted by NN	3274	4157	0	0	7431
	Cases correctly predicted by MLE	2902	2293	-	-	5195
2	Cases correctly predicted by NN	3261	4151	0	0	7412
	Cases correctly predicted by MLE	2749	2289	-	-	5038
1	Cases correctly predicted by NN	3254	4152	0	0	7406
	Cases correctly predicted by MLE	3661	0	-	-	3661
Learning set:	Actual number of cases	3725	4184	129	154	8192

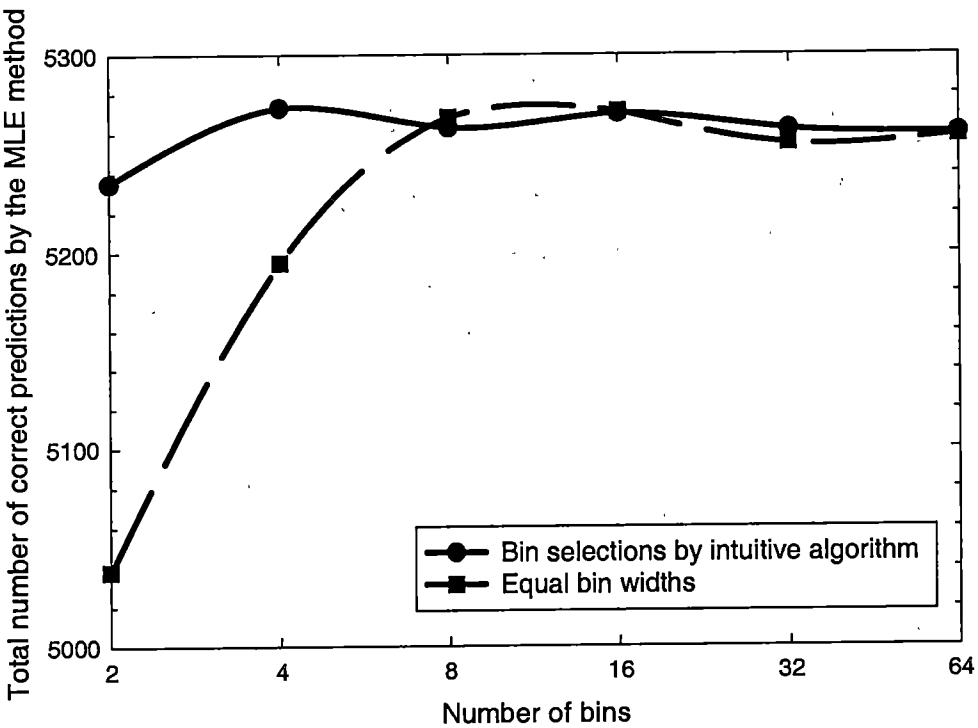


Figure 5.10: Total number of correct predictions by the MLE method versus the number of bins

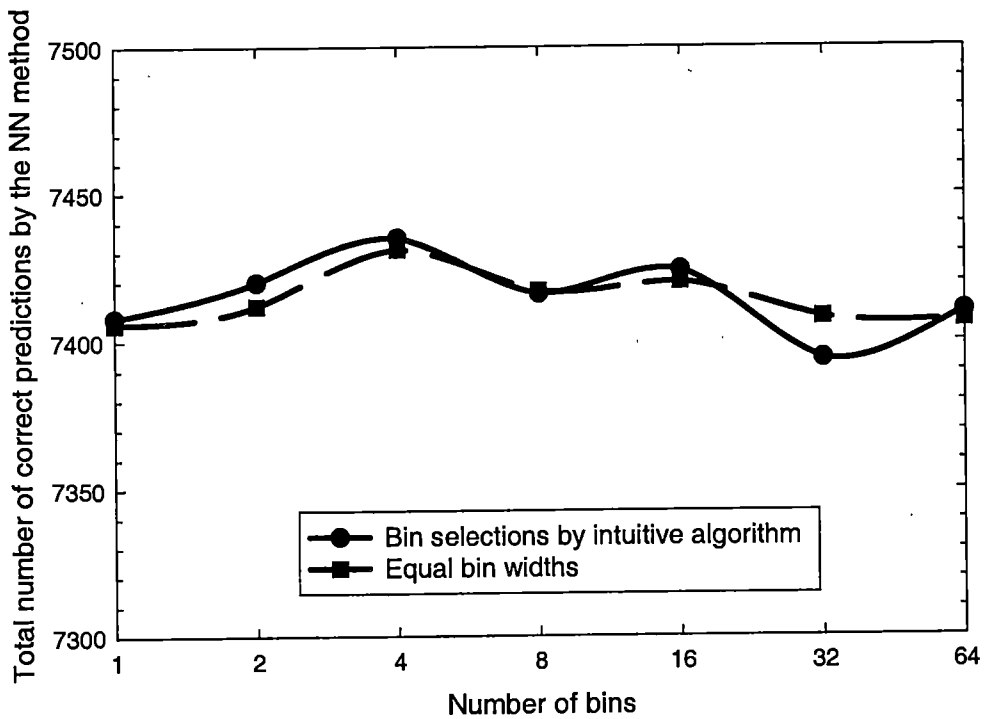


Figure 5.11: Total number of correct predictions by the NN method versus the number of bins

Table 5.15: Results for bin selection determined by the intuitive algorithm for various number of bins: A: Fluorophore A (R6G), B: Fluorophore B (TRITC), C: Background, AB: Both fluorophore.

Number of temporal bins	Categories	A	B	C	AB	Total
	Actual number of cases	3662	4233	131	166	8192
64	Cases correctly predicted by NN	3262	4148	0	0	7410
	Cases correctly predicted by MLE	2966	2294	-	-	5260
32	Cases correctly predicted by NN	3220	4174	0	0	7394
	Cases correctly predicted by MLE	2970	2292	-	-	5262
16	Cases correctly predicted by NN	3254	4170	0	0	7424
	Cases correctly predicted by MLE	2974	2296	-	-	5270
8	Cases correctly predicted by NN	3250	4166	0	0	7416
	Cases correctly predicted by MLE	2955	2308	-	-	5263
4	Cases correctly predicted by NN	3282	4153	0	0	7435
	Cases correctly predicted by MLE	2943	2330	-	-	5273
2	Cases correctly predicted by NN	3267	4153	0	0	7420
	Cases correctly predicted by MLE	2891	2344	-	-	5235
1	Cases correctly predicted by NN	3253	4155	0	0	7408
	Cases correctly predicted by MLE	3662	0	-	-	3662
Learning set:	Actual number of cases	3725	4184	129	154	8192

comparable to the results of those for the equal bin width selections as shown in figure 5.11.

The results above show that, under the current conditions, the total number of bins within the time window needs to be at least 8 if the widths of each bin is chosen to be equal and the MLE method is used to analyze the data. If the intuitive algorithm is applied, the required resolution of bins by the MLE method can be further reduced. In addition, the selections of bin-width within the time window influence more to the

MLE method than the NN analysis. From the point of view of the NN analysis, the bursts belonging to molecule type either A or B are clearly separated by the information of fluorescence brightness (or total number of photons in each burst), other than the timing information. Hence, the selection of the number of bins are not as important to the NN analysis as to the MLE analysis.

In the 1 bin case, only the NN method gives the consistent result with respect to other bin selections. The MLE method predicts all the events as the passage of molecule type B. This is another evidence to show that the classification of bursts by the NN analysis depends almost entirely on the brightness of the molecules in this case.

In summary, the NN analysis is superior than the MLE method in that it is able to grasp the information difficult to be formulated and to be implemented into the statistics, such as the information of the brightness of bursts. The selection of small number of bins, such as 4 bins, is sufficient to achieve the required accuracy, and will save time in the data collection stage.

5.6 Experiment Improvements

5.6.1 Motivation for Fast SMD

SMD is believed to be an enabling technology for searching for rare molecules due to its sensitivity. To accomplish the task, it is necessary to process a very large number of molecules and spectroscopically distinguish the rare molecules from the majority species. Rapid detection of molecules is necessary if large numbers of fluorescent-labeled

molecules are to be processed within a reasonable time. In addition, rapid detection requires the transit time of molecules to be minimized when the molecules pass through the probe region, and the capability to distinguish the simultaneous passage of two or more molecules. For example, for the assay of 1 micro-liter of solution at a concentration of 1 pico-molar, 6×10^5 molecules must be counted. If the time required for each detection is 100 μs, and if molecules were processed at a mean rate of 10^3 s^{-1}, then the assay would take about 10 minutes, and the Poisson probability that a photon burst is due to two or more molecules passing simultaneously through the detection zone would be ~5%.

5.6.2 Details of experimental improvements

In our lab, Li *et al.* had constructed an experimental system enabling the single molecule detection with an efficiency of ~80% and with a transit time of $2\delta t$ =1.0 ms. Improvements in the speed of SMD are considered in this work. Detection of BODIBY-TR dye molecule with a transit time of $2\delta t$ =176 μs has been achieved in the recent measurements. Improvements of the SMD apparatus will be introduced in this section, and the experimental results are shown in section 5.6.3.

The experimental setup is shown in figure 5.12. The laser used in the experiment is a synchronously-pumped dye laser (Coherent 702-1), which provides 8 ps pulses at 76 MHz and 585 nm. The horizontally polarized beam is expanded and collimated with a pair of doublet lenses and focused with another doublet to a beam waist of 3.5 μm, as measured in air by scanning a 1.0 μm pinhole through the waist. When focused

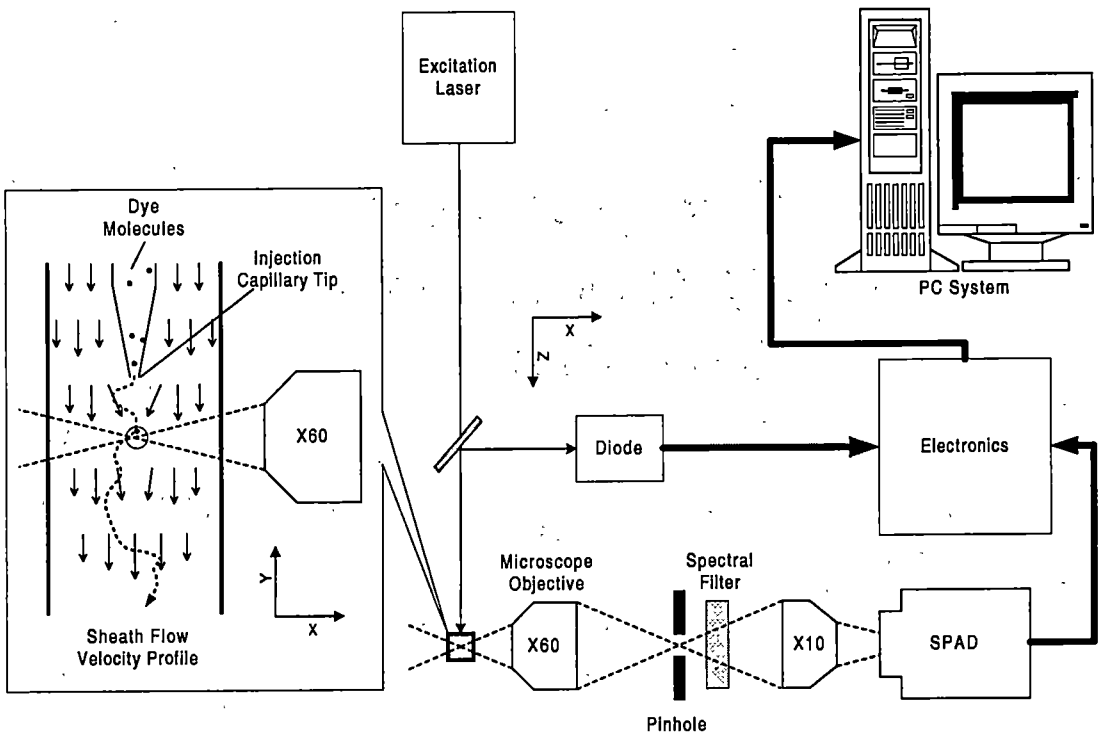


Figure 5.12: Apparatus for efficient single-molecule detection

into water, the waist would theoretically be smaller by a factor of 1.33 due to the refractive index of water, if it were not for aberrations introduced at the glass walls of the square-bore flow cell. The fluorescence is collected using a $\times 60$ microscope objective (Nikon CF Plan Achromat, model 79173), which has a numerical aperture of NA= 0.85, and an adjustable cover-glass aberration correction. When micro-molar concentration dye is introduced through the injection capillary, the image appears as a slightly fuzzy circular disk and hence there appears to be no advantage to the use of crossed slits for the spatial filter. Instead, a round pinhole of radius $r_0 = 250$ μm is used as a spatial filter. This just accommodates the image. Within the approximation of geometric optics and no aberrations, the pinhole defines a double-conic shaped region of maximum light collection efficiency $\hat{C} = (1 - \cos\alpha)/2 = 0.12$, with radius $s_0 = r_0/60 = 4.2$ μm, conic half-angle $\alpha = \sin^{-1}(\mathrm{NA}/1.33) = 40°$ and half-depth $s_0 \tan\alpha = 5.0\mu$m (see section 3C of reference [11]). The intersection of the flow stream and laser beam would be completely contained within the double cone.

Following the pinhole, the collected light passes through a Raman notch filter (Kaiser), which rejects Rayleigh scattered laser light at 585 nm with an optical density of 6.0 while transmitting fluorescence with an efficiency of 0.9, and then a band-pass interference filter (Omega Optical). The interference filter pass band was selected to reject the major Raman scatter component from the solvent, which is the 0-H stretching mode at 3000 cm^{-1} and which falls at 709 nm for an excitation wavelength of 585 nm. The light passed by the pinhole and filters is then imaged by a Newport M-10 microscope

objective to a disk of diameter $\sim 50\mu$m at the center of a single photon avalanche diode (SPAD).

In preliminary experiments to determine the complete optical throughput, a highly attenuated portion of the dye laser beam was focused into the flow cell so as to directly pass into the collection objective, through the pinhole and towards the SPAD. When the dye laser was tuned to 600 nm, the weak broad-band fluorescence within the dye laser beam was identified as the major source of background. Hence to reduce the background count rate, the laser beam is passed through a narrow-band interference filter, placed before the focusing lens.

In Li's experiment [10], single molecules of S-101 in aqueous solution are efficiently detected as they pass from a 0.8 μm microcapillary injection tip into a sheath flow, which carries them through an elliptical cylindrical excitation volume of 1.9 pl with a mean transit time of 1.0 ms. Analytical calculations and Monte Carlo simulations [22] of the experiment indicated that the 0.8 μs dead time of the avalanche photodiode was the most significant factor limiting the time required for the detection of each molecule. Although photo-multiplier tubes exhibit considerably less dead time, use of a high quantum efficiency (> 0.6) avalanche photodiode (EG&G Canada SPCM-200) minimizes the number of molecular excitations required for detection and hence the photodegradation probability.

A new active-quenching circuit has been developed and custom modified for sub-nanosecond time-gated detection [24]. Best timing performance and background after-

pulse rate are obtained with a dead time of 65 ns, which enables instantaneous photon count rates of up to 1.9×10^6 s^{-1} with only 10% loss due to pile-up. The previous experimental results were obtained at excitation intensities well below saturation, because of the high rate of promptly scattered photons from solvent molecules ($\sim 3 \times 10^5$ s^{-1}), which were removed by the time-gating circuit but nevertheless gave significant pile-up at the passively-quenched avalanche photodiode. The reduced dead time of the new actively-quenched detector enables higher laser intensities to be used, which in turn enables single molecules to be detected with shorter transit times and faster solution flow rates.

The Monte Carlo simulation indicates that for faster sheath flow rates less diffusional spreading occurs over the same distance downstream from the microcapillary injector and hence efficient detection can still be accomplished with an excitation volume of smaller cross-section. Diffusion of the sample in the sheath flow is reduced and the volumetric rate at which the sample is processed is increased. For our present experiments, we have accordingly reduced the excitation volume to $\sim$ 0.25 pl while increasing the laser excitation intensity by a factor of ~ 5.

To accomplish single molecule detection within faster transit times and higher instantaneous fluorescence photon rates, the components of the data acquisition system have been replaced. Previously a time-to-amplitude converter (TAC) with 2.5 μs dead time and an analog-to-digital converter with 1 μs conversion time were used to provide software controllable time-gating, while most of the Raman background was removed

by a temporal pre-filter (a sub-nanosecond anti-coincidence circuit) with a dead time of $\sim$ 10 ns. During the passage of a single molecule, the instantaneous rate of the counts that pass the pre-filter was $\sim$ 20 photons in 1.0 ms, and hence the 2.5 μs TAC dead time gave $<$ 5% loss. For detection of molecules within shorter transit times, significantly higher instantaneous fluorescence rates occur. Therefore, the output signal from the temporal pre-filter is connected directly to a 100 MHz multi-channel scalar (Ortec ACE-MCS) and time-gating is accomplished solely by the hardware.

5.6.3 Fast SMD results and discussions

Figure 5.13 shows an example of photon bursts of BODIPY-TR(D-6116) dye obtained in a bulk flow cell with the capillary injector, under the following conditions: Laser power P = 23 mW; circular beam waist ω_0 = 3.5 μm; spatial filter object space diameter = 8.4 μm. The molecular transit time obtained from the auto-correlation function, shown in figure 5.14, is $2\delta_t$ = 176 μs. The dwell time of the multi-channel scalar is 50 μs and the data is processed by a simple sliding sum filter with a bin time of 200 μs to yield the graph in figure 5.13. The background bursts in the figure is obtained by shutting off the flow in the capillary, and hence no dye molecule will diffuse into the probe region.

For the experimental parameters used, the peak excitation probability per laser pulse, which occurs when a molecule is in the center of the sample volume, is P_{excite} = 0.44. This indicates a moderate level of saturation of the absorption of the molecule. The mean number of excitations per molecule is $\sim$ 4500, comparable to that in reference

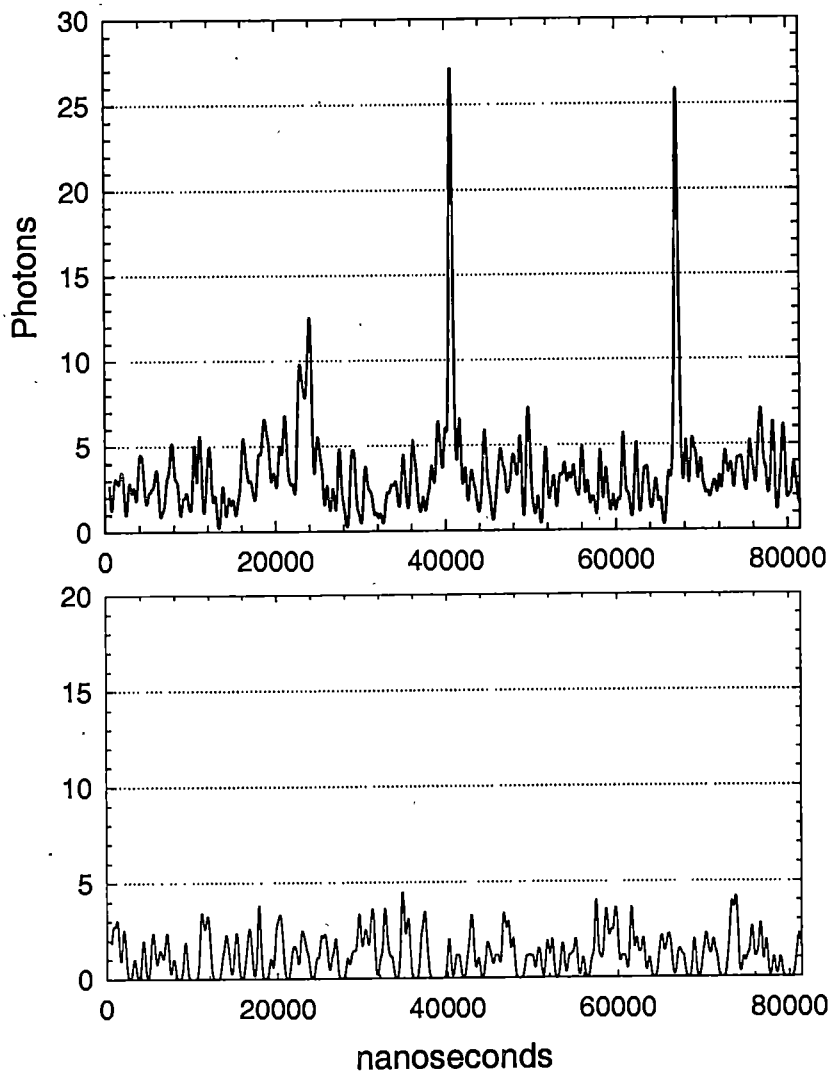


Figure 5.13: Photon bursts from single molecules of BODIPY-TR(D-6116) and background from pure water

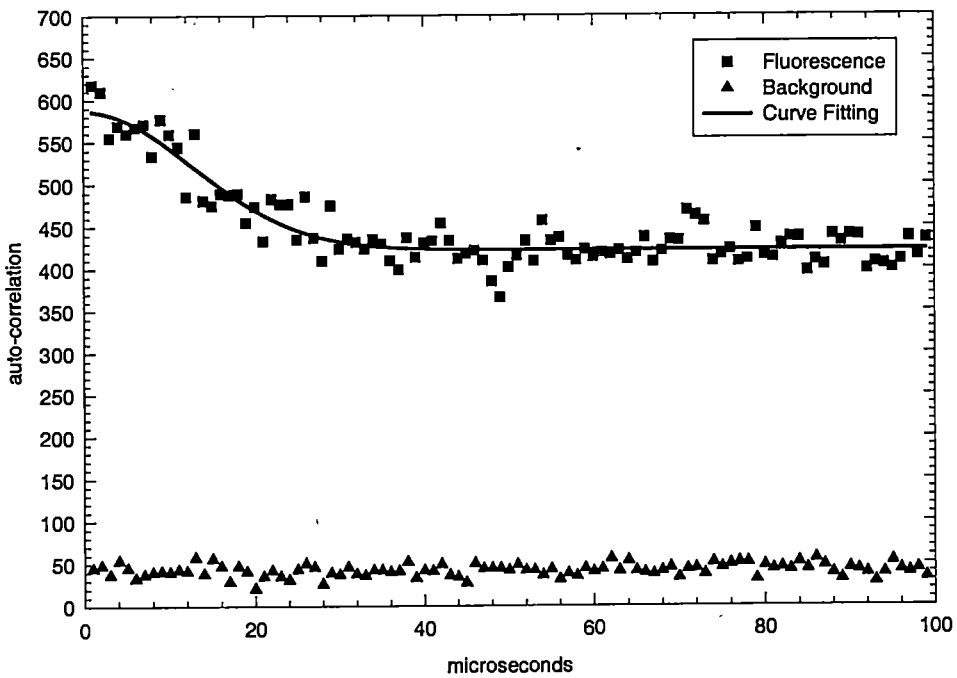


Figure 5.14: Auto-correlation function

[10], and hence the burst amplitudes are also of comparable size as shown in figure 5.15.

5.7 Fast data collection scheme

Section 2.1 explained that the nanosecond time interval between the laser excitation pulse and the fluorescence photon can be used to sort each photon into an appropriate bin for use for identification of molecules. The usual experimental hardware for acheiving this involves the use of a time-to-amplitude converter (TAC) and pulse height analyzer, or TAC and analog to digital converter [10], or time to digital converter [28]. The time interval is typically measured with a precision of 8 to 14 bits, i.e., with 256 to 16,184 channels of resolution. Adjacent channels may be arbitrarily grouped in the software to produce a smaller number of bins. Such hardware invariably introduces a dead time of the order of a millisecond and is not suitable for experiments for fast single molecule identification. This section discusses alternate experimental hardware for directly measuring and sorting the time interval into a small number (2 to 4) of bins. The hardware utilizes fast coincidence or time-gate circuits, with widths that may be varied by adjusting the threshold of the discriminators and the lengths of co-axial cables. Thus, bins with desired widths and temporal positions can be obtained. However, the bins need not be contiguous, and also they need not be non-overlapping.

As implemented in the fast SMD experiment, a multi-channel scalar (MCS)(ACE-MCS) was used to collect the data. The main disadvantage of this scheme is that the

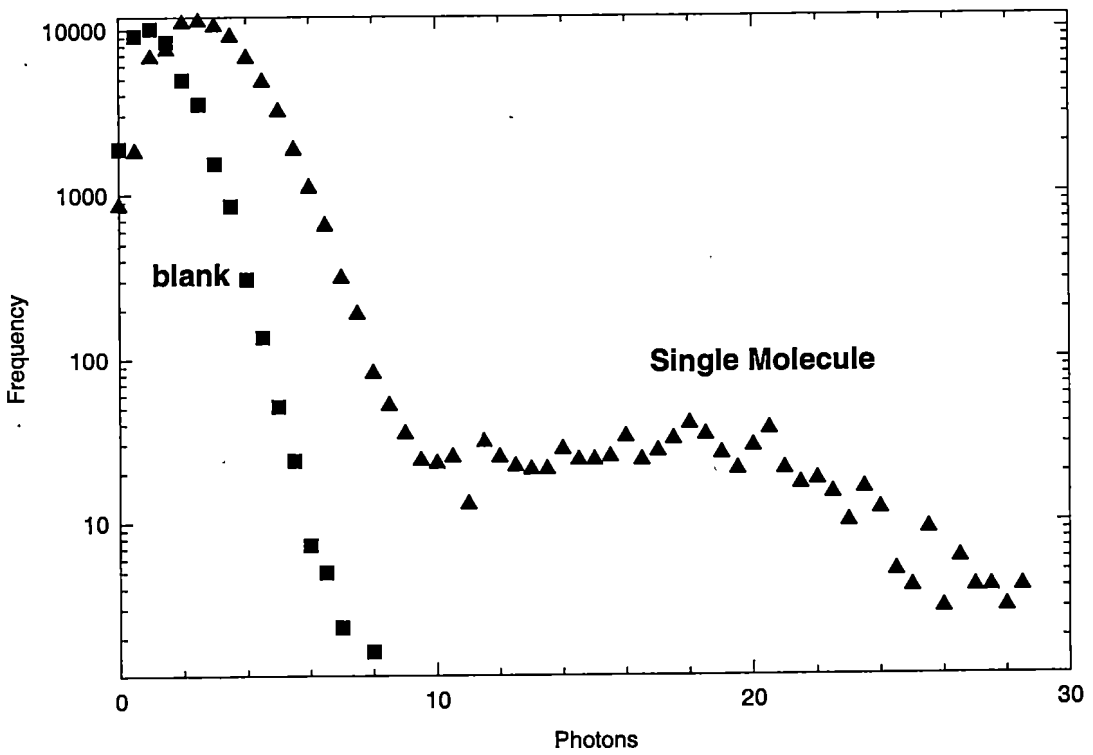


Figure 5.15: Distribution of photons

data acquisition is not continuous, and the analysis of the data is not performed at the exactly same time as it is acquired. The display of the photon bursts is delayed. For each run of data collection, the acquisition is inactive for a period of time while the Fortran program reads in the 4096 point data file generated by the MCS and processes it with the simple sliding sum algorithm. Therefore, this inevitable loss of data, or interrupted data collection must be accounted for in the algorithm for obtaining the auto-correlation function. Moreover, in order to implement the single molecule identification experiment, multiple bins within the time window must be implemented. The ACE-MCS can only collect data from one bin at a time, and a MCS with multiple inputs is required.

To overcome the above limitations, a new data collection scheme was constructed using a National Instruments PC-TIO data acquisition card. The PC-TIO card is a timing and digital I/O interface for the ISA bus of IBM compatible PCs, with up to 10 10 MHz counters or clocks, which may be read under interrupt control. For the reason that two counters of the PC-TIO card are used to alternatively count TTL at low or high voltage output for each channel and one counter is required to provide the clock signal, only 3 channels is able to be constructed on one PC-TIO card. It is possible to read the counters and continuously stream data to memory for real time analysis in such a way as to emulate a 3-channel MCS. In this work, a C program, which was written by L. Davis for emulating a 1-channel MCS with the PC-TIO card and which was used in reference [27], was modified and extended to give a 3-channel MCS.

As the output signal from the prefilter circuits are negative and relatively small($\sim 1V$),

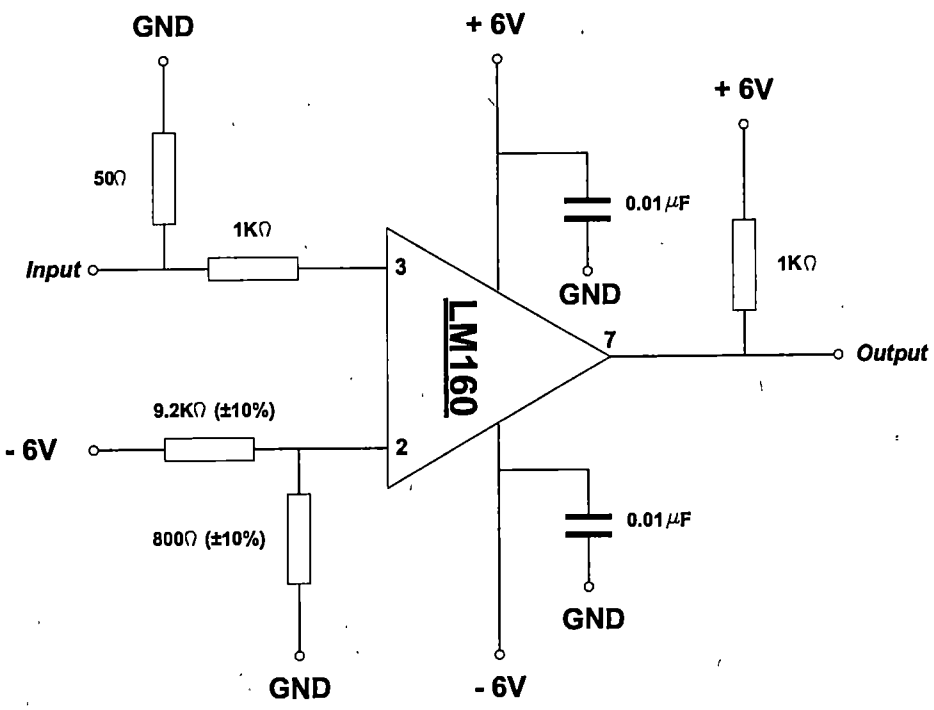


Figure 5.16: Comparing amplifier

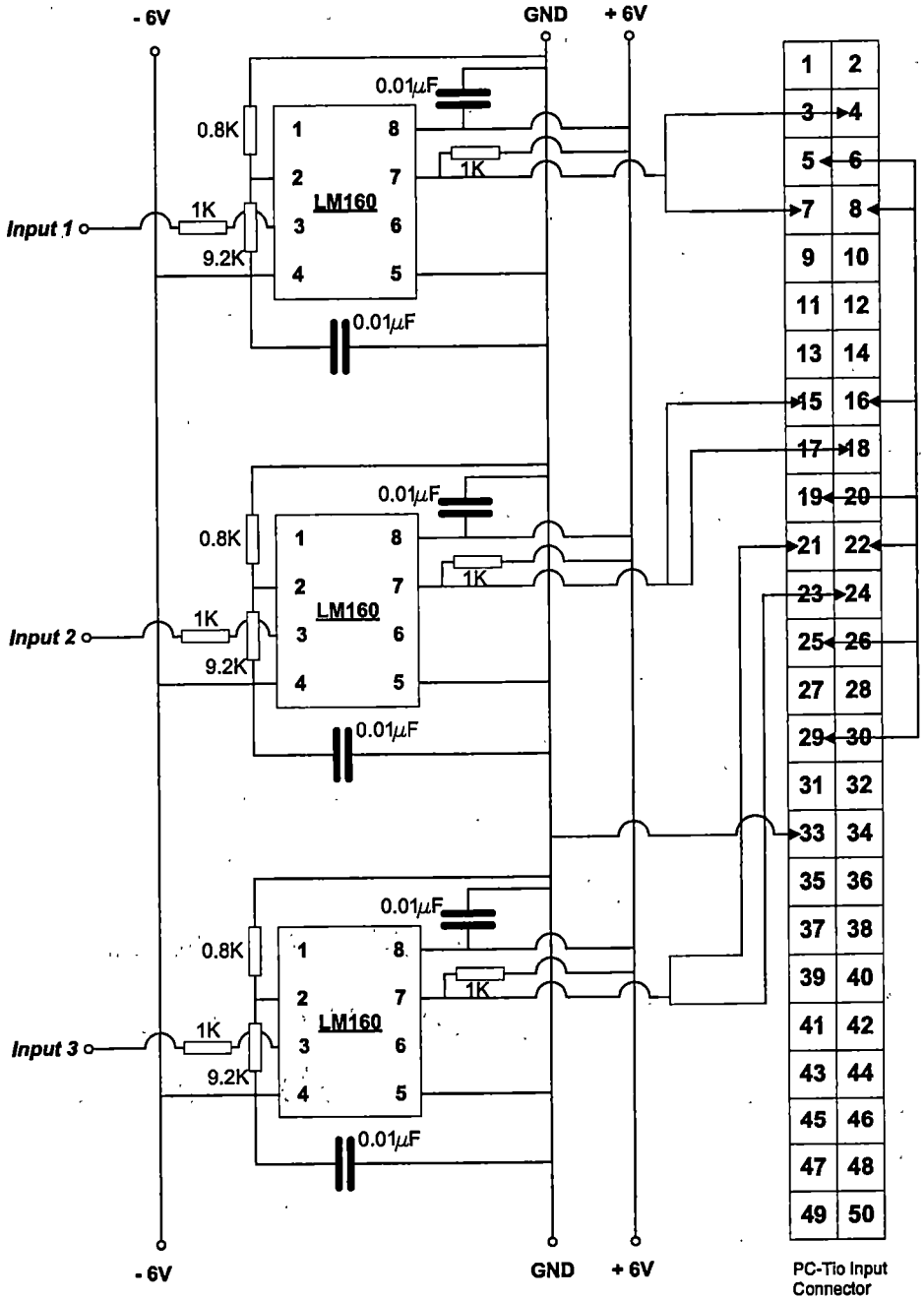


Figure 5.17: Pulse conditioning circuit and connection to PC-Tio board

yet the PC-TIO card needs a driven signal larger than $2.4V$, a fast invert amplifier circuit is required to be used as the interface between the prefilter circuit and the data acquisition card. A comparator circuit is chosen for such purpose, and the diagram of such a circuit is illustrated in figure 5.16. The +/− 6V power is supplied by the Nuclear Instrument Module (NIM).

Figure 5.17 illustrates the block diagram of the connection between the PC-TIO card and 3 comparator circuits. Counter 1 of the PC-TIO card is used to provide a 10MHz clock signal to all the other gates. Counters 2 and 3 are used by channel 1, counters 6 and 7 are used by channel 2, and counters 8 and 9 are used by channel 3. Among all the counters, 2, 6, and 8 are configured to collect data during the high voltage clock cycle, and 3, 7, and 9 are configured to collect data during the low voltage cycle. The ground of all 3 channels is connected to pin 33 of the PC-TIOcard.

5.8 Simulation with calibration runs

As described in chapter 4, the neural networks require learning prior to the testing, and hence a calibration run needs to be preformed to train the neural networks before any real time detection in the SMD experiment. Therefore, a corresponding procedure is proposed to implement the NN analysis for the experiment, and in such a experiment,

- the experiment conditions need to be kept the same for all the runs below,
- run 1: only molecule A goes through the probe region, the corresponding bursts are collected, and the desired output for the NN training is set to be molecule

type A,

- run 2: only molecule B goes through the probe region, the corresponding bursts are collected, and the desired output for the NN training is set to be molecule type B,
- run 3: no molecule goes through the probe region and the desired outputs of the corresponding collected bursts are set to be type background,
- the data of the above runs are combined, and are used as the training data for the NN analysis(note that the NN software can be configured to read in the training data in a random order),
- the trained neural networks are outputted as a C subroutine, and the subroutine can be implemented into the data analysis of the experiment,
- run 4: finally the real time detection can be performed by using the trained network.

Clearly, the case for the passage of both types of molecules is beyond the ability of this procedure. Therefore, the concentration of the solution needs to be sufficiently diluted and the flow rate needs to be kept sufficietly fast to avoid such events from happening, and minimize the error where both types of molecule are present in the probe region.

According to the procedure described above, a SMD experiment with calibration runs is simulated with the improved version of the SMD simulation. Most of the pa-

Table 5.16: Parameters in the simulation with calibration

Parameters Used in each run	Run 1 Bursts from A&Prompt	Run 2 Bursts from B&Prompt	Run 3 Bursts from Only Prompt
Fraction of type A among the total	1.0	0.0	0.333
Concentration of solution (mol/picoliter)	1.0×10^{-13}	2.0×10^{-13}	0.0×10^{-13}

rameters are the same as those in section 5.4.1, and other changed parameters are listed in table 5.16.

In the simulation, data for the run 1 and the run 2 are collected by using different solutions containing molecule A and B respectively, this is acheived by varying the value of fraction of As among all the molecules. In third run, only pure water goes through the probe region, and data for the background is collected. All the above 3 runs are subjected to a fixed amount of time. A value of 2×10^9 total time steps are used in the simulation, and it corresponds to 1220 seconds worth of data collection in the calibrations. As the result, there are totally 4513 bursts from molecule A, 5240 bursts from B, and 163 from the background. These data are combined and used as the training data of the NN analysis in a 64 bins case. The trained network is then implemented into the real time simulation, and the 8192 bursts generated from the simulation with the solution of mixture of A and B are tested by the network. The results of identification of molecules by the NN and the MLE methods are shown in tables 5.17 and 5.18, respectively.

Table 5.17: Identification of molecule by the NN method, simulating a real time experiment

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3650	3254(396)	89.2
molecule B	4240	4124(116)	97.3
Prompt	145	0(145)	0
both A and B	157	0(157)	0
Total	8192	7378(814)	90.1

Table 5.18: Identification of molecule by the MLE method, simulating a real time experiment

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3650	2912(738)	79.8
molecule B	4240	2394(1846)	56.5
Prompt	145	0(145)	0
both A and B	157	0(157)	0
Total	8192	5306(2886)	64.8

The results from the NN and MLE methods are comparable with the results in sections 5.4 and 5.5. The NN method again outperforms the MLE method, and gives above 90% correct predictions of individual molecules. In comparison, the MLE method gives only 64.8% correct predictions, and more than 1/3 of the events are predicted wrong. Therefore, it is an applicable approach to implement the NN analysis into a real time single molecule identification experiment.

5.9 Simulation with overlapped bins

With the hardware discussed in section 5.6, the sorting bins within the time window are selected by varying the thresholds of discriminators and the lengths of co-axial cables. The widths and positions of each bin may be monitored on an oscilloscope during experimental setup. The bins may be selected in such a way that they overlap, and hence a photon may contribute a count to two or more bins. According to the assumption made in the MLE method, introduced in section 3.2, the spectroscopic bins are not allowed to overlap each other, and each bin should be statistically independent of all others. Therefore, with overlapping bins, the MLE method based upon the multinomial distribution would no longer be valid. It would be too complicated to implement the MLE analysis to consider the case of overlapped bins, and hence only the NN analysis is used in this section.

Note that the overlapping of bins may provide extra sorting space for photons, and hence further reduce the required number of bins within the time window. This could also speed up the data analysis. For example, using 2 overlapped bins rather than 3 non-overlapping contiguous bins reduces the data storage and analysis time requirements. In addition, in the 2-bin example above, the photon count in the overlapped bin can not be obtained as the arrival time of each photon with respect to the laser pulse is no longer available in the experiment. The photon counts in bin 1 and 2 will not be sufficient to give the number of photons falling within the overlapped bin.

As before, in the NN analysis, a 2-layer back-propagation neural network with δ

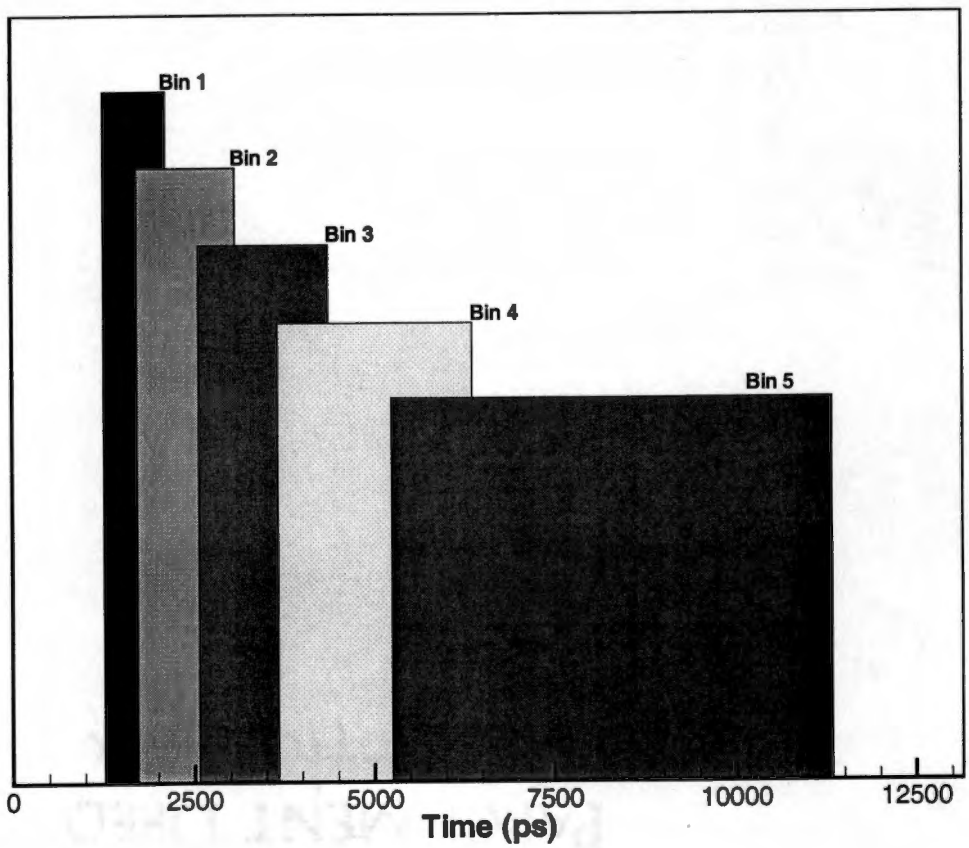


Figure 5.18: The 5 overlapping bins within the time window

Table 5.19: Identification of molecule by the NN for 5 overlapping bins

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3662	3264(398)	89.1
molecule B	4233	4158(75)	98.2
Prompt	131	0(133)	0
both A and B	166	0(166)	0
Total	8192	7422(770)	90.6

learning rule and sigmoid activation function is used for data analysis. The 7 inputs of the network include the photon counts in each bin, the peak amplitude and the duration of each burst, and the hidden layer contains 14 nodes. The bins and the corresponding widths are arbitrarily chosen to be as shown in figure 5.18. To simulate a 5 overlapped bins case, the same set of 16384 burst data as in section 5.4 are used. The first 8192 data are used to train the network, and the remaining 8192 sets of data are used as the testing data. The result is shown in table 5.19.

To simulate situation where overlapped bins exist, a similar approach with calibration runs as in section 5.8 is carried out. The bursts data for training and testing the neural network in section 5.8 are used for the overlapped bins case, respectively. The result is shown in table 5.20.

The results show that the NN analysis still gives a good result under the condition where the bursts data from each bin are dependent to each other. This is because the classification of molecules by the NN analysis is based on not only the timing information but also the information of fluorescence brightness of dye molecules.

Table 5.20: Identification of molecules by the NN method in the case of a selection of 5 overlapping bins by the procedure introduced in section 5.8

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3650	3237(413)	88.7
molecule B	4240	4155(85)	98.0
Prompt	145	0(145)	0
both A and B	157	0(157)	0
Total	8192	7392(800)	90.2

Table 5.21: Identification of molecules by the NN analysis for the case with 3 overlapping bins within the time window by the procedure introduced in section 5.8

Type of burst	Actual Occasion	Correct prediction (incorrect prediction)	% of the correct predictions
molecule A	3650	3248(402)	89.0
molecule B	4240	4145(95)	97.8
Prompt	145	0(145)	0
both A and B	157	0(157)	0
Total	8192	7393(799)	90.2

A case with 3 overlapped bins within the time window is also simulated as there are only 3 bins available in the currently developed experiment in our lab. The widths of the 3 bins are chosen arbitrarily, and are illustrated in figure 5.19. The same set of parameters in section 5.8 are used here, and the result is shown in table 5.21.

5.10 Time consumption by the NN analysis

This section will investigate the total time used for performing the NN analysis. In general, the time is composed of that taken during the calibration experiments, i.e., the

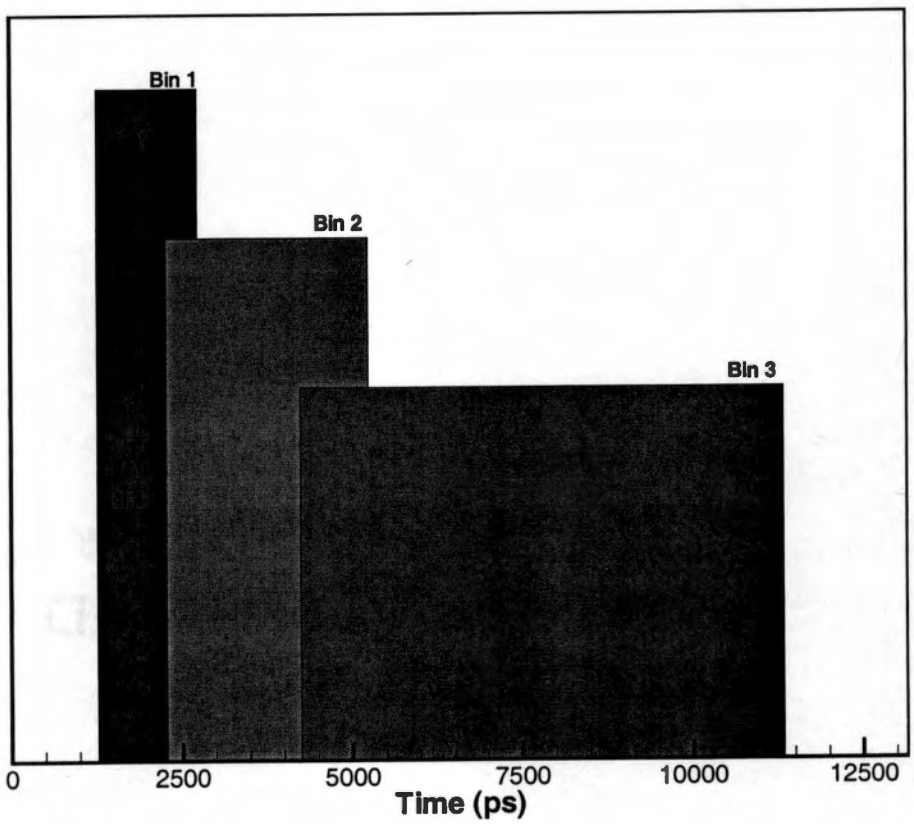


Figure 5.19: 3 overlapping bins within the time window

time spent on training neural networks, and that taken during the real time testing. According to the procedure proposed in section 5.8, the total time that the calibration experiments take is typically 3 times of the training time used in a single run. In this section, onlt a single run of the NN training is simulated, and the time spent in such a run is recorded.

The hardware setup and other conditions for training the neural network are listed as below:

- Hardware: Pentium II 300MHz, 128MB memory, windows 95 operating system,
- Software: NeuralWare Professional II/Plus,
- Achetecheture of the networks: 2-layer back-propagation networks, δ learning rule, sigmoid activation function, number of nodes in inputs, outputs, and hidden layers are defined in table 5.11,
- Total learning times: 100000.

The results for selections of different number of input and hidden nodes are shown in table 5.22.

Table 5.22 shows that if a small number of bins is chosen within the time window, e.g., less than 16 bins, the training process takes no more than 1 minute, and the asymptotic limit of root mean square error of the network can be reached with the current total number of trainings, i.e., 100000.

The time taken by the NN analysis during the real time testing is comparable to

Table 5.22: Training time for the networks define in table 5.11

Number of bin in time window	Number of input units	Number of units in the hidden layer	Number of output units	Total Units	Training Time (s)
64	66	132	4	202	363
32	34	68	4	106	175
16	18	36	4	58	93
8	10	20	4	34	70
4	6	12	4	22	62
2	4	8	4	16	61
1	3	6	4	13	60

that of the MLE method according to the investigation on the SPARC workstations for both methods. The costs of predictions for both the methods are around a few clock cycles per identification, and the NN analysis normally takes one or two more cycles than the MLE method.

Chapter 6

Summary and Future Work

This work is the first that incorporates the NN method in the research area of the data analyses of SMD and SMI experiments and simulations. Results show that the NN analysis for categorization can be superior to the traditional MLE method, particularly in cases where fluorophores that are subject to be classified have distinctive photophysical properties. In addition, a procedure of how to implement the NN analysis in real-time for experiments is proposed. The results of the simulation with the NN procedure are promising.

In this dissertation, the prior version of the simulation code for single molecule detection and identification experiments has been extended to include the fluorescence probability as the convolution of the exponential decay and instrument impulse function. The selection of the time window has been implemented by monitoring the variation of signal and noise as anticipated to occur in experiments. The so-called intuitive algorithm

for selecting bin widths within the time window has also been extended accordingly.

Several approaches to optimize the SMI have been evaluated theoretically, for example, errors of predictions of fluorescence bursts have been investigated, (1) by adjusting the start and end points of the time window, (2) by changing the number of sorting bins within the time window, and (3) by selecting the bin widths. Results show that several bin-width selections are favorable, for example, equal bin widths or bin widths chosen by the intuitive algorithm. The investigation clarifies that the choice of a set of optimal bin widths is largely determined by the specific experimental conditions.

A number of cases have been simulated by using the improved version of the simulation program. Results confirm that at most 8 bins are required to sort photons within the time window given all bin widths are equal; more bins, e.g., 16, 32 and 64 bins yield hardly any improved identification. Results also show that the number of bins within the time window can be further reduced if all the widths were determined by the intuitive algorithm. The simulations with bin widths determined by the intuitive algorithm generally yield improved results compared to those with equal widths when the total number of bins within the time window is less than 8.

Simulation results provide a useful direction for the SMD and SMI experiment. Fluorophores need to have as distinctive photo-physical properties as possible, and such properties may be chosen, for example, the different fluorescence lifetimes or different fluorescence quantum efficiencies of dyes. Furthermore, properly setting the width of the time window will reduce the portion of background photons in each burst, and

substantially improve the precision of identification of molecules. In the single molecule identification experiment, the concentration of the solution containing dye molecules needs to be further diluted to avoid the cases where multiple molecules are present at the probe region simultaneously. For diluted samples, a reduction of the error is expected for both the MLE and the NN methods.

The future work may include finding suitable dyes that have large signal strength at 585 nm excitation wavelength, otherwise, other excitation wavelengths need to be explored, and corresponding spectral filters are required. To further suppress the background photon counts, an experimental setup with a probe volume in the order of femtoliters needs to be pursued, including the use of confocal epi-illumination. In addition, the new data collection scheme using the PC-TIO card and the pre-amplifier circuitry needs to be evaluated together prior to experiments. Currently, 3 bins are configured for the PC-TIO card and the circuitry, but only one bin of data collection is available from the signals of the SPAD and the PD. Therefore, extra hardware is required for implementation for the multiple bin data acquisition. Incorporation of the NN method in the data acquisition of SMD experiments can be done according to the procedure proposed in this work. Finally, with the suggested preparation work, the real-time single molecule identification experiment is expected to show the simulated, predicted results.

Bibliography

Bibliography

[1] L. M. Davis, J.W.K. Williams, and D.T. Lamb, "Computer simulation of gene detection without PCR by single molecule detection," BiOS Europe '98 Conference, Stockholm 8–12 September 1998, pp. 282–293, SPIE Proceedings Vol. 3570, 'Biomedical Sensors, Fibers, and Optical Delivery Systems,' N.I. Croitoru, M. Miyagi, F. Baldini,I. Lundstroem, O.S. Wolfbeis, M. Frenz, and R. Pratesi, Editors, Jan. 1999.

[2] U. Lieberwirth, J. Arden-Jacob, K. H. Drexhage, D. P. Herten, R. Muller, M. Neumann, A. Schulz, S. Siebert, G. Sagner, S. Klingel, M. Sauer, and J. Wolfrum, "Multiplex dye DNA sequencing in capillary Gel electrophoresis by diode laser-based time-resolved fluorescence detection", *Anal. Chem.* **70**, pp. 4771-4779, 1998.

[3] J.H. Werner, H. Cai, P.M. Goodwin, and R.A. Keller, "Current status of DNA sequencing by single molecule detection," BiOS '99 Conference, San Jose, January 23–29, 1999, paper 40 in SPIE Proceedings Vol. 3602, 'Advances in Fluorescence Sensing Technology IV,' J.R. Lakowicz, S.A. Soper, and R.B. Thompson, Editors,

Apr. 1999.

[4] K. Dorre, S. Brakmann, M. Brinkmeier, K. Han, K. Riebeseel, P. Schwille, J. Stephan, T. Wetzel, M. Lapczyna, M. Stuke, R. Bader, M. Hinz, H. Seliger, J. Holm, M. Eigen, and R. Rigler, "Techniques for single molecule sequencing", *Bioimaging* **5**, pp. 139–152, 1997.

[5] S. A. Soper, L. M. Davis, and E. B. Shera, "Detection and identification of single molecules in solution", *J. Opt. Soc. Am. B* **9**, pp. 1761–1769, 1992.

[6] C. Zander, M. Sauer, K. H. Drexage, D. S. Ko, A. Schultz, J. Wolfrum, L. Brand, C. Eggling, and C. A. M. Seidel, "Detection and characterization of single molecules in aqueous solution", *Appl. Phys. B* **63**, pp. 517–523, 1996.

[7] M. Sauer, J. Arden-Jacob, K. H. Drexhage, F. Gobel, U. Lieberwirth, K. Muhlegger, R. Muller, J. Wolfrum, and C. Zander, "Time-resolved identification of individual mononucleotide molecules in aqueous solution with pulsed semiconductor lasers", *Bioimaging* **6**, pp. 14–22, 1998.

[8] M. Sauer, J. Wolfrum, "Single molecule detection in biology with multiplex dyes and pulsed semiconductor lasers", *Applied Fluorescence in Chemistry, Biology and Medicine*, W. Rerrig, B. Strehmel, S. Schrader, and H. Seifert Eds., **Springer-Verlag**, Chap.2, 1999; and references therein.

[9] J. Enderlein, P. M. Goodwin, A. Van Orden, W. P. Ambrose, R. Erdmann, and R. A. Keller, "A maximum likelihood estimator to distinguish single molecules by their fluorescence decays", *Chem. Phys. Lett.* **270**, pp. 464–470, 1997.

[10] L.Q. Li and L. M. Davis, "Rapid efficient detection of single chromophore molecules in aqueous solution", *Appl. Opt.* **32**, pp. 806–820, 1993.

[11] D. H. Bunfield, and L. M. Davis, "Monte Carlo simulation of a single-molecule detection experiment," *Appl. Opt.* **37**, pp. 2315–2326, 1998.

[12] D. H. Bunfield, "Simulation of a single molecule detection experiment", M.S. thesis (University of Tennessee, Knoxville, Tennessee, 1997)

[13] M. Kollner, "How to find the sensitivity limit for DNA sequencing based on laser-induced fluorescence", *Appl. Opt.* **32**, pp. 806–820, 1993.

[14] L.M. Davis, Y. Sun, and B. Whitehead, in Advances in Fluorescence Sensing Technology IV, Proc. SPIE **3602B**, 379–381 (1999)

[15] L. M. Davis, and D. H. Bunfield, "Spectroscopic identification of individually detected fluorescent molecules", *The Fifth International Conference on Methods and Applications of Fluorescence Spectroscopy, Berlin, 1997. Book of Abstracts*, W. Rettig, ed. (Springer-Verlag, Berlin, 1997), p.27

[16] M. Kollner and J. Wolfrum, "How many photons are necessary for fluorescence-lifetime measurements?", *Chem. Phys. Lett.* **200**, pp. 199–204, 1992.

[17] S. A. Soper, E. B. Shera, L. M. Davis, H. L. Nutter, and R. A. Keller, "The photophysical constants of several fluorescent dyes pertaining to ultrasensitive fluorescence spectroscopy", *Photochem. and Photobiol.* **57**, pp. 972–977, 1993.

[18] B. Krose, and P. Van Der Smagt, "An Introduction to Neural networks", 8th ed., (University of Amsterdam, Netherland, 1996)

[19] Daniel T. Gillespiem "A theorem for physicist in the theory of random variables", *Am. J. Phys.* **50**, pp. 520-533, 1983.

[20] M. Eigen and R. Rigler, "Sorting single molecules: Application to diagnostics and evolutionary biotechnology," Proc. Nati. Acad. Sci. USA **91** , 5740 (1994).

[21] Y.H. Lee, R.G. Maus, B.W. Smith, and J.D. Winefordner, "Laser-induced fluorescence detection of a single molecule in a capillary," Anal. Chem. **66**, 4142 (1994).

[22] L.M. Davis and L.Q. Li, "Monte Carlo model of a single molecule counting experiment,"in *Laser Applications to Chemical Analysis*, 1994 Technical Digest Series, Vol.5 (Optical Society of America, Washington, DC, 1994) pp.206-209.

[23] S. Baker and R. D. Cousins, *Nuclear Instrumental Methods* **221**, 437 (1984)

[24] A. Spinelli, L.M. Davis, and H. Dautet, "Actively-quenched single photon avalanche diode for high repetition rate time-gated photon counting," Rev. Sci. Instrum.,**67**, 55 (1996)

[25] K. Ghiggino, M. Harris, and P. Spizzirri, "Fluorescence lifetime measurements using a novel fiber-optic laser scanning confocal microscope", Rev. Sci. Instrum.,**63**, 2999 (1992)

[26] D. Magde, and M. W. Windsor, "Picosecond flash photolysis and spectroscopy:3,3'-Diethyloxadicarbocyanine iodide(DODCI)," Chem. Phys. Lett., **27**, 31 (1974)

[27] J.C. Fister, L.M. Davis, S.C. Jacobson, and J.M. Ramsey, "High Sensitivity Detection on Microchips," Digest of Topical Meeting on Laser Applications to Chemical Analysis, Optical Society of America, **3**, 58 (1996)

[28] See webpage at http://www.picoquant.com/timeharp.htm.

Vita

Yuxing Sun earned his B.S. and M.S. degree in physics at Peking University, China, in 1992 and 1995, respectively. From 1995 to 2000, he worked as a research assistant at the Center for Laser Application at The University of Tennessee Space Institute (UTSI), and completed his Ph.D degree in physics in May 2000.