Extension of multidimensional microscopy to ultra-sensitive applications with maximum-likelihood analysis

Lloyd M. Davis*, Guoqing Shen
Center for Laser Applications, University of Tennessee Space Institute, Tullahoma, TN 37388

ABSTRACT

Multidimensional fluorescence microscopy is finding service in forefront biological studies that require separation of images from different fluorophores. For example, commercial microscopes are available with multi-band analog detectors and user-friendly software for "linear unmixing" of species with overlapping emission spectra. To extend such techniques to ultrasensitive and single-molecule applications, we have developed a custom-built microscope, which incorporates two tunable-wavelength picosecond dye lasers for pulse-interleaved laser excitation, angle-tuned reflection of the laser beams from narrow-band Raman notch filters to introduce epi-illumination and provide strong rejection of scattered laser wavelengths, diffraction-limited confocal imaging with 3-dimensional piezo-scanning, an adjustable prism spectrometer for high-throughput resolution of collected fluorescence into 4 spectral bands, and a 4-channel high-quantum efficiency avalanche diode for sub-nanosecond-resolved single-photon detection. Custom software enables multi-band fluorescence correlation spectroscopy and identification of photon bursts for single-molecule detection. For unmixing of spectrally-overlapping signatures for ultrasensitive molecular imaging applications, we find that maximum-likelihood analysis can out-perform least-squares-based linear unmixing in the regime of low photon numbers per spectral/temporal channel. Also, the likelihood surface provides the confidence of the parameter estimates and the covariance of the species contributions. Monte Carlo simulations show that bias in the results of the analysis, which stems from the constraint that photon numbers should be positive, becomes more pronounced at low signal levels, for both maximum-likelihood and least-squares based unmixing.

Keywords: Fluorescence, microscopy, spectral-FLIM, maximum-likelihood, multidimensional microscopy, photon counting

1. INTRODUCTION

Advances in modern fluorescence spectroscopy have enabled it to become a powerful analytical tool for studying molecular structure and dynamics in biological, chemical and medical research. Developments in sample preparation and image processing have been coupled with recent capabilities for acquiring multidimensional data and simultaneously reporting distributions of several different molecules with multiple fluorophore labels, for studies in living cells and tissues [1 - 12]. Hence molecular interactions and processes can be directly monitored and studied in their natural environments.

A major goal of multidimensional fluorescence spectroscopy is to separate the contributions from multiple fluorophores that overlap spectrally and spatially. In practice, it is not possible to directly resolve species with overlapping excitation and emission spectra. For example, there are several intrinsic fluorescent proteins of interest in cellular bioimaging, with excitation and emission spectra that are overlapping so that their fluorescence cannot be separated by optical means alone [3]. Fortunately, species with overlapping spectra can be separated by multidimensional spectroscopy with the use of mathematical algorithms. Commercial multidimensional microscopes, such as the Zeiss LSM 510 META, have been successfully applied in bioimaging applications with multi-labeled biological samples [5, 12].

Multidimensional fluorescence spectroscopy, however, is often required to operate at low signal levels due to particular experimental conditions. In many cases, it is not possible to increase signal levels by simply increasing the excitation power or collection time because of factors such as sample damage, movement, or change, especially for living samples.

*Corresponding author: e-mail: ldavis@utsi.edu; Tel.: (931)393-7335; Fax: (931)393-7218; http://ldavis.utsi.edu/

Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIV,
edited by Jose-Angel Conchello, Carol J. Cogswell, Tony Wilson, Proc. of SPIE Vol. 6443, 64430N, (2007)
1605-7422/07/$18 · doi: 10.1117/12.702217

A similar challenge also exists in multi-dimensional single-molecule detection experiments, in which the fluorescence signal is limited by a maximum signal rate due to saturation or the dead time of the detection electronics, and also by a finite observation window due to a limited diffusional residence time and eventually photobleaching [13].

For measurements with low signal levels, the use of least-squares methods for curve-fitting of experimental data suffers from its limitations. The assumption of equal weights across the measured data set becomes increasingly invalid when there are large differences in signal at different data points. Weights should be used that are inversely proportional to the variance at each data point, but when the measured signal is zero, the variance is also zero, so that the weight at that point becomes infinite. In this case, the weighted least squares (WLS) curve fitting is forced to return a fitting function that is zero. Least squares curve fitting implicitly assumes that the measured data are distributed about the fitting function with a Gaussian distribution and hence it fails to correctly fit shot-noise limited data when counts are low. It has been pointed out that the least-squares method is not appropriate to analyze fluorescence lifetime data at low photon counts, as often occurs in fluorescence lifetime imaging microscopy (FLIM), because the underlying assumption of a Gaussian distribution instead of the actual Poisson distribution in the photon histograms is no longer correct [14, 15]. It was reported that WLS fitting of exponential decay profiles to extract the fluorescence lifetime fails when the total number of photons is less than about 1000, but maximum likelihood estimation (MLE) gives stable results [15]. A similar problem has also been considered in single-molecule localization [16] and fluorescence resonance energy transfer (FRET) experiments [17]. The use of the least-squares criterion results in reduced molecule localization accuracy. In the time-resolved FRET experiments, MLE still provides a statistically significant parameter fluctuation at the 1-σ level when the number of photons is less than about 100. MLE provides a proper approach for problems with low signal levels [18] and can account for Poisson, multinomial, or other statistics. MLE provides optimal properties in data analysis: sufficiency, consistency, efficiency, and parameterization invariance [19].

Unfortunately, few studies have been reported on linear unmixing methods for multidimensional microscopy at low photon counts. The purpose of the present work is to develop and study MLE-based unmixing methods for multidimensional fluorescence spectroscopy at low signal levels. The performance of MLE unmixing has be evaluated at various signal levels by use of Monte Carlo simulations and experiments for the case of separation of three different fluorophores, which have spectrally overlapping excitation and emission spectra. The results show that MLE can out-perform WLS-based linear unmixing in the regime of low signal levels. Useful unmixing results for the three overlapping species can be obtained with signals of several hundred photons. The techniques used for determining errors and covariance of parameter estimates can be used in other applications with low signal levels. Such studies are important for advancing the scope of multidimensional spectroscopy to ultra-sensitive detection applications.

2. MAXIMUM LIKELIHOOD UNMIXING

The principle of MLE was primarily developed by Fisher in the 1920's. It states that the desired probability distribution is the one that makes the observed data "most likely". In other words, if measured data are to be described by a given statistical model that depends on a set of parameters, then the likelihood function is the probability to observe the particular measured data set for the assumed model and parameters, and the parameters that give rise to the greatest value of the likelihood function are the best estimates for those parameters [18].

For a set of n independent measurements, $\mathbf{m} = m_1, m_2, ...,$ and m_n, the likelihood to observe the measured data set $\mathbf{m}$ is simply the product of the likelihood functions for each of the measurements. If the likelihood functions all depend on the same model function, which in turn depends on a set of parameters, $\mathbf{f} = f^{(1)}, f^{(2)}, ..., f^{(k)}$, with $k < n$, then the likelihood function for the set of independent measurements may be expressed as the joint probability density

$$L(\mathbf{m};\mathbf{f}) = \prod_{i=1}^{n} l(m_i;\mathbf{f}), \tag{1}$$

where $l(m_i;\mathbf{f})$ is the likelihood function for the i-th measurement, i.e., the probability to obtain the measurement m_i for the given statistical model and set of parameters $\mathbf{f}$. The set of parameters $\mathbf{f}^*$ that results in the greatest value of $L(\mathbf{m};\mathbf{f})$ is the maximum-likelihood solution. It has been proven that in the limit of a large number of measurements, $\mathbf{f}^*$ converges without bias to the true value, and furthermore, there are no other methods of estimation that are more precise [20]. Also, for convenience, the entropy or logarithmic likelihood function

$$S(\mathbf{m};\mathbf{f}) = \ln L(\mathbf{m};\mathbf{f}) = \sum_{i=1}^{n} \ln l(m_i;\mathbf{f}), \tag{2}$$

is often used in calculations rather than $L(\mathbf{m};\mathbf{f})$, as the products are replaced by a summation.

In fluorescence measurements, signal acquisitions are often based on photon-counting. From a statistical point of view, photon-counting can be regarded as a random process, and the number of photons detected in a given measurement will be described by the Poisson distribution with a given mean. The probability to record the set of measured photon numbers, **m**, is then

$$L(\mathbf{m};\mathbf{f}) = \prod_{i=1}^{n} \exp\left[-\mu_i(\mathbf{f})\right] \mu_i(\mathbf{f})^{m_i} / m_i! \tag{3}$$

where $\mu_i(\mathbf{f})$ is the number of photons expected for the i-th measurement. The logarithmic likelihood function for the Poisson distribution is

$$S(\mathbf{m};\mathbf{f}) = \sum_{i=1}^{n} m_i \ln \mu_i(\mathbf{f}) - \mu_i(\mathbf{f}) - \ln(m_i!). \tag{4}$$

For the case of linear unmixing of the spectral signatures from k species, $\mathbf{m} = m_1, m_2, \ldots,$ and m_n represents the number of photons measured in each of the n spectral channels. Also, provided there are no interactions between the various species that are present, the mean number of photons expected in the i-th spectral channel, μ_i, is the sum of $\mu_i^{(j)}$, the mean number of photons expected in that channel from each of the j=1,…, k species:

$$\mu_i = \sum_{j=1}^{k} \mu_i^{(j)}. \tag{5}$$

Calibration measurements may be made to determine the normalized spectral signatures of each of the species, and also quantitative measures of the signals from known quantities of each species. For example, if only species X is present and a spectral signature $\mathbf{m}^{(x)}$ is collected for a large number of photons, then the normalized spectrum for species X is given by

$$s_i^{(x)} = m_i^{(x)} \Big/ \sum_{i=1}^{n} m_i^{(x)}. \tag{6}$$

Also, if the quantity of species X present is known, then the total number of photons collected from a given level of irradiance provides the information needed to quantify X for a particular measured signal level. The spectral signature of the background and the total background for a given level of irradiance can also be measured. Background can be accounted for in the analysis by treating it in the same way as another species. Note that the i = 1, …, n spectroscopic channels can denote different excitation wavelengths, emission bands, or temporal channels that capture sub-nanosecond timing information such as the fluorescence lifetime, or some combination of these. Once the normalized spectral signatures are known, then the mean number of photons expected in each spectral channel may be expressed as

$$\mu_i = M \sum_{j=1}^{k} f^{(j)} s_i^{(j)}, \tag{7}$$

where

$$M = \sum_{i=1}^{n} m_i \tag{8}$$

is the total number of measured photons, and $f^{(j)}$ is the fraction of the fluorescence signal due to species j, with $0 \le f^{(j)} \le 1$ and

$$\sum_{j=1}^{k} f^{(j)} = 1. \tag{9}$$

The set of species signal fractions $\mathbf{f} = f^{(1)}, f^{(2)}, \ldots, f^{(k)}$, are the unknown parameters to be determined by the linear unmixing. The MLE determination of the signal fractions is obtained by substituting Eqn. (7) into Eqn. (3) or (4) to obtain the likelihood function or the entropy and then finding the set of parameters **f*** that provide the largest likelihood or entropy. In principle, this can be done by solving the set of k equations $\partial L / \partial f^{(j)} = 0$, with $\partial^2 L / \partial f^{(j)2} < 0$, or equivalently $\partial S / \partial f^{(j)} = 0$, with $\partial^2 S / \partial f^{(j)2} < 0$. However, these sets of equations are in general analytically intractable and a numerical approach is required.

In previous work on the unmixing of exponentially-shaped spectra from species with differing fluorescence lifetimes, it was found that the entropy function is well approximated by a quadratic function over the entire interval $0 \le f^{(j)} \le 1$, so that a non-iterative solution with accuracy close to the precision afforded by the statistics could be very quickly obtained with a non-iterative algorithm [21]. It may prove useful to use such an approach for an initial view of the MLE-based unmixing, or as an initial estimate for an iterative solution. However, with the use of look-up tables and modern computers, it is possible to compute the likelihood function $L(\mathbf{m};\mathbf{f})$ at a grid of points throughout the range of possible signal fractions and to globally search for the maximum within the milliseconds of time that it takes to acquire the data for each pixel. For the numbers of photons of interest in this work, the statistical precision of the un-mixed component fractions is quite modest, ~ 10%, and hence one only needs to find the optimum values of the signal fractions to a precision of ~ 0.01. The grid of points is thus taken to be

$$f_{l_j}^{(j)} = 0.01 \times l_j \text{, for } l_j = 0, 1, 2, \ldots, 100 \text{, for each of the } j = 1,\ldots, k \text{ species.} \tag{10}$$

To avoid numerical underflow, the likelihood function may be scaled by an appropriate constant value, ξ. The MLE of the unmixing fractions is then given by the set of values $\mathbf{f}^* = f_{l_1}^{(1)}, f_{l_2}^{(2)}, \ldots, f_{l_k}^{(k)}$ at which the scaled likelihood $L'(\mathbf{m};\mathbf{f}) = \xi L(\mathbf{m};\mathbf{f})$ is a maximum. In order to estimate the precision of any given single estimate, the sum of the scaled likelihood over the entire grid of points is found

$$L'_\Sigma(\mathbf{m}) = \sum_{l_1=0}^{100} \sum_{l_2=0}^{100} \cdots \sum_{l_k=0}^{100} L'(\mathbf{m}; f_{l_1}^{(1)}, f_{l_2}^{(2)}, \ldots, f_{l_k}^{(k)}) . \tag{11}$$

This value is then used to determine the normalized probability density function at each grid point

$$P(\mathbf{m};\mathbf{f}) = L'(\mathbf{m};\mathbf{f}) / L'_\Sigma(\mathbf{m}) . \tag{12}$$

The normalized probability density function provides a robust measure of the error of an estimate. The value of $P(\mathbf{m};\mathbf{f}^*)$ equals the confidence that the actual signal fractions are within one grid-spacing of the set of values $\mathbf{f}^*$. Contours of the function $P(\mathbf{m};\mathbf{f})$ around the coordinates $\mathbf{f}^*$ at which the maximum occurs then provide a measure of the errors. The sum of the values of $P(\mathbf{m};\mathbf{f})$ within a given contour provides the confidence that the actual signal fractions are given by $\mathbf{f}^*$ within an error defined by that contour. For Gaussian distributed data, 68% of the distribution falls within plus or minus one standard deviation of the mean [22]. Therefore, we define the 1-σ confidence interval as the contour that encloses a sum of 0.68 of the values of $P(\mathbf{m};\mathbf{f})$.

In this paper, we investigate MLE-based unmixing of the signals from 3 particular fluorophores, Texas Red (*X*), Alexa 610 (*Y*), and Alexa 633 (*Z*). The fluorescence signal is collected in 4 emission wavelength bands and due to two excitation wavelengths provided by pulse-interleaved excitation with time-resolved photon detection, resulting in $n = 8$ spectroscopic channels. The relative response of the 8 spectroscopic channels can be adjusted by selection of the emission wavelength bands and the excitation wavelengths and laser powers. The spectroscopic signatures of each of the fluorophores determined from calibration measurements are $s_i^{(x)}$, $s_i^{(y)}$, $s_i^{(z)}$, which are normalized such that

$\sum_{i=1}^{8} s_i^{(x)} = \sum_{i=1}^{8} s_i^{(y)} = \sum_{i=1}^{8} s_i^{(z)} = 1$. The fractional contribution of fluorophores *X*, *Y*, and *Z* to the total signal are $f^{(x)}$, $f^{(y)}$, and $f^{(z)} = 1 - f^{(x)} - f^{(y)}$, where $0 \le f^{(x)} \le 1$ and $0 \le f^{(y)} \le 1$. Background is negligible and not included in the analysis, although it is considered in other analysis [21]. The likelihood function is found by insertion of the spectroscopic signatures $s_i^{(x)}$, $s_i^{(y)}$, and $s_i^{(z)}$ into Eqn. (7), which is then inserted into Eqn. (3).

3. SIMULATIONS

Monte Carlo simulations were developed to study the performance of the MLE-based unmixing at various signal levels. The simulation uses random numbers to generate data based upon a known mixture of fluorophores with known spectra. The data is analyzed by the above-described MLE-based unmixing algorithm and the estimated signal fractions are compared with the known fractions. The simulation was written in C/C++ and run on a personal computer (PC).

Fig. 1a shows the relative excitation and emission spectra of the three fluorophores selected for this study, Texas Red (*X*), Alexa 610 (*Y*), and Alexa 633 (*Z*). The spectra exhibit significant overlap and it is clear that the fluorescence from each species cannot be separated by physical means alone. The two selected excitation wavelengths are 580 and 632 nm, and the four emission bands are 585 – 618 nm, 619 – 631 nm, 633 – 649 nm, and 650 - 690 nm. Therefore, there are eight spectral channels (four emission bands with two excitation wavelengths). For the simulations, the laser powers are taken to be equal and differences in the molar absorptivities of the three fluorophores are ignored so that for each fluorophore, the expected relative response of a spectral channel is given by the product of the value of the relative excitation spectrum at the given laser excitation wavelength and the integral of the relative emission spectrum within the given emission band. The expected relative responses of the 8 spectral channels are normalized to find the spectroscopic signatures of the 3 species, as shown in Fig. 1b.

The excitation wavelengths and emission bands have been selected to minimize the overlap of the spectroscopic signatures. Here, this has been achieved by use of the 'Solver' feature of Microsoft Excel to adjust the excitation wavelengths and the emission bands so as to minimize the value of $\sum_{i=1}^{8} s_i^{(x)} s_i^{(y)} + s_i^{(y)} s_i^{(z)} + s_i^{(z)} s_i^{(x)}$, which gives a measure of the overlap of the spectroscopic signatures.

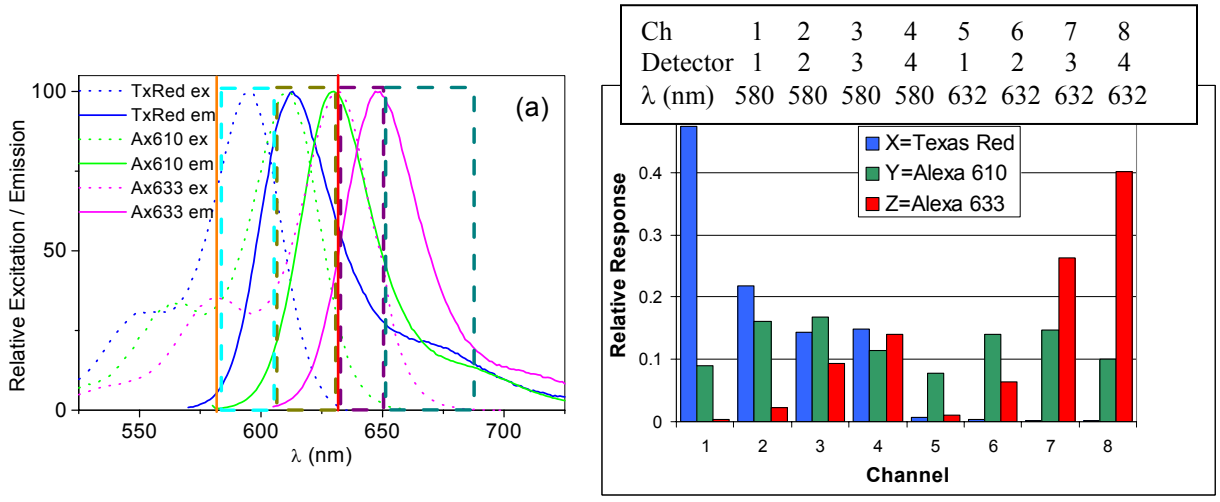

Figure 1. (a) Excitation spectra (dotted lines) and emission spectra (solid lines) of the three fluorophores, the 2 excitation laser wavelength (vertical solid lines), and the 4 emission bands (dashed boxes) into which fluorescence is collected. (b) The resulting normalized spectroscopic signatures of the three fluorophores.

4. EXPERIMENTS

The fluorophores used in this work (Texas Red, Alexa 610, and Alexa 633) were purchased from Invitrogen. The excitation peaks of the fluorophores are at 595, 612, and 653 nm, and the emission maxima occur at 613, 630, and 669 nm, respectively, as seen in Fig. 1. The concentrations of stock solutions are determined from the molar absorptivities using an absorption spectrophotometer and the stock solutions were serially diluted using purified water (Barnstead, NANOpure Infinity) to concentrations of ~ 100nM. Known mixtures of the samples were made for use in evaluation experiments.

A custom-built confocal epi-illumination microscope for single-molecule detection and fluorescence correlation spectroscopy was used for the experiments. Parts of the microscope are shown schematically in Fig. 2.

Two synchronously pumped dye lasers (Coherent 700) running Rhodamine 6G and DCM-special dyes were used for sample excitation. The dye lasers were pumped by a Spectra-Physics Vanguard modelocked frequency-doubled Nd:YAG laser, which produces 2W at 532nm with a pulse repetition rate of 76 MHz and pulse width of 10 ps. The dye laser wavelengths are continuously tunable from 560 – 605 nm and 600 – 680 nm for Rhodamine 6G and DCM special dyes, respectively. In this study, the two excitation wavelengths were tuned to 580 and 632 nm and the laser beams were attenuated to provide excitation powers at the samples of 30μW for each of the wavelengths.

As shown in Fig. 2a, two narrow-band Raman notch filters (RNF1 and RNF2) (Kaiser Optical Systems, HNPF – 585 – 1.0 and HNPF – 647 – 1.0) were used in the microscope. There are two purposes for using Raman notch filters. First, the dye laser beams contain a very small fraction of broadband incoherent fluorescence light that overlaps the collected fluorescence emission band. The Raman notch filters are angle tuned to efficiently reflect the 580 nm and 632 nm laser wavelengths, but not the broadband incoherent light, which would otherwise give rise to the predominant background contribution. Second, the Raman notch filters block specularly and elastically scattered light (i.e., Rayleigh scatter) from the sample with an optical density of > 6, while also providing ~ 90% transmission for other nearby wavelengths. Before each laser beam enters the microscope, it passes through two lenses, which are used to expand the beam so as to fill the back aperture of the water-immersion microscope objective lens (Olympus UPlanApo, 60×, 1.2 N.A. UPLAPO60XW). The objective focuses each beam to a diffraction-limited spot with waist of <0.5 μm, as determined by later fluorescence correlation spectroscopy calibration measurements. The two pairs of beam expanding lenses may be adjusted in position so as to give fine adjustment to the position of the beam waists, so that they may be overlapped with each other and with the confocal volume from which fluorescence is collected. The same objective collects fluorescence from the focal region, and the fluorescence passes through the Raman notch filters and is focused by a 250 mm focal length tube lens to a spatial filter with pinhole size of 150 μm. After the pinhole, the fluorescence passes through another lens, which almost collimates the beam. As shown in Fig. 2b, the fluorescence is then spectrally dispersed at a Brewster prism, passes through a second prism, and is split by mirrors with adjustable positions into 4 adjustable spectral bands. The lens after

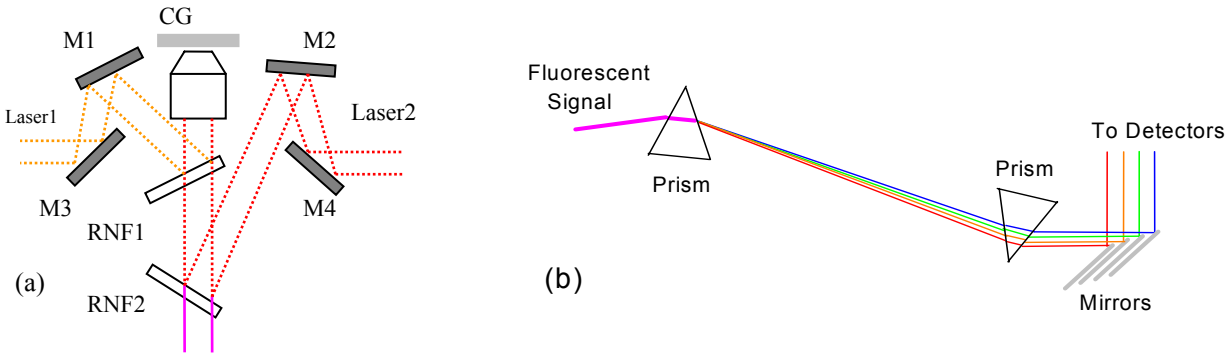

Figure 2. (a) A schematic of the excitation part of the microscope, including the objective. CG: Cover Glass, M1, M2, M3, and M4: Mirrors, RNF1 and RNF2: Raman Notch Filters for Laser1 and Laser2. (b) The prism spectrometer with the detection part of the microscope.

the spatial filter almost collimates the fluorescence but is adjusted so as to focus the fluorescence at the location of the mirrors, to improve spectral resolution. Each of the spectral bands is coupled into a fiber optic (Oz Optics, N.A. 0.29, 100 μm core /140 μm cladding) by use of a packaged optical lens assembly (Thorlabs F260FC-B) and the purpose of the second prism is to remove astigmatism to improve efficiency of coupling to the fiber.

The fibers pass the fluorescence to a 4-channel single-photon counting module (Perkin Elmer SPCM-AQ4C), which provides >50% single-photon detection efficiency. The TTL pulses from the photon counting module are sent to a PC-based time-to-digital converter (TDC) and photon time-stamping card (PicoQuant, TimeHarp 200) through a 4-channel router (PicoQuant, PRT400). The start signal for the TDC is provided by the single-photon detectors and a small part of one of the laser beams is directed to a fast photodiode (Hewlett Packard, 5082-4203), which provides a 76 MHz stop signal for the TDC. The electronics records the arrival time of each detected photon with respect to a preceding laser pulse with 34 ps digitization, and with respect to a 20 MHz clock. The electronics also records marker pulses from the piezo-scanner to enable collected photons to be registered within an image. Although the system is in principle capable of acquiring photon timing information to distinguish species with different fluorescence lifetimes, the SPCM-AQ4C detector was found to exhibit a count-rate dependent time-walk, which must be improved before such measurements become practical. This problem can be solved by modifying the detector electronics [23], but in this work only the excitation and emission wavelength signatures, and not the temporal profiles, are used for linear unmixing. The typical photon counts acquired for each data set are in the range of 1000 photons collected in 6 milliseconds, with ~0.5% of photons due to background. The electronics enable hundreds of sets of data to be rapidly acquired, for statistical analysis of stability and precision of unmixing of species contributions.

5. RESULTS AND DISCUSSION

Simulations were conducted for a variety of signal fractions and levels. Fig. 3 presents an example of results, where the known fractions of photons from fluorophores were set to be $f^{(x)} = 0.4$, $f^{(y)} = 0.3$, and $f^{(z)} = 0.3$. Fig. 3a, 3b, and 3c present results for MLE-based unmixing for a total photon signal of 50, 500, and 3000 photons respectively. In each figure, results are shown for 100 individual measurements. The scatter of the results of the unmixing (pink points) is indicative of the statistical precision of the analysis. The oval shape formed by the points indicates the covariance between the fitting parameters. A measure of the statistical precision of the analysis and covariance between the fitting parameters can be obtained from each individual measurement by viewing a contour plot of the normalized likelihood function (i.e., Eqn. (12)). This is shown in Fig. 3d for the unmixing of the data from the blue point in Fig. 3b. The black contour in Fig. 3d is the 1-σ confidence interval. There is a 68.3 percent confidence that the actual values of the component fractions lie within the contour. The shape and size of this contour are consistent with the scatter of points in Fig. 3b, obtained from many simulated measurements. The results indicate that for the chosen example of three species with significant spectral overlap, MLE-based unmixing can recover the signal fractions to useful precision even from a single measurement with a moderately low number of photons.

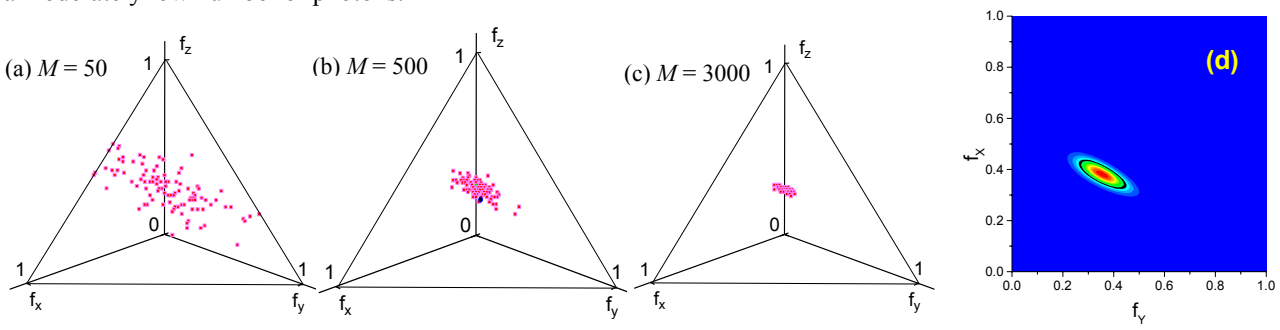

Figure 3. Simulation results for MLE-based unmixing of data for which the known signal fractions are $f^{(x)} = 0.4$, $f^{(y)} = 0.3$, and $f^{(z)} = 0.3$, for a total photon count of (a) $M = 50$, (b) $M = 500$, and (c) $M = 3000$. (d) The likelihood surface for a selected point, shown in blue in Fig. 3b. The black contour presents the 1-σ confidence interval.

Note that the fractions of signal from each species are constrained to be between 0 and 1. These constraints are equivalent to the requirement that the numbers of photons be positive and not greater than the total number of observed

photons. While the constraints ensure that every un-mixing result is physically possible, the constraints can introduce bias in the mean results from many measurements. In particular, when the contribution from one of the species is actually zero, the estimate can only be greater than or equal to zero, so with statistical errors, the mean of many measurements will be greater than zero, especially when the number of photons is small. We have found that such statistical bias can be avoided by relaxing the constraints. In this case, the mean is unbiased but individual unmixing results sometimes predict negative signal contributions.

To further investigate the relationship between the statistical errors of MLE-based unmixing and the total number of photons, a series of simulations was preformed for a wide range of total numbers of photons. Fig. 4 summarizes the results. The dotted horizontal lines show the actual signal fractions: $f^{(x)} = 0.4$, $f^{(y)} = 0.3$, and $f^{(z)} = 0.3$. The solid points show the mean estimated signal fractions of X (red square), Y (green triangle), and Z (blue circle) obtained from 100 individual simulated measurements. The standard deviations of the estimated fractions of photons from X, obtained from the 100 results, are also plotted. As expected, the precision of the unmixing estimates becomes better when the number of photons is increased. Interestingly, the MLE-based unmixing provides useful estimates of component fractions for significantly lower signal levels than are usually used in multi-dimensional microscopy. Also, the average of the results of the 100 simulated measurements provides a useful measure of the signal fractions for total photon counts as low as 100 or even 30 photons per measurement. An averaging effect is often provided in imaging applications, where information from nearby pixels is assimilated into the image.

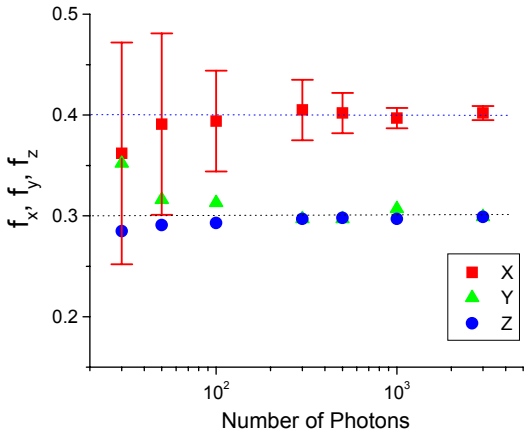

Figure 4. The estimated component fractions from fluorophores X, Y, and Z for various total number of photons.

To investigate the performance of MLE-based unmixing in an imaging application, the Monte Carlo simulations were extended to produce simulated images. Fig. 5 presents an example of results. The total image is shown in Fig. 5a. Each of the 100 × 100 pixels of Fig. 5 consists of photon counts collected within 8 spectroscopic channels, but only the total photon counts are shown in the image. The spectra at each of the pixels of Fig. 5a are formed by combining the spectra from the images shown in Figs. 5b, 5c, and 5d, which correspond to the known contributions from the three fluorophore X, Y, and Z. Fig. 5e presents the color scale, which shows that the number of photons from each species varies from 0 to about 1000 at the peak of an image feature. A pixel-by-pixel MLE-based unmixing of the spectra from Fig. 5a results in the estimated species contributions shown in Figs. 5f, 5g, and 5h. The features in the images of each of the species are usefully resolved, and the images are seen to reproduce the actual images with good fidelity. Comparison of the estimated and actual images of X, Y, and Z components show that the estimated images give very similar features, even in the low photon count regions (100 – 200 photons).

Simulations were also used to investigate the configuration of a single excitation wavelength of 580 nm and 8 emission bands, each of 11 nm width starting from 595 nm. This configuration is similar to that of the Zeiss's LSM 510 META microscope, which has 32 detectors each with a 10.7 nm bandwidth, covering the 380 – 720 nm emission range, but with a maximum output of eight spectral channels. The simulation results show that this configuration provides inferior but somewhat similar capability for unmixing Texas Red, Alexa 610, and Alexa 633 to that of the case of 2 excitation

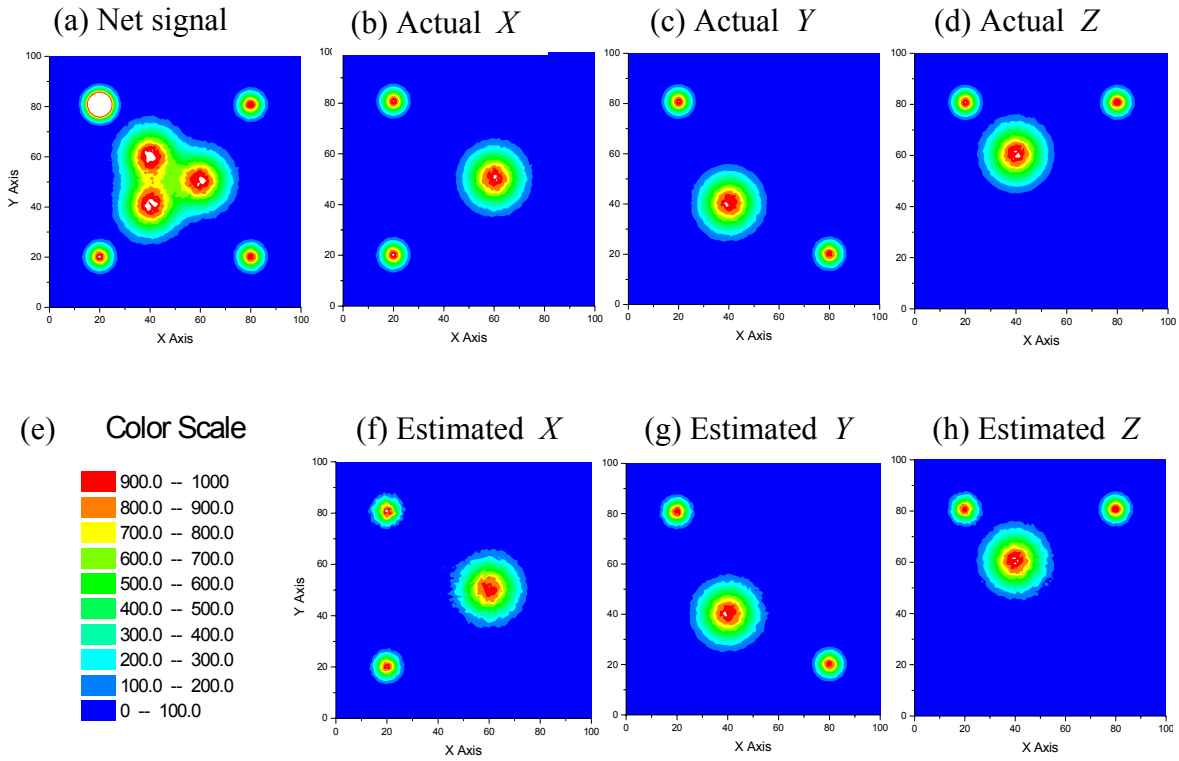

Figure 5. An application of MLE-based unmixing to image recovery: Panel (a) The net image was generated by combining images from species *X*, *Y*, and *Z*, shown in panels (b), (c) and (d). Panel (e) shows the color code. Panels (f), (g), and (h) show the estimated contributions from *X*, *Y*, and *Z*, obtained by MLE-based unmixing of the spectral data at each pixel of the image (a).

wavelengths and 4 emission channels. Fig. 6 presents an example of the component fractions of 100 individual measurements for a total signal of 300 photons, with results from the 8 detectors with single excitation wavelength shown in pink and results from the 4 detectors with two excitation wavelengths shown in blue. The green spot is the true signal fractions.

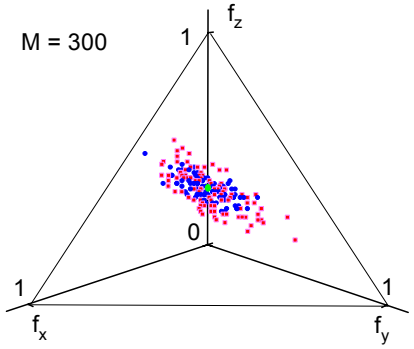

Figure 6. The distribution of the component fractions obtained from 2 excitation wavelengths with 4 detection channels (blue), and 1 excitation wavelength with 8 detection channels (pink).

In the experiments, seven different mixtures of Texas Red (*X*), Alexa 610 (*Y*), and Alexa 633 (*Z*), were prepared from serial dilution of mixtures of stock solutions, whose concentrations were determined to be 85nM, 72nM, and 110nM. These concentrations were determined using an absorption spectrophotometer, assuming extinction coefficients of 160000, 100000, and 239000 $M^{-1}cm^{-1}$ for Texas Red, Alexa 610, and Alexa 633, respectively. The expected fraction of fluorescence signal from each species in each of the mixtures depends upon the relative fluorescence brightness of each

species, which is difficult to measure accurately, and was estimated to be $X:Y:Z$ =1:2:1. Table 1 lists the calculated expected fluorescence contribution from each species in the mixture. The multi-channel emission spectra from the samples were recorded under the same laser irradiance conditions as used to acquire the spectroscopic signatures from each of the pure fluorophores. MLE-based unmixing was then used to estimate the fraction of signal from each species. Table 1 also lists the mean and standard deviation of MLE-based unmixing estimates from 100 measurements for each sample, with each measurement containing 500 photons. The results of the MLE-based unmixing estimates are seen to be in reasonably good agreement with the expected species contributions.

Table 1. Summary of experimental results for 7 mixtures at a total photon count of M = 500.

Sample	Texas Red		Alexa 610		Alexa 633	
	Expected	Estimated	Expected	Estimated	Expected	Estimated
S1	0	0.0±0.03	0.53	0.46±0.13	0.47	0.54±0.04
S2	0.38	0.36±0.08	0	0.0±0.05	0.62	0.64±0.06
S3	0.37	0.35±0.09	0.63	0.64±0.15	0	0.01±0.02
S4	0.23	0.20±0.07	0.4	0.43±0.12	0.37	0.37±0.03
S5	0.16	0.13±0.07	0.29	0.27±0.11	0.55	0.60±0.05
S6	0.17	0.23±0.09	0.57	0.48±0.14	0.26	0.29±0.04
S7	0.37	0.32±0.09	0.34	0.35±0.12	0.29	0.33±0.05

Fig. 7 presents the MLE-based unmixing results of 100 individual measurements of sample S7 for a total photon count of 500 in each measurement. In Fig. 7a, the scatter of the points gives the statistical precision and the distribution of the scattered points is indicative of the covariance of the fractional components. Fig. 7b shows the normalized likelihood surface for a selected point in Fig. 7a (which is shown in blue). The black contour in Fig. 7b is the 1-σ confidence interval. The elongated shape of the oval contour indicates that the sum of $f^{(x)}$ and $f^{(y)}$ has a relatively small uncertainty, i.e., $f^{(z)} = 1 - f^{(x)} - f^{(y)}$ can be determined to good precision, but the fractional contributions from species X and Y cannot be independently determined to high precision. The larger uncertainty in experiments than in the simulations is mostly because the experimental spectroscopic signatures exhibit greater overlap than those of Fig. 1b. A greater number of photons is required to reliably unmix the three fluorophores in our experiments, but improvements to the experimental parameters to provide greater spectroscopic distinction between the species should be possible.

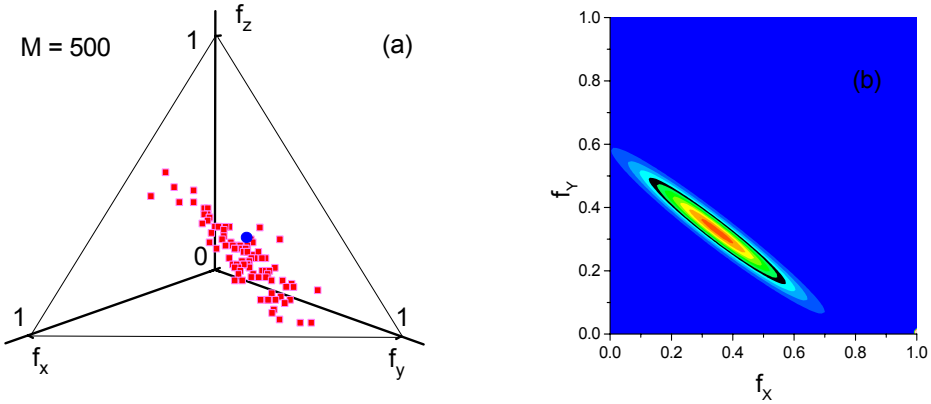

Figure 7. Results from MLE-based unmixing of experimental data: (a) The results from 100 measurements, each with a total photon count of 500. (b) The normalized likelihood surface from a selected measurement, which corresponds to the blue point in Fig 8a. The black contour shows the 1-σ confidence interval.

It should be noted that greater spectroscopic distinction between species may be possible at the expense of decreased sensitivity, by selection of a laser wavelength to minimize the excitation of one of the species. Also, improvements in the

timing performance of the single photon detectors should enable time-resolved measurements to utilize differences in fluorescence lifetimes to help unmix the contributions from different species.

Finally, the analyses of all simulation and experimental data were also performed using the weighted least-squares (WLS) based linear unmixing method, to compare its performance with that of MLE-based unmixing. As expected, the uncertainty of the component fractions increases with decreasing photon counts and is greater than that offered by MLE-based unmixing. However, simulations show that both estimators give similar predictions when the photon counts per spectroscopic channel increase above about 100. For 8 spectroscopic channels, when the total number of photons is below about 1000, MLE begins to provide better performance than WLS.

6. CONCLUSIONS

Monte Carlo simulations and experiments are used to study the use of multidimensional spectroscopy and maximum-likelihood based data analysis methods for unmixing the contributions from spectrally overlapping fluorophores in the limit of low photon counts. Studies focus on the capabilities for unmixing the signals from three fluorophores that have significant spectral overlap. A maximum-likelihood based procedure for estimating the species contributions and the errors and covariance of the contributions is described. Experiments are performed using a custom-built microscope, which incorporates two tunable excitation wavelengths and four emission bands. This system also provides the capabilities for multi-band fluorescence correlation spectroscopy and identification of photon bursts for single-molecule detection. However, a count-rate dependent time-walk in the single-photon counting timing module make fluorescence lifetime imaging applications presently impractical and improvements in the detector electronics are needed. The simulations and experiments show that MLE-based unmixing can outperform the usually used weighted-least squares analysis method when the number of photons collected in each spectroscopic channel is below about 100. The results also show that MLE-based unmixing provides a useful means to extend the application of multidimensional spectroscopy to situations in which the number of photons is limited, due to weak fluorescence or fast data collection times. The signals from three fluorophores with significant spectral overlap can be unmixed with useful precision even when the total number of photons in the measurement is as low as several hundred. Best performance is obtained when the spectroscopic channels are selected so as to provide the least overlap between the spectroscopic signatures of the fluorophores. MLE-based unmixing is expected to have application in ultrasensitive fluorescence lifetime imaging microscopy (FLIM) and spectral lifetime imaging microscopy (SLIM).

ACKNOWLEDGEMENT

This work was supported by research grant 1R03EB004586-01A1 from the National Institutes of Health.

REFERENCES

1. S.J. Morris, “Real-time multi-wavelength fluorescence imaging of living cells,” *Biotechniques*, **8**: 296-308 (1990).
2. H. Tsurui, H. Nishimura, S. Hattori, S. Hirose, K. Okumura, and T. Shirai, “Seven-color fluorescence imaging of tissue samples based on Fourier spectroscopy and singular value decomposition,” *J. Histochem. Cytochem.* **48**: 653–662 (2000).
3. M.E. Dickinson, G. Bearman, S. Tille, R. Lansford and S.E. Taylor, “Multispectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy,” *Bioimaging*, **31**: 1274 - 1278 (2001).
4. Y. Hiraoka Y, T. Shimi, T. Haraguchi, “Multispectral imaging fluorescence microscopy for living cells,” *Cell Struct. Funct*., 27: 367–374 (2002).
5. K. Weisshart, V. Jüngel, and S.J. Briddon, “The LSM 510 META - ConfoCor 2 System: An Integrated Imaging and Spectroscopic Platform for Single-Molecule Detection,” *Curr. Phar. Biotech.,* **5**: 135-154 (2004).
6. J.P. Knemeyer, D.P. Herten, and M. Sauer, “Detection and identification of single molecules in living cells using spectrally resolved fluorescence lifetime imaging microscopy,” *Anal. Chem.*, **75**: 2147-2153 (2003).
7. L. Zhu, W. Stryiewski, S. Lassiter, and S.A. Soper, “Fluorescence multiplexing with time-resolved and spectral discrimination using near-IR detector,” *Anal. Chem.*, **75**: 2280-2291 (2003).
8. T. Zimmermann, J. Rietdorf, and R. Pepperkok, “Spectral imaging and its applications in live cell microscopy” *FEBS Lett*. **546**: 87–92 (2003).

9. D.J. Stephens and V.J. Allan, "Light Microscopy Techniques for Live Cell Imaging," *Science*, **300**: 82-86 (2003).
10. R.C. Ecker, R. de Martin, G.E. Steiner, and J.A. Schmid, "Application of spectral imaging microscopy in cytomics and fluorescence resonance energy transfer (FRET) analysis," *Cytometry Part A*, 59: 172–181(2004).
11. R. Lansford, G. Bearman, and S.E. Fraser, "Resolution of multiple green fluorescent protein color variants and dyes using two-photon microscopy and imaging spectroscopy," *J. Biomed. Optics*, **6**: 311-318 (2001).
12. V. Seyfried, H. Birk, R. Storz, and H. Ulrich, "Advances in multispectral confocal imaging," SPIE Progress in biomedical optics and imaging, **5139**: 147 – 157 (2003).
13. C. Eggeling, J. Widengren, R. Rigler, and C. A. M. Seidel, "Photobleaching of fluorescent dyes under conditions used for single-molecule detection: evidence of two-step photolysis,'' *Anal. Chem.,* **70**: 2651–2659 (1998).
14. M. Maus, M. Cotlet, J. Hofkens, T. Gensch, F.C. De Schryver, J. Schaffer, and C.A.M. Seidel, "An experimental comparison of the maximum likelihood estimation and nonlinear least-squares fluorescence lifetime analysis of single molecules," *Anal. Chem.* **73**: 2078 - 2086 (2001).
15. P. Hall and B. Selinger, "Better estimates of exponential decay parameters", *J. Phys. Chem*., **85**: 2941 – 2946 (1981).
16. R.J. Ober, S.Ram, and E.S. Ward, "Localization Accuracy in Single-Molecule Microscopy," *Biophys. J*., **86**: 1185-1200 (2004).
17. G.F. Schoder and H. Grubmuller, "Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments,", *J. Chem. Phys.*, **119**: 9920 – 9924 (2003).
18. S. M. Kay, "Fundamentals of statistical signal processing, estimation theory", *Prentice Hall*, Englewood Cliffs, New Jersey, 1993.
19. I.J. Myung, "Tutorial on maximum likelihood estimation," *J. Math. Psych*., **47**: 90–100 (2003).
20. H. Cramer, "Mathematical Methods of Statistics," *Princeton Press*, New Jersey (1946).
21. L.M. Davis and L.R. Middendorf, "Increased throughput DNA sequencing by spectroscopically resolving the contributions of different fluorescent dyes," *US Patent Application publication No. 09/688,178* (2000).
22. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, "Numerical recipes in C," 2nd ed., *Cambridge University Press*, New York (1992).
23. I. Rech, I. Labanca, M. Ghioni, and S. Cova, "Modified single photon counting modules for optimal timing performance," *Rev. Sci. Instru*., 77: 033104 (2006).