Eur Biophys J (1991) 20:247–262

European Biophysics Journal

Maximum likelihood method for the analysis of time-resolved fluorescence decay curves

Željko Bajzer*, Terry M. Therneau, Joseph C. Sharp, and Franklin G. Prendergast

Department of Biochemistry and Molecular Biology and Department of Health Science Research, Mayo Foundation, Rochester, MN 55905, USA

Received November 14, 1990/Accepted in revised form September 18, 1991

Abstract. The usefulness of fluorescence techniques for the study of macromolecular structure and dynamics depends on the accuracy and sensitivity of the methods used for data analysis. Many methods for data analysis have been proposed and used, but little attention has been paid to the maximum likelihood method, generally known as the most powerful statistical method for parameter estimation. In this paper we study the properties and behavior of maximum likelihood estimates by using simulated fluorescence intensity decay data. We show that the maximum likelihood method provides generally more accurate estimates of lifetimes and fractions than does the standard least-squares approach especially when the lifetime ratios between individual components are small. Three novelties to the field of fluorescence decay analysis are also introduced and studied in this paper: *a)* discretization of the convolution integral based on the generalized integral mean value theorem: *b)* the likelihood ratio test as a tool to determine the number of exponential decay components in a given decay profile; and *c)* separability and detectability indices which provide measures on how accurately, a particular decay component can be detected. Based on the experience gained from this and from our previous study of the Padé-Laplace method, we make some recommendations on how the complex problem of deconvolution and parameter estimation of multiexponential functions might be approached in an experimental setting.

Key words: Maximum likelihood – Fluorescence decay

1 Introduction

The extraction of fluorescence decay lifetimes from time-correlated single photon counting (TC SPC) measurements is a problem with a long history, a consequence of the difficulties involved. The issue centers on two numerically ill-conditioned problems: deconvolution with respect to the instrument reponse function (IRF), and parameter estimation of multiexponential functions. Favorable characteristics of the problem relate to the possibility that the level of noise in the data can be controlled, and to the fact that one can obtain dense sampling of the fluorescence intensity decay curve.

The essential property of any method for the analysis of multicomponent decay curves determined experimentally, and contaminated by noise, is the ability to separate the individual components (*separability, or resolvability*). For a given level of noise, a given number of sampled data, and a given algorithm, there is a *critical lifetime ratio* and a *critical fraction ratio* beyond which the corresponding components cannot be resolved. Thus, if the lifetime ratio is smaller than the critical lifetime ratio, or if the fraction ratio is greater than the critical fraction ratio, the corresponding two components are likely to be detected by the algorithm as one component. It is also essential for the method to be *accurate*, i.e. the values obtained for the parameters must be sufficiently close to the "true" values. In ideal circumstances the latter can be known from experiments in which the contribution of each component has been determined prior to creation of a mixture of components. In the case of synthetic data the values of the parameters have been selected.

These considerations become particularly significant for the interpretation of the fluorescence intensity decays in macromolecules. For example, the interpretation of heterogeneous tryptophan fluorescence lifetimes in a protein bearing a single tryptophan residue in terms of discrete exponential terms is frequently held to imply unique, distinguishable conformational states of the protein. Moreover, the actual values of the lifetimes given by the analysis may be used to infer the physical nature of the interaction of the indole moiety with its environs. Given the usefulness of fluorescence techniques for the study of macromolecular structure and dynamics, especially in the picosecond to nanosecond timescale, it is

* On leave of absence from Rudjer Bošković Institute, Zagreb, Croatia, Yugoslavia

Offprint requests to: F. G. Prendergast

easy to understand why the experimenter needs accurate data on the fluorescence decay constants.

Many methods of multicomponent decay curve analysis and deconvolution have been applied to estimate decay lifetimes and fractions from the TCSPC data: *i*) The method of moments (Isenberg and Dyson 1969; Small et al. 1989); *ii*) the modulating functions method (Valeur and Moirez 1973; Szabo and Bramal 1983); *iii*) the exponential series method (Ware et al. 1973; Siemiarczuk et al. 1990); *iv*) the least squares method with iterative reconvolution (LS) (Grinvald and Steinberg 1974; Periasamy 1988; Wong and Harris 1989); *v*) the Laplace transform method (Gafni et al. 1975; Ameloot and Hendrickx 1983); *vi*) the Fourier transform method (Andre et al. 1979; Good et al. 1984); *vii*) the phase plane method (Demas and Adamson 1971; Jezequel et al. 1982); *viii*) the integral equation method (Provencher 1982a; Gregory and Zhu 1990); *ix*) the maximum likelihood method (Heinde et al. 1977; Hall and Selinger 1981; 1984); *x*) the global analysis approach (Eisenfeld and Ford 1979; Knutson et al. 1983; Ameloot et al. 1986; Beechem and Gratton 1988; Janssens et al. 1990); and most recently *xi*) the maximum entropy method (Livesey and Brochon 1987; Siemiarczuk et al. 1990), and *xii*) the generalized Padé-Laplace method (GPL) (Bajzer et al. 1990a). The very fact that so many methods have been proposed, and used, suggests that none of them is completely satisfactory.

O'Connor et al. (1979) performed a relatively detailed comparison of nine techniques applied to the analysis of measured fluorescence decay curves for well characterized fluorophores. They recommended the LS method because it "can be used with no loss in accuracy to fit any chosen section of the decay curve". This statement ignores the observation that some parts of the curve carry more information on a given component than others. In this paper we show explicitly that loss of accuracy occurs if the tail of decay curve is omitted.

The most complete studies, by simulation, considered seven different methods, and were performed by McKinnon et al. (1977), but only for two-component decays. They concluded that the LS method is "marginally better in its sensitivity to noise and its ability to separate components when tested on simulated data with a rapidly decaying lamp profile", but is superior to other methods when there is a long tailed instrument response function. However, since present day laser technology offers a very narrow decaying IRF, there is no need to discriminate between the methods on the basis of their ability to accommodate a long tailed IRF.

The preference for LS is reiterated in the book of O'Connor and Phillips (1984). Most recently the LS method was highly regarded in "Recommended Methods for Fluorescence Decay Analysis" (Eaton 1990). The effectiveness of the LS approach is further enhanced by the use of "global analysis". The latter states that the separability attained by the LS method can be further improved by use of simultaneous analysis of several data sets (Knutson et al. 1983). Thus, today the LS approach is generally considered the "gold standard", and if another method is to be considered a replacement, there must be compelling justification. By use of numerous simulations and an elaborate procedure for comparison of different methods, we have recently shown (Bajzer et al. 1990a) that under certain conditions the GPL method provides more accurate estimates than the LS method. Similarly, Small et al. (1989) recently pointed out advantages of the method of moments over the standard LS approach. These most recent findings, and in general the multiplicity of the methodologies which have already been developed, suggest clearly the uncertainty investigators still feel regarding the validity of any one method of analysis, irrespective of the statistical criteria used to buttress the results obtained.

In the present paper we revisit the maximum likelihood method. This method was introduced for fluorescence data analysis by Hall and Selinger (1981; 1984), but apparently it has not been developed subsequently. These authors did not include convolution with the IRF in their work, and they basically limited their studies to one- and two-component decays, with the exception of some formal developments in their Appendix (Hall and Selinger 1984). They were primarily concerned with the development of a special algorithm for maximizing the likelihood function which would take advantage of analytical properties of multiexponential functions. In their first paper, Hall and Selinger (1981) provided analytical solutions for the case of a single exponential model using least squares, the method of moments, and the maximum likelihood estimators. These estimators were compared and contrasted, and the authors concluded that the maximum likelihood estimator provided the best choice. An important and key difference between the LS and ML methods for the analysis of time correlated photon counting fluorescence decay data is the assumption of Poisson statistics in the ML method to treat the noise in the data, versus the assumption of Gaussian statistics used in the LS analysis. We show in this paper how this difference in treatment of noise affects the analysis.

Thus, the rationale for this paper may be summarized as follows: *a*) The interpretation of the fluorescence intensity decay time in terms of protein structure and dynamics depends on the accuracy of the methods used for data analysis; *b*) There are several indications the LS method used as a "golden standard" is not always the most accurate; *c*) The ML method was applied to the analysis of very particular cases of fluorescence decays with encouraging results (Hall and Selinger 1984); *d*) Our preliminary results have shown (Bajzer et al. 1990b), that when the estimates from the ML and LS methods are compared for accuracy in parameter recovery (lifetimes and fractions), for small lifetime ratios (say 1.3) the ML method is superior; *e*) The maximum likelihood approach has been thoroughly studied in the statistical literature and is known to be most powerful method for parameter estimation; *f*) Someone accustomed to the LS method will have no difficulty in using the ML method because the practical procedures of analysis are exactly the same and for all practical purposes the methods are computationally equivalent.

The main *goals* of this paper are: *1*) To formulate and develop the ML method for fluorescence decay analysis

with any number of components and with the instrument response function taken into account. *2)* To investigate how accurate are the ML and the LS parameter estimates under various conditions of number of components, lifetimes ratios, fraction ratios, channel widths and signal to noise ratios.

In Sect. 2 of the paper we accomplish the first main goal. The maximum likelihood method with iterative reconvolutions is presented in Sect. 2.1 and in Sect. 2.2 the properties of the ML estimator are considered. Specifically, we introduce "separability" and "detectability" indices which provide information a priori, whether (and how accurately) a particular decay component can be detected by use of the maximum likelihood method. We also consider the likelihood ratio test as a tool to determine how many exponential decay components are required to adequately fit the fluorescence intensity decay data. This test has a long history in statistics, but apparently has never been systematically employed in the field of fluorescence intensity decay analysis. Subsection 2.3 is devoted to a new discretization scheme of the convolution integral based on the generalized integral value theorem and the minimax principle.

Section 3 of this paper is devoted to the second main goal. By using over 7 000 simulated fluorescence intensity decay profiles, we have tested the performance of the ML and LS methods under different conditions with regard to number of decay components, lifetime ratios, fraction ratios, channel width and signal to noise ratio. We also tested the performance of the modified least squares method (LSM) which includes weights defined by the model function instead by measured data as in the standard LS method (Awaya 1979; Hall and Selinger 1981; Provencher 1982a; Selinger and Harris 1983). This approach has not been widely used but is known to provide more accurate estimates than the standard LS method. From our results it is clear that tests by simulation offer much more flexibility than can tests on real data. As a consequence, some hidden features of the method being considered can be revealed. We have therefore devoted a good deal of the present paper to testing by simulation in order to thoroughly evaluate the methodology. Application of our ML method the analysis of fluorescence decay in proteins has been reported elsewhere (Haydock et al. 1990).

To compare the ML method with the LS approach and the GPL method we use the average relative errors per parameter (Eisenfeld and Ford 1979; Bajzer et al. 1990a). The results show that the maximum likelihood method generally provides most accurate estimation of lifetimes and fractions. In conclusion, we offer some recommendations on how to approach the analysis of fluorescence intensity decay curves by use of the maximum likelihood method in combination with other methods and how to design the additional experiments to increase the amount of information in data necessary to obtain most accurate estimate of parameters. The appendix is devoted to the comparison of the ML the LS and the LSM methods from a computational standpoint, and to the computational aspects of the new discretization scheme.

2 Theory

2.1 The maximum likelihood method

The maximum likelihood method is considered to be the most powerful statistically based method of parameter estimation (see for example Kendall and Stuart (1979); Meyer (1975); Lehmann (1981)). It can be outlined briefly as follows.

Let us assume that we have n independent observations (counts) $c_1, \ldots, c_n$. If the probability of i-th observation is $p(c_i; \boldsymbol{\theta})$, where $\boldsymbol{\theta}=(\theta_1, \ldots, \theta_m)$ is a parameter vector, the joint probability, which is called the *likelihood function* can be written as

$$L(c_1, \ldots, c_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} p(c_i, \boldsymbol{\theta}). \tag{1}$$

The ML method then provides as an estimate of the "true" value of the parameter vector $\boldsymbol{\theta}$, a particular vector, say $\hat{\boldsymbol{\theta}}$, which maximizes the likelihood function. In the case of time correlated single photon counting measurements the observed counts c_i can be assumed to follow the Poisson distribution,

$$p(c_i; \boldsymbol{\theta}) = \mathrm{e}^{-\langle c_i \rangle} \langle c_i \rangle^{c_i} / c_i\,!\,, \tag{2}$$

where $\langle c_i \rangle$ is the expected value of the number of counts in the i-th channel. This expected value for the fluorescence decay process is modeled by (O'Connor and Phillips 1984):

$$\langle c_i \rangle = \int_{(i-1)h}^{ih} I_0(t)\,\mathrm{d}t + b_i \equiv g_i(\boldsymbol{\theta}),\ i=1, \ldots, n\,, \tag{3}$$

where h denotes channel width (time calibration) and b_i denotes the number of background counts in the i-th channel determined as an average dark count per channel (O'Connor and Phillips 1984). $I_0(t)$ is the "observed" fluorescence decay function expressed as a convolution of the instrument response function $R(t)$ and the actual fluorescence intensity decay law $I(t)$:

$$I_0(t) = \int_0^t R(t+\delta-u)\, I(u)\,\mathrm{d}u = \int_0^t R(u+\delta)\, I(t-u)\,\mathrm{d}u, \tag{4}$$

$$R_i = \langle R_i \rangle + v_i,\ \langle R_i \rangle = \int_{(i-1)h}^{ih} R(t)\,\mathrm{d}t,\ i=1, \ldots, n, \tag{5}$$

$$I(t) = \sum_{k=1}^{N} A_k\, \mathrm{e}^{-t/\tau_k},\ A_k = f_k/\tau_k,\ \sum_{k=1}^{N} f_k = f,\ N=(m-1)/2. \tag{6}$$

R_i is the observed number of counts in channel i obtained by measurement of the instrument response function (it is assumed that background count per channel is determined as above and subtracted). $\langle R_i \rangle$ and v_i are the corresponding expected value and noise respectively. The parameter vector $\boldsymbol{\theta}$ is given by lifetimes and fractions, $\theta_{2k-1}=\tau_k$, $\theta_{2k}=f_k$, $k=1, \ldots, N$, and by the zero-time shift, $\theta_m=\delta$.

The log-likelihood function $\ln L$ attains its maximum for the same value of $\boldsymbol{\theta}$ as the likelihood function, L. It is customary to determine $\hat{\boldsymbol{\theta}}$ by minimizing $-\ln L$, which in the case of Poisson distribution is equivalent to minimiz-

ing the Poisson deviance (McCullagh and Nelder 1983):

$$D(\boldsymbol{\theta})=2\sum_{i=1}^{n}\{c_i\ln[c_i/g_i(\boldsymbol{\theta})]-[c_i-g_i(\boldsymbol{\theta})]\}\,. \tag{7}$$

In the standard LS approach used for fluorescence decay analysis (Grinvald and Steinberg 1974; O'Connor and Phillips 1984) which was recently recommended by the International Union of Pure and Applied Chemistry (Eaton 1990), the function to be minimized is formally quite different:

$$\chi^2=\sum_{i=1}^{n}[c_i-g_i(\boldsymbol{\theta})]^2/c_i\,. \tag{8}$$

This is the deviance for a Gaussian distribution of the observed counts, with a standard deviation σ_i of each count estimated as $\sqrt{c_i}$ [1]. When the number of counts is large, however, (7) and (8) are numerically close, as a consequence of the central limit theorem by which the Poisson distribution can be approximated by the Gaussian distribution (Meyer 1975). Thus, from a theoretical point of view, the ML *estimator will differ from the LS estimator by the extent to which the low-count region of the fluorescence decay curve determines the values of parameters.*

When low count numbers are involved, several authors (Awaya 1979; Provencher 1982a; Selinger and Harris 1983) proposed the following modification of deviance (8):

$$\tilde{\chi}^2=\sum_{i=1}^{n}[c_i-g_i(\boldsymbol{\theta})]^2/g_i(\boldsymbol{\theta}) \tag{9}$$

which is a Gaussian deviance with standard deviation $\sigma_i=\sqrt{g_i(\boldsymbol{\theta})}$. We will refer to this approach as the modified least-squares method (LSM). The minimization of deviance (9) yields less biased estimates of parameters than the minimization of deviance (8) (Awaya 1979; Selinger and Harris 1983). However, as shown by Awaya (1979, 1980) for analysis of low radioactivity decay curves, the minimization of Poisson deviance (7) yielded more accurate estimates than those obtained by minimization of (8) or (9). Results of our simulations (Sect. 3) have shown the same.

To compare the deviances (7), (8) and (9) we expanded the expressions under summation in a Taylor series with respect to the quantity $\xi_i=(c_i-g_i)/g_i;\ g_i\equiv g_i(\boldsymbol{\theta})$:

$$D(\boldsymbol{\theta})=\chi^2+\sum_{i=1}^{n}g_i[2\xi_i^3/3+O(\xi_i^4)]=\tilde{\chi}^2-\sum_{i=1}^{n}g_i[\xi_i^3/3+O(\xi_i^4)]. \tag{10}$$

Clearly, for very small ξ_i (large g_i and close to c_i) the three deviances are numerically very close. However when small numbers of counts (say <100) are included, ξ_i becomes relatively large and these deviations may differ substantially.

[1] If in a certain channel j, $c_j=0$, the corresponding term in the sum is omitted. Such cases occur very infrequently since the background is included in the number of counts (see (3)). Usually $n=512$, and in simulations we performed the number of counts was zero, at most, in a few channels, so that neglecting the corresponding channels cannot significantly affect the estimation of parameters

The minimization of deviances (7), (8) or (9) is known to be a difficult numerical problem with solutions often very dependent on the initial guess for parameter values. The Nelder-Mead simplex minimization algorithm (Nelder and Mead 1965) has been advocated (see e.g. Phillips and Eyring 1988) as being robust with respect to starting parameter values and has recently been used in fluorescence decay analysis (Wong and Harris 1989). This was an incentive to apply it in our investigations – the implementation of Press et al. (1986) was used. Besides the simplex algorithm we also employed the modified Levenberg-Marquardt algorithm (More 1977; Morris 1981) known as one of the best presently available.

In the standard LS method the goodness of fit is usually assessed by the value of the reduced chi-square, $\chi_v^2=\chi^2/v$, $v=n-m$ and by inspection of residuals: $[c_i-g_i(\boldsymbol{\theta})]/\sqrt{c_i}$. In the ML approach $D(\boldsymbol{\theta})/v$ corresponds to the reduced chi-square, and normally distributed residuals can be defined (McCullagh and Nelder 1983) as: $\sqrt{2}[c_i\ln(c_i/g_i(\boldsymbol{\theta})-c_i+g_i(\boldsymbol{\theta})]^{1/2}\,\mathrm{sign}[c_i-g_i(\boldsymbol{\theta})]$.

2.2 Properties of maximum likelihood estimator

Two aspects of the maximum likelihood estimator will be of interest to us. The first concerns estimation, i.e., how "close" is the ML estimator to the truth; and the second has to do with testing, i.e., evaluation of whether a given component is actually present. The properties of the ML estimator related to these aspects are summarized in basic statistical texts (Cox and Hinkley 1974; Meyer 1975; Kendal and Stuart 1979; Lehman 1981).

The uncertainty of the maximum likelihood estimator $\boldsymbol{\theta}$ is related to the concept of *information about the parameter vector* $\boldsymbol{\theta}$ which we explicitely define for the benefit of the reader. For a given model function $g_i(\boldsymbol{\theta})$ the information about $\boldsymbol{\theta}$ contributed by the i-th number of counts c_i (with an error distributed according to the Poisson distribution) is the matrix $\mathbf{I}_i(\boldsymbol{\theta})$ defined by elements:

$$I_{ik}=\left\langle\frac{\partial}{\partial\theta_j}\ln p(c_i;\boldsymbol{\theta})\frac{\partial}{\partial\theta_k}\ln p(c_i;\boldsymbol{\theta})\right\rangle=\frac{1}{g_i(\boldsymbol{\theta})}\frac{\partial g_i(\boldsymbol{\theta})}{\partial\theta_j}\frac{\partial g_i(\boldsymbol{\theta})}{\partial\theta_k}\,. \tag{11}$$

The total information about $\boldsymbol{\theta}$ contributed by the set of counts $\{c_1,\ldots,c_n\}$ is

$$\mathbf{I}(\boldsymbol{\theta})=\sum_{i=1}^{n}\mathbf{I}_i(\boldsymbol{\theta})\,. \tag{12}$$

For any non-biased estimator $\boldsymbol{\theta}_e$ of $\boldsymbol{\theta}$, and any vector $\boldsymbol{a}=(a_1,\ldots,a_m)$ of constants, irrespective of the observed data set and the estimation method, the following inequality is valid:

$$\mathrm{var}(\boldsymbol{a}\cdot\boldsymbol{\theta}_e)\geq\boldsymbol{a}[\mathbf{I}(\boldsymbol{\theta})]^{-1}\boldsymbol{a}^T\,. \tag{13}$$

This is a special case of "information (or Cramér-Rao) inequality". A biased estimator can easily have smaller variance, for instance $\hat{\boldsymbol{\theta}}/10$ is biased towards zero, but with 1/100th the variance.

We found this inequality especially useful for obtaining some information a priori on the possible separability of

two components. Sandor et al. (1970) defined the *index of separability* $\bar{S}_{kl}$ for components with lifetimes τ_k and τ_l as

$$\bar{S}_{kl}=\frac{|\tau_k-\tau_l|}{[\mathrm{var}(\tau_k)+\mathrm{var}(\tau_l)]^{1/2}} \tag{14}$$

The higher the value of this index the higher is the probability that the components can be separated. Now, by using the inequality (13) we can calculate the upper bound of the index of separability:

$$\bar{S}_{kl}\leq\frac{|\tau_k-\tau_l|}{(I'_{kk}+I'_{ll})^{1/2}}\equiv s_{kl},\ I'_{kk}=[\mathbf{I}^{-1}]_{2k-1,\,2k-1},\ k=1,\ldots,N \tag{15}$$

Thus, s_{kl} can be calculated for a given instrument response function and assumed values of parameters (including the value of the background counts per channel), and provides information on whether two components can be separated before estimation of the parameters is attempted. It is clear that if $s_{kl}<1$, the two components are likely to be nonseparable.

In a similar way we introduce the *index of detectability*, $\bar{D}_k$, for the fraction f_k:

$$\bar{D}_k=\frac{f_k}{[\mathrm{var}(f_k)]^{1/2}}\leq\frac{f_k}{(I''_{kk})^{1/2}}\equiv d_k,\ I''_{kk}=[\mathbf{I}^{-1}]_{2k,\,2k}, \tag{16}$$
$$k=1,\ldots,N.$$

Again, the upper bound of the index of detectability d_k can be calculated before estimation of the parameters is attempted, and thus provides the information a priori on whether a component is detectable. For $d_k<1$ a considered component is likely to be undetectable.

As an overall error measure we have used the average relative error in parameters,

$$E=\frac{1}{2N}\sum_{k=1}^{N}\left(\left|\frac{\tau_k-\tau'_l}{\tau_k}\right|+\left|\frac{f_k-f'_l}{f_k}\right|\right) \tag{17}$$

where τ'_l are estimated lifetimes and f'_l are corresponding estimated fractions. The index l is such that for a given k, $|\tau_k-\tau'_l|$ is minimal. This error measure successfully characterizes the overall accuracy of parameter estimation when the true values of parameters are assumed (Eisenfeld and Ford 1979; Bajzer et al. 1990a).

The maximum likelihood estimate is related to the information matrix in a key way: As the information about $\boldsymbol{\theta}$ increases the bias of $\hat{\boldsymbol{\theta}}$ tends to zero and the covariance matrix $\mathbf{V}$ of $\hat{\boldsymbol{\theta}}$ tends to $\mathbf{I}^{-1}$. By the information inequality (13), $\hat{\boldsymbol{\theta}}$ is asymptotically fully efficient, i.e. no other unbiased estimator can have a smaller variance. We will measure the information about $\boldsymbol{\theta}$ with

$$M=(\mathrm{Trace}\,\mathbf{I}^{-1})^{-1}, \tag{18}$$

which is approximately the reciprocal value of the sum of the variances of $\hat{\boldsymbol{\theta}}$, so that it increases as the variances decrease. The covariance matrix $\mathbf{V}$ of $\hat{\boldsymbol{\theta}}$ can be also approximated by the inverse Hessian matrix:

$$\mathbf{V}\approx[\mathbf{H}(\hat{\boldsymbol{\theta}})]^{-1},\ H_{jk}(\boldsymbol{\theta})=\frac{1}{2}\frac{\partial^2 D(\boldsymbol{\theta})}{\partial\theta_j\,\partial\theta_k}. \tag{19}$$

As the information about each parameter increases, it is also true that ML estimates will assume a Gaussian distribution, and confidence limits for each parameter can be computed as $\pm 1.96\sigma(\theta_j)$, $\sigma(\theta_j)=\sqrt{V_{jj}}$.

For a given measured fluorescence intensity decay profile one does not know a priori the number (N) of multiexponential components present. To obtain this information we could use some other method of analysis which intrinsically provides this number (e.g. the generalized Padé-Laplace method (Bajzer et al. 1990a)), but we can also rely on the *likelihood ratio test*. This test is based on the ratio of likelihood functions at the maximum for two corresponding nested models (e.g. two- and one-component models). The logarithm of this ratio equals the difference of related deviances, and the following applies:

Assume that a sequence of models with 3, 5, 7, ... parameters has been fit to a given data set and that $D_3, D_5, D_7, \ldots$ are the resulting deviances. If the true model has r parameters then D_r-D_{r+2}, the *apparent* improvement in fit, will be distributed as a chi-square distribution on two degrees of freedom, χ^2_2. The difference $D_{r-2}-D_r$ will be distributed as a random variable which is stochastically larger than a χ^2_2; how much larger depends both on the lack of fit and the amount of information.

The practical test of whether a particular fluorescence decay profile can be described by $N-1$ or N components may be then formulated as follows: First an arbitrary cutpoint is chosen, for example $Pr(X>9.2)=0.01$ when X is chi-squared on two degrees of freedom. Then, if for the difference of deviances the inequality

$$\Delta_N\equiv D_{2N-1}-D_{2N+1}\geq 9.2 \tag{20}$$

is valid, the model with $N-1$ components is rejected as unacceptable. The probability that a model with $N-1$ components will be erroneously rejected, i.e., that too many components are chosen, is then 0.01.

Based on simulations (described below), in Fig. 1 we show how the difference of deviances varies with the lifetime ratio for 2, 3 and 4 exponential components. This figure clearly reveals that the threshold in lifetime ratio for detection of the exact number of components increases with the number of components, which reflects the known fact that the problem of fitting data with multiexponential functions becomes increasingly more difficult as the number of exponentials increases.

2.3 *Discretization of the convolution integral*

Standardly accepted discretization of the convolution integral (4) for fluorescence intensity decay analysis (Grinvald and Steinberg 1974) is based on the trapezoid rule for integration. McKinnon et al. (1977) proposed in the supplement of their paper another discretization scheme based on piecewise linear approximation of the IRF and explicit analytical integration of the resulting integrals. They found this scheme to be more accurate when short lifetimes are involved (i.e. when h/τ_k is not very small). Recently, this was independently confirmed by numerous simulations (Periasamy 1988)[2]. However Isenberg (1983)

[2] We note that in these simulations the counts of IRF, R_i, were inappropriately simulated by $R((i-1)h)$ instead by the integral in (5)

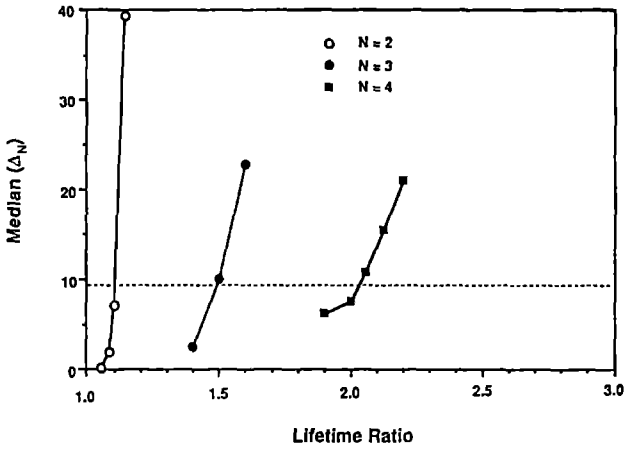

Fig. 1. The median of difference between deviances for $N-1$ and N components given by (17) as a function of lifetime ratio and N. For a given N and the lifetime ratio the median is determined from differences obtained by applying the ML method to the analysis of 101 simulated data sets corresponding to different realizations of the Poisson noise. The lifetime ratio is defined as τ_2/τ_1 for $N=2$, $\tau_2/\tau_1=\tau_3/\tau_2$ for $N=3$, and $\tau_2/\tau_1=\tau_3/\tau_2=\tau_4/\tau_3$ for $N=4$ where $\tau_1=1$ ns, zero time shift $\delta=0$, and the components were characterized by equal fractions. The simulations were based on measured IRF comprising 512 data points, 40 000 counts at peak and channel width of 0.025 ns/channel. The peak value of simulated fluorescence intensity decay profile was limited to 40 000 counts + noise, and the background level was set to zero. The full width half maximum of IRF was 0.09 ns

indicated that Simpson integration is more precise than the discretization scheme of McKinnon et al. We present here a new scheme whereby the maximal numerical integration error is minimized. In addition we explicitly derive the expressions with $\delta \neq 0$.

Our aim is to express the expected number of counts $\langle c_i \rangle = g_i(\boldsymbol{\theta})$ as a function of parameters τ_k, f_k, $k=1, \ldots, N$, and δ, and observed counts of IRF, R_i. Equations (3–6) imply:

$$g_i(\boldsymbol{\theta}) = \sum_{k=1}^{N} A_k \int_{(i-1)h}^{ih} e^{-t/\tau_k} \int_0^t R(u+\delta)\, e^{u/\tau_k}\, du\, dt + b_i , \quad (21)$$

and by partial integration with respect to t we obtain:

$$g_i(\boldsymbol{\theta}) = \sum_{k=1}^{N} f_k [e^{-h(i-1)/\tau_k} G_{i-1} - e^{-hi/\tau_k} G_i + R_i^{\delta}] + b_i , \quad (22)$$

$$G_i = \int_0^{ih} R(u+\delta)\, e^{u/\tau_k}\, du , \; R_i^{\delta} = \int_{(i-1)h}^{ih} R(u+\delta)\, du . \quad (23)$$

By using the generalized mean value theorem of integral calculus (cf. Bronshtein and Semendyayev 1985) we can express G_i through R_i^{δ},

$$G_i = \sum_{j=1}^{i} F_j , \; F_j = \int_{(j-1)h}^{jh} R(u+\delta)\, e^{u/\tau_k}\, du = e^{h(j-1+\eta_j)/\tau_k} R_j^{\delta} ,$$
$$0 \leq \eta_j \leq 1 . \quad (24)$$

Here the value of $\eta_j, j=1, \ldots, n$ depends on h, δ and $R(t)$, and generally is not known. In the absence of appropriate information about the function $R(t)$, we choose a value of 0.5 for η_j to reduce the largest possible error (minimax principle). So, for $\eta_j = 0.5$ the largest possible error in estimation of the integral in (24) is given by the factor $\exp(0.5h/\tau_k)$. For any other choice of the value of η_j (independent on j) the corresponding factor could be larger. This approach is more conservative than the above mentioned approaches because it makes no assumption regarding the shape on the integrand. In particular, it does not require linear approximation of $R(t)$ in the interval $[hi-h, hi]$ and subsequent replacement of $R(hi)$ by $\langle R_i \rangle$ as is required in the approach of McKinnon et al. (1977). The same replacement is presumably also used in the Simpson integration of Isenberg (1983).

Now, the expression (22) with (24) for $\eta_j = 0.5$ becomes

$$g_i(\boldsymbol{\theta}) = \sum_{k=1}^{N} f_k [\alpha_k \sum_{j=1}^{i-1} e^{h(j-i)/\tau_k} R_j^{\delta} + \beta_k R_i^{\delta} + \varepsilon_k'] + b_i , \quad (25)$$

where

$$\alpha_k = 2 \sinh\left(\frac{h}{2\tau_k}\right), \; \beta_k = 1 - e^{h/(2\tau_k)} . \quad (26)$$

and ε_k' is the approximation error due to choice of η_j as 0.5. For small h/τ_k, $\alpha_k \approx h/\tau_k$, $\beta_k \approx h/(2\tau_k)$. Formally, the expression (25) is equal to the standard formula of Grinvald and Steinberg (1974) when $\delta=0$, $\varepsilon_k'=0$, $\alpha_k = h/\tau_k$ and $\beta_k = h/(2\tau_k)$. This means that Grinvald-Steinberg discretization scheme and our approach should yield numerically close results when the smallest τ_k is a few orders of magnitude larger than channel width h.

In order to obtain an expression for $g_i(\boldsymbol{\theta})$ in terms of the measured counts, R_i, when δ is assumed to be different from zero, the following expressions are derived directly from definitions of R_i^{δ} and $\langle R_i \rangle$:

$$R_i^{\delta} = Q(ih+\delta) - Q(ih-h+\delta), \; Q(x) = \int R(x)\, dx,$$
$$Q(ih) - Q(ih-h) = \langle R_i \rangle . \quad (27)$$

The zero-time shift δ always can be written in the form

$$\delta = lh + \varepsilon, \; -h/2 \leq \varepsilon \leq h/2 , \quad (28)$$

where l is an integer and R_i^{δ} can be expressed as a difference of functions of the form $Q(x+\varepsilon)$. Then, by using the Taylor expansion of $Q(x+\varepsilon)$ with respect to ε and

$$R(x) = dQ(x)/dx = [Q(x+h) - Q(x)]/h + O(h) \quad (29)$$

(27) and (5) lead to

$$R_i^{\delta} = P_i + v_{i+l} + (v_{i+l+1} - v_{i+l})\, \varepsilon/h + \varepsilon O(h) + O(\varepsilon^2) ,$$
$$P_i = R_{i+l} + \varepsilon (R_{i+l+1} - R_{i+l})/h \quad (30)$$

Now, $R_j^{\delta}, j=1, \ldots, i$ in (25) can be replaced by P_j yielding an expression prepared for actual numerical evaluation. This replacement formally requires ε_k' to be replaced by a certain ε_k which is a combined approximation error arising from *a*) noise v_i, i.e. replacement of $\langle R_i \rangle$ by R_i, *b*) the choice of η_i as 0.5, and *c*) the error, of the order $O(h^2)$, in approximating R_i^{δ}. A computationally convenient recursive form of (25) analogue to that of Grinvald and Steinberg (1974), is derived in the Appendix.

In the following we compare the performance of the proposed discretization scheme with those of Grinvald and Steinberg (GS) and McKinnon, Szabo and Miller (MSM), by the use of simulations based on the IRF given

Table 1. Comparison of discretization schemes[a]

τ_1[b]	E_{med} GS	E_{med} MSM	E_{med}	M[c]
0.05	0.073	0.018	0.038	1
0.2	0.017	0.015	0.010	0.4
0.6	0.020	0.030	0.017	0.03
1.0	0.054	0.068	0.049	0.002
1.5	0.264	0.336	0.263	0.00006

[a] GS – scheme of Grinvald and Steinberg (1974); MSM – scheme of McKinnon et al. (1977); the approach presented here is unlabeled
[b] τ_1 is in nanoseconds; $\tau_2 = 2$ ns, $f_1 = 0.1$, $f_2 = 0.9$
[c] M is normalized to the largest value

by: $R(t') = at'^2 e^{-0.4t'}$ (t' in channels; O'Connor and Phillips 1984). The parameter a was chosen to yield 20 000 counts at the peak of the IRF and the channel width was $h = 0.05$ ns yielding FWHM = 0.12 ns. The background in fluorescence intensity decay functions was $b_i = 2$ counts and we assumed no zero-time shift. The data for the fluorescence intensity decay were simulated by use of (3), (4) and (6). With the given analytical form of $R(t)$, the respective integrations were performed analytically to obtain $g_i(\boldsymbol{\theta})$. The latter was then contaminated with the Poisson noise. Similarly, the data, R_i, for IRF were obtained from (5) and subsequently Poisson noise was added. The simulated data were analyzed by the ML method and the accuracy in the estimated parameters was measured by the median E_{med} of the relative errors E (17) for 101 simulated data sets. The results shown in Table 1 indicate that the proposed discretization scheme is a good compromise yielding equal or better accuracy than either the GS or the MSM approach, except for $\tau_1 = h$. In this case our approach performes better than the GS scheme, but not as good as MSM scheme. As seen from Table 1 the latter becomes less accurate as the information about parameters, M, decreases. Note that the decrease of information about parameters is well correlated with the decrease of the lifetime ratio τ_2/τ_1, as one would expect.

3 Results and discussion

Extending our preliminary work (Bajzer et al. 1990b), we have here tested the maximum likelihood method for analysis of time-resolved fluorescence decay curves on synthetic data. We generated data for various lifetimes and fractions (2, 3 and 4 components) using (25) for the number of counts with various measured instrument response functions standardly defined on 512 channels. The number of counts is normalized so that the counts in the peak channel (CPC) are equal to the counts in the peak channel of the measured instrument response function. According to obervations in measurements on our single photon counting system, the background level was set to be 0.01% of CPC (if not otherwise stated). For each set of lifetimes, fractions and instrument response function, 101 synthetic data sets were generated corresponding to different realizations of the Poisson noise (a noise generator from Press et al. (1986) was used). As in Sect. 2.3 the overall accuracy of lifetimes and fractions was characterized by the median, E_{med}, of the average relative errors E_i, $i = 1, \ldots, 101$ (17) corresponding to 101 synthetic data sets. The least-square methods were tested on the same synthetic data. We found that for two-component decays the simplex algorithm was robust with respect to the starting parameter values (chosen randomly) and the parameter estimates obtained by this algorithm were equal to those obtained by the modified Levenberg-Marquardt algorithm. However, for the three-component decays we observed that the simplex minimization was less robust and was worse for the four-component decays. This finding agrees with the observation of Olsson and Nelson (1975) that the simplex algorithm is most effective for less than six parameters.

Several factors influence the accuracy of the estimation of lifetimes and fractions. Besides the obvious signal to noise ratio (characterized by the CPC), and the ratios of lifetimes and fractions, the "tail" of the fluorescence intensity decay curve is also important. This latter factor of influence we characterized by the ratio r_t and the average number of counts, m_c, in the last 30 channels of the decay curve:

$$r_t = t_{max}/\tau_{max}, \quad m_c = \sum_{i=1}^{30} c_{n-30+i}/30 . \tag{31}$$

Here t_{max} is the maximal time at which the counts are recorded (number of channels × channel width) and τ_{max} is the largest lifetime involved. The results shown in Table 2 clearly demonstrate that *the longer the tail* (larger r_t and channel width h) *the more accurate are the lifetime and fraction estimates* (smaller E_{med}) for all of the three methods we compared: ML, LS and LSM. If m_c is fairly large (> 100) all of these methods are equally accurate as expected from our theoretical considerations. However, in the case of long tail with small number of counts ($m_c = 10$) *the ML estimates are significantly more accurate than those obtained from application of the LS and the LSM techniques.* The error E (ML) was smaller than E(LS) in 86% of the time (in 101 simulations) and smaller 63% of the time than E(LSM). In both cases this difference is statistically significant: by the pair test for E the probability that the result of 63% or more in comparison with 50% is observed by chance is equal or smaller than 0.01. The significance of difference in accuracy of the compared methods is also confirmed by the confidence intervals for E_{med}.

The question may be posed as to whether the effects of the discretization error and Poisson noise in the IRF could disturb the presented accuracy in such a way that the ML and the LSM estimates would then appear equally accurate. To answer this question we generated synthetic decay data with Poisson noise ($\tau_1 = 1$ ns $\tau_2 = 1.5$ ns $f_1 = f_2 = 0.5$, $h = 0.025$ ns, CPC = 20 000, $r_t = 7.9$) using the synthetic IRF (without noise) given in Sect. 2.3. The data were then analyzed assuming the discretized form of the decay function (25) with the same synthetic IRF, but now contaminated by Poisson noise. The results obtained were: E_{med} (ML) = 0.041 and E_{med} (LSM) = 0.066 with 95% confidence intervals [0.034, 0.052] and [0.050, 0.085] respectively. In 68% of time E(ML) was smaller than E(LSM) and this is statistically significant. Clearly, E_{med}

Table 2. The accuracy of the ML, LS and LSM estimates for selected channel widths and lifetime ratios

IRF[a] h/FWHM	Lifetimes and tail[b]			E_{med}			Information indices		
	τ_2	m_c	r_t	ML	LS	LSM	s_{12}	d_1	M
0.010/0.096	1.5	910	3.1	0.13	0.13	0.13	3.6	2.8	12
0.025/0.130	1.5	10	7.9	0.032[c]	0.106	0.049	15	11	178
0.010/0.096	3.0	1899	1.5	0.02	0.02	0.02	16	26	61
0.025/0.130	3.0	169	3.9	0.005	0.005	0.005	117	113	302

[a] Two measured IRF are characterized by the number of counts in the peak channel being 20 000, by the channel width h expressed in nanoseconds and by the full width at the half of maximum (FWHM) expressed in nanoseconds

[b] $\tau_1 = 1$ ns; τ_2 is expressed in nanoseconds. The fractions were $f_1 = f_2 = 0.5$

[c] The 95% confidence intervals for the medians are: ML: [0.024, 0.042]; LS: [0.093, 0.119]; LSM: [0.040, 0.055]

has increased for both methods due to the effects of the discretization approximation and the noise in the IRF (cf. second row in Table 2), but the accuracy of the ML method is still significantly higher. To eliminate any possibility that this result is an artifact, we have also tested the robustness of the used Levenberg-Marquardt minimizer. Twenty one randomly chosen starting parameter values ($f_1, f_2 \in [0, 1]$; $\tau_1, \tau_2 \in [0.1, 10]$) were used to analyze 5 realizations of the synthetic decay data. The resulting parameter estimates were consistently equal to the third decimal place.

The lower bounds of the index of separability and the index of detectability (Table 2) basically reflect the behavior of the E_{med} for the ML method. Along the same lines, the increase of M generally reflects the increase of accuracy although a deviation from such behavior is noted in Table 2.

With the above result we have clearly demonstrated the advantage of the ML method. This advantage is further illustrated by numerous two-component simulations performed to characterize the behavior of the ML and the standard LS methods under various circumstances and from different aspects, to wit:

– *The shape of the distributions of error E.* For very close values of lifetimes this distribution is broader for the ML method than is the corresponding distribution found for the LS method. At the same time, it is shifted towards smaller values with respect to the distribution for the LS method. A typical example is shown in Fig. 2a. For higher lifetime ratios the distribution of E for the ML method is also clearly shifted, but appears somewhat narrower than that given by the LS method (Fig. 2b).

– *The behavior of the E_{med} for increasing lifetime ratios, channel widths and r_t.* The curve corresponding to a two-component decay in Fig. 1 suggests that the critical lifetime ratio under specified conditions is approximately 1.1. Additionally, we performed simulations with $\tau_1 = 1$ ns, $\tau_2 = 1.11, 1.12, 1.13$ ns, $f_1 = f_2 = 0.5$, $h = 0.025$ ns, 40 000 CPC and a constant background level of 4 counts. We found that the median of Δ_2 for $\tau_2 = 1.11$ was 7.2 while for $\tau_2 = 1.12$ it was 11.4, i.e. just above the threshold of 9.2

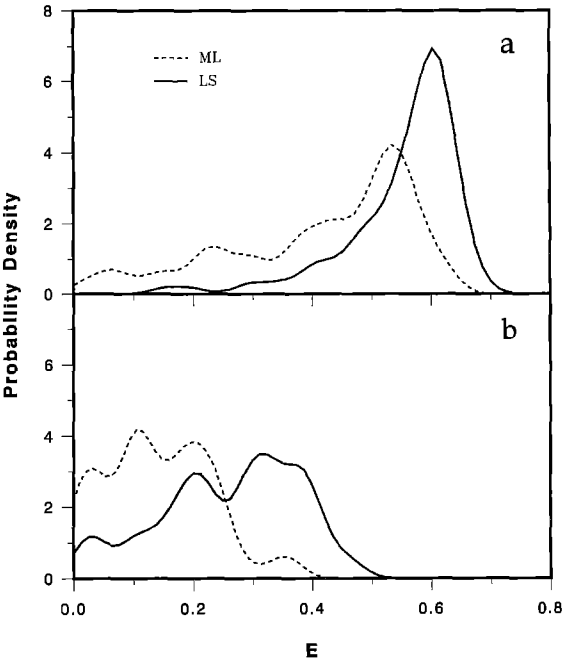

Fig. 2 a, b. The distributions of the average relative errors E for the ML and LS methods. The distributions are based on the analysis of 101 samples of two-component decay curves generated by use of measured IRF (512 channels, CPC = 40 000, $h = 0.025$ ns, and the full width half maximum was 0.09 ns) and parameters: $\tau_1 = 1$ ns, $f_1 = f_2 = 0.5$, $\delta = 0$; **a** $\tau_2 = 1.16$ ns, **b** $\tau_2 = 1.3$ ns

for Δ_2 (cf. Sect. 2.2). Thus, we can estimate the critical lifetime ratio to be 1.12 for the above conditions. At this lifetime ratio two components are almost unresolved as is seen from Table 3. The value of s_{12} is lower than one, in accordance with considerable departure of the estimated lifetimes from the trues ones. Similarly, the upper bound on the index of detectability, d_1, is much smaller than one, indicating substantially inaccurate estiamtes of fractions. For a lifetime ratio of 1.2 the estimates of parameters for the ML method (Table 3) are much closer to the true values, although the fractions are still considerably biased towards unequal values. The results of the LS method at $\tau_2/\tau_1 = 1.2$ are almost as inaccurate as the results of the Ml method for a lifetime ratio of 1.12 (Table 3).

If the lifetime ratio is increased (by increasing τ_2 while t_{max} is kept constant), both methods provide increasingly more accurate estimates (Fig. 3), to the point where for all practical purposes both methods provide satisfactorily accurate estimates (say at lifetime ratio of 1.8 when $E_{med} < 0.02$ for both methods). The effect of increasing the channel width, h, while keeping r_t constant is also shown in Fig. 3. By comparing the dashed and solid lines one can see that the ML method is less sensitive than the LS analysis to changes in h. Again, this can be attributed to the fact that in the ML method Poisson statistics are adequately taken into account even for the low number of counts at the tail of the decay curve. In Fig. 4 we show that the lifetimes estimated by the LS method are evidently biased towards lower values while for the ML method this is less emphasized.

The importance of the tail of the fluorescence decay curve is more comprehensively illustrated in Fig. 5. It is

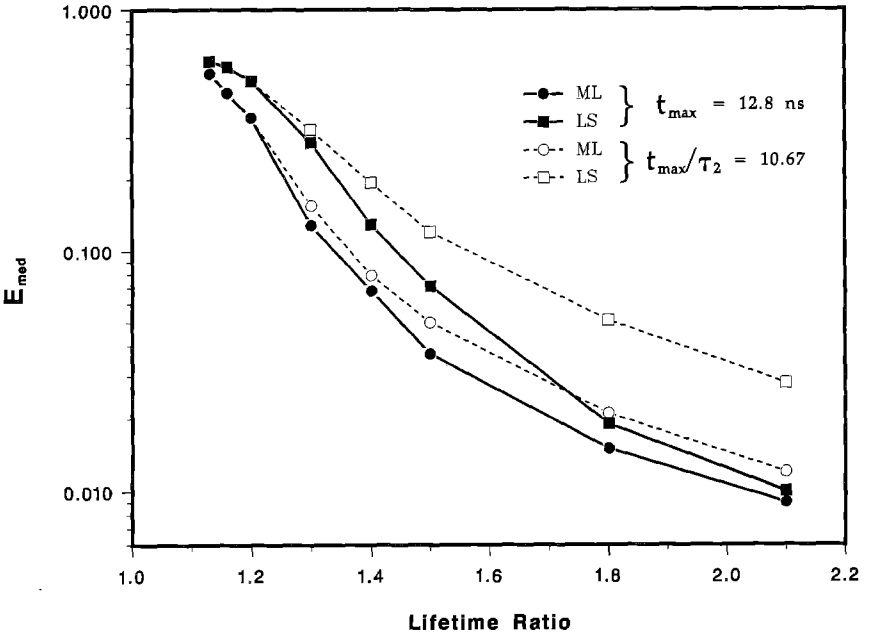

Fig. 3. The medians of the average relative errors E for the ML and LS methods as a function of lifetime ratio τ_2/τ_1 for the two-component decay. Solid lines correspond to the case where t_{max} is constant ($h=0.025$ ns), and dashed lines correspond to the case where $r_t=t_{max}/\tau_2$ is constant (variable h). The IRF was the same as in Fig. 2, except for the variation in h. The other parameters were: $\tau_1=1$ ns, $f_1=f_2=0.5$, $\delta=0$

Table 3. Recovered lifetimes and fractions of selected two-component decay simulations

	τ_1	τ_2	f_1	f_2	R.D.[a]	E_{med}	s_{12}	d_1	d_2
Exact	1.00	1.12	0.50	0.50	1.20		0.52	0.37	0.37
ML[b]	0.7	1.066	0.02	0.98	1.02	0.55			
S.D.[c]	0.2	0.007	0.04	0.04					
Exact	1.00	1.20	0.50	0.50	1.38		1.50	1.50	2.11
ML	0.9	1.14	0.2	0.8	1.16	0.36			
S.D.	0.1	0.03	0.2	0.2					
LS	0.76	1.122	0.07	0.93	1.12	0.51			
S.D.	0.09	0.008	0.04	0.04					
Exact	1.00	1.80	0.009	0.991	1.13		1.65	0.69	75.5
ML	0.4	1.797	0.003	0.997	1.03	0.32			
S.D.	0.2	0.002	0.002	0.002					
Exact	1.00	1.80	0.04	0.96	1.06		7.3	3.04	73.0
ML	0.9	1.794	0.030	0.970	1.02	0.09			
S.D.	0.1	0.005	0.009	0.009					
LS	0.88	1.792	0.030	0.971	1.11	0.10			
S.D.	0.08	0.004	0.007	0.007					

[a] R.D. denotes the reduced deviance given by $D(\theta)/\nu$ in case of the ML method, and by χ^2/ν in case of the LS method and for the case of "exact" parameters

[b] Displayed parameters for a given method correspond to that simulation (among 101 with the given "exact" parameters) for which the average relative error achieves its median value

[c] The standard deviations (S.D.) in parameters are estimated from the inverse Hessian of the Poisson deviance for the ML method (see 19), and from the inverse Hessian of χ^2 (8) for the LS method

clear that there is an optimal value of r_t (≈ 6.6 at given lifetime ratio of 1.3) for which the error E_{med} for both ML and LS estimates, is the lowest. This is consistent with the result of Hall and Selinger (1981), i.e. that for a one-component decay there is an optimal channel width. For higher values of r_t the accuracy of the ML estimator decreases significantly less than the accuracy of the LS estimator. Again, this can be understood as the impact of the low-count region of the fluorescence decay curve on the accuracy of estimation. Below the optimal r_t the error E increases rapidly for both methods. In Fig. 5 we have also displayed the indices s_{12} and d_1 which predict the behavior of E_{med} in the case of the ML method (the greater these indices are the smaller is the error). In practice one can use these indicates to assess the best values for r_t; in other words, if we assume a certain value for τ_{max} (obtained in preliminary analysis) we can calculate s_{12} and d_1 for different values of t_{max}.

– *The behavior of the E_{med} for various fraction ratios.* In Fig. 6 we illustrate how both the ML and LS methods become less accurate as the fraction ratio f_2/f_1 departs from 1. For the ML method E_{med} is noticeably symmetric with respect to $f_2/f_1=1$. This is not the case for the LS method where noticeable asymmetry can be expected because of the longer tail of the fluorescence intensity decay curve for higher f_1 (as τ_1 was smaller than τ_2), which magnifies the shortcomings of the LS method at the low-count region of the fluorescence intensity decay function. The upper bounds of detectability indices d_1 and d_2, also shown in Fig. 6, behave in agreement with E_{med}. The critical fraction ratio can be assessed in the same way as the critical lifetime ratio, i.e. by using the likelihood ratio test. For $\tau_2/\tau_1=1.8$, $r_t=7.1$ and 40 000 CPC, we found that the critical fraction ratio is approximately 110 ($f_1=0.009$). For this ratio $\Delta_2=10.9$, but the actual estimates of τ_1 and f_1 are considerably biased (Table 3). This is well reflected by the values of s_{12} and d_1. For a fraction ratio of 24 ($f_1=0.04$), the parameter estimates are considerably more accurate (Table 3) for both the ML and LS methods.

– *The dependence of E_{med} on the signal to noise ratio.* As mentioned above we characterize the signal to noise ratio by the number of counts in the peak channel of the fluorescence intensity decay profile. In Table 4 we show how E_{med} decrease with an increase in the CPC (three different measured IRF were used) and we compare four methods: ML, LS, LSM and the generalized Padé-

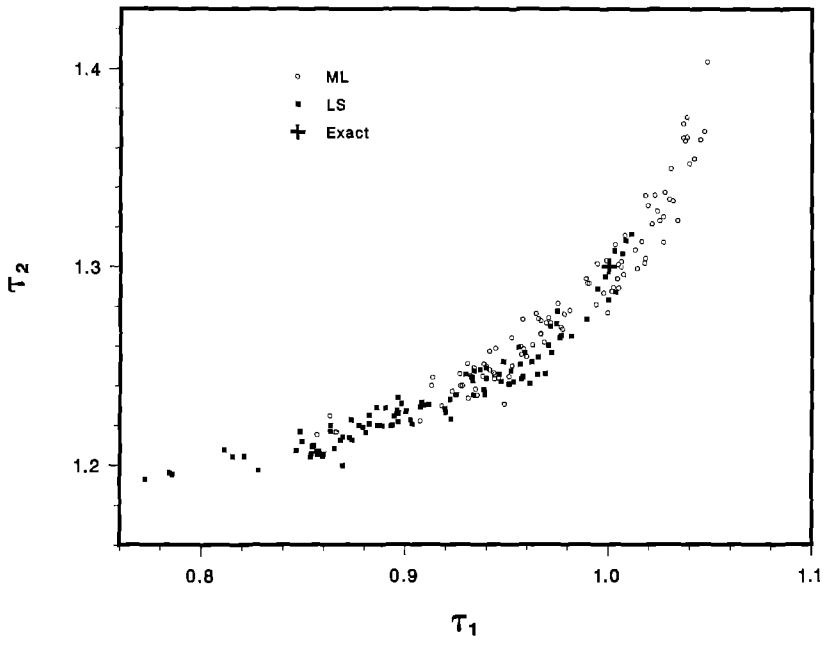

Fig. 4. The distribution of lifetimes in τ_2 versus τ_1 scattergram. The simulated two-component decays were generated with IRF and parameters as in Fig. 2, except for the lifetimes being here: $\tau_1 = 1$ ns, $\tau_2 = 1.3$ ns

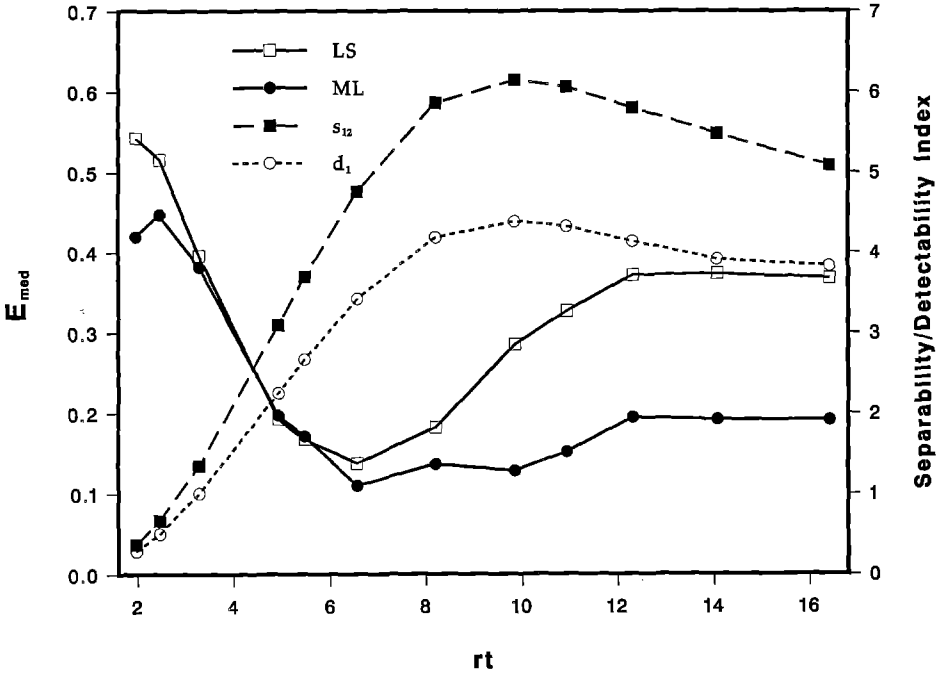

Fig. 5. The medians of the average relative errors E for the ML and LS methods as a function of r_t, and the corresponding lower bounds of separability and detectability indices (right scale) in two-component decays. The IRF was the same as in Fig. 2. The value of r_t was varied by changing τ_2 and τ_1 while keeping their ratio constant at 1.3. Other parameters required for simulation were: $f_1 = f_2 = 0.5$, $\delta = 0$

Laplace method. The simulations were defined by $\tau_1 = 1$ ns, $\tau_2 = 1.3$ ns, $f_1 = f_2 = 0.5$, $h = 25$ ps. The lower bounds of the index of separability and the index of detectability correlate well with E_{med}. As expected the amount of information about the parameters, M, increases with the CPC. The comparison of the four approaches (based on E_{med} with 95% confidence limits and on the percentage p defined in Table 4) shows that the ML method provides the most accurate and the standard LS method the least accurate results (except for the highest CPC, where, interestingly, LS performs better than GPL).

All presented simulations of two-component decays demonstrate just how complex is the problem of determining lifetimes and corresponding fractions. This complexity increases markedly for three- or four-component decays. Now, we present simulations which illustrate some of the features pertinent to three- and four-component decays.

– *The behavior of E_{med} for increasing lifetime ratios in the case of three-component decays.* First we consider the case where the lifetime ratio is given by $\tau_2/\tau_1 = \tau_3/\tau_2$. The result of simulations shown in Fig. 7 (curves a) reveals

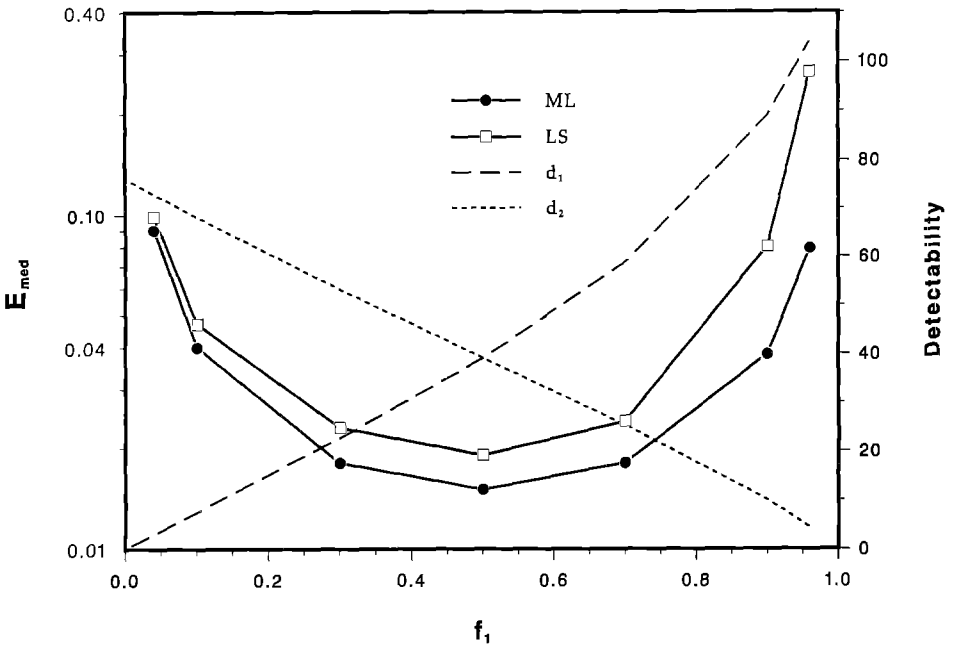

Fig. 6. The median of average relative error E as a function of fraction f_1 in two-component decays for the ML and LS methods. The upper bound of detectability indices are also shown as a function of f_1 (right scale). The parameters were: $\tau_1 = 1$ ns, $\tau_2 = 1.8$, $f_2 = 1 - f_1$, $\delta = 0$. The IRF is the same as in Fig. 2

Table 4. The accuracy of the ML, LS, LSM and GPL estimates for increasing signal to noise ratio

	ML	LS		LSM		GPL		Information indices		
CPC[a]	E_{med} C.I.[b]	E_{med} C.I.	p[c]	E_{med} C.I.	p	E_{med} C.I.	p	s_{12}	d_1	M
10	0.25	0.50	87	0.42	74	0.23	53	3.1	2.2	9
	0.21	0.47		0.38		0.18				
	0.30	0.52		0.48		0.29				
40	0.12	0.19	70	0.16	65	0.14	52	6.2	4.4	36
	0.10	0.16		0.14		0.11				
	0.15	0.23		0.19		0.17				
200	0.054	0.056	59	0.060	60	0.089	69	14	10	178
	0.041	0.046		0.042		0.071				
	0.063	0.076		0.068		0.097				

[a] CPC is the number of counts in the peak channel given in thousands

[b] C.I. is the 95% confidence interval for the corresponding medians

[c] p is the percentage expressing how many times in 101 simulations E for the ML method is smaller than E for other method corresponding to the label in the first row of the table

that a ratio of 2 and higher the ML and LS methods coincide in accuracy (see also Table 5). For the ratio of 1.8, E(ML) is smaller than E(LS) in 65% times showing the significant difference in accuracy. The critical lifetime ratio was estimated to be approximately 1.5 (see Fig. 2) with the median of Δ_3 being 10 (see Table 5 for the representative estimates of lifetimes and fractions). This value of the critical lifetime ratio can be expected on the basis of greater complexity of the three-exponential compared to the two-exponential model. However, in some cases two of the three components with the lifetime ratio smaller than 1.5 *can* be resolved. Thus, if we choose $\tau_1 \leq 1$ ns, $\tau_2 = 1.25$ ns, $\tau_3 = 5$ ns $f_1 = f_2 = f_3 = 1/3$, and define the lifetime ratio to be τ_2/τ_1, the critical lifetime ratio was estimated to be approximately 1.25 with the median of Δ_3 being 10.3. The results of simulations for increasing τ_2/τ_1 are also shown in Fig. 7 (curves b) and typical estimates of parameters at the critical lifetime ratio are displayed in Table 5. The values of upper bounds on indices of separability and detectability generally reflect the accuracy of estimated parameters; examples are shown in Table 5.

– *The behavior of E_{med} for increasing lifetime ratios in the case of four-component decays.* We assume that the lifetime ratio is given by $\tau_2/\tau_1 = \tau_3/\tau_2 = \tau_4/\tau_3$. Simulations showed that the ML and LS methods display rather similar accuracy, except for the lifetime ratio of 2 (see Fig. 8). The critical lifetime ratio under the conditions defined in the figure caption of Fig. 8 was estimated to be 2.05 with the median of Δ_4 being 10.8. Typical estimates of parameters at the critical lifetime ratio as well as for the ratio 2.5 are shown in Table 5. The upper bounds on indices of separability and detectability are also shown. As in the two- and three-component cases they reflect the accuracy of the estimated parameters.

In all of the results presented thus far, we have assumed that the zero-time shift is negligible ($\delta = 0$). If δ is of the order of a few channels, our simulations show that the accuracy of the estimated lifetimes and fractions does not differ from the case when $\delta = 0$. However, we noticed that it was important to start the minimization with an initial value of δ not far from the true one. Therefore it appears useful for the analysis of experimental data to obtain the initial value of parameters by use of a method which does not involve the numerical minimization of non-linear functions (e.g. the method of moments, the GPL method, Laplace transform method).

Concluding this section, we want to emphasize that in all the cases we have studied here, the common criteria of

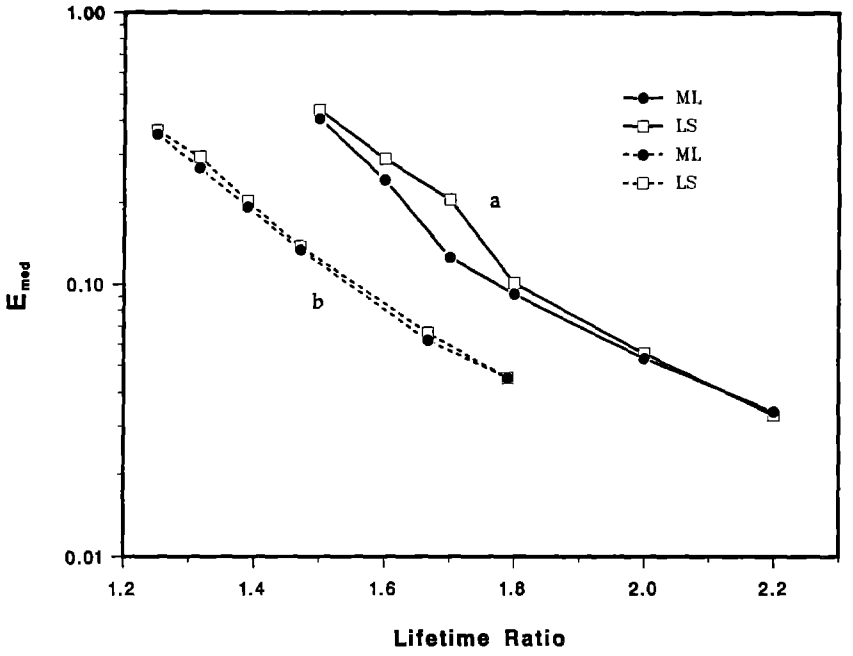

Fig. 7. The medians of the average relative errors E for the ML and LS methods as a function of lifetime ratio for the three-component decay. The curves (a) correspond to lifetime ratio defined as $\tau_2/\tau_1=\tau_3/\tau_2, \tau_1=0.8$ ns, and the same IRF as in Fig. 2 were used. The curves (b) correspond to lifetime ratio defined as τ_2/τ_1, $\tau_2=1.25$ ns, $\tau_3=5$ ns and the IRF used was the one described in Fig. 2 with $h=0.04$ ns. In both cases the fractions were $f_1=0.334$, $f_2=f_3=0.333$, and $\delta=0$

Table 5. Recovered lifetimes and fractions of selected simulations with 3 and 4 decay components

Index	τ_1 s_{12}	τ_2 s_{23}	τ_3 s_{34}	τ_4	f_1 d_1	f_2 d_2	f_3 d_3	f_4 d_4	R.D.[a]	E_{med}
Exact	0.50	0.75	1.125		0.334	0.333	0.333		1.15	
Index	0.8	1.2			0.9	31	12			
ML[b]	0.3	0.60	1.1		0.03	0.51	0.40		1.05	0.41
S.D.[c]	0.1	0.04	0.02		0.04	0.02	0.04			
LS	0.3	0.56	1.06		0.01	0.52	0.47		1.01	0.43
S.D.	0.2	0.02	0.01		0.02	0.01	0.02			
Exact	0.50	1.0	2.0		0.334	0.333	0.333		1.12	
Index	8.9	10.8			8.9	18	12			
ML	0.49	0.92	1.94		0.30	0.33	0.37		1.05	0.053
S.D.	0.02	0.08	0.03		0.05	0.03	0.02			
LS	0.51	1.10	2.04		0.37	0.33	0.30		1.06	0.056
S.D.	0.01	0.09	0.06		0.03	0.02	0.04			
Exact	1.0	1.25	5.0		0.334	0.333	0.333		1.14	
Index	0.5	9.0			0.38	0.38	58			
ML	1.07	1.8	5.1		0.60	0.08	0.32		1.06	0.36
S.D.	0.04	0.9	0.1		0.09	0.08	0.02			
LS	0.7	1.14	4.97		0.03	0.63	0.335		1.04	0.37
S.D.	0.4	0.04	0.03		0.08	0.08	0.004			
Exact	0.2	0.41	0.84	1.72	0.25	0.25	0.25	0.25	1.17	
Index	1.6	1.7	4.5		3.2	4.1	2.8	7.0		
ML	0.19	0.3	0.72	1.69	0.2	0.2	0.32	0.27	1.01	0.18
S.D.	0.04	0.1	0.06	0.03	0.2	0.2	0.04	0.01		
LS	0.18	0.26	0.66	1.63	0.1	0.3	0.33	0.29	1.13	0.22
S.D.	0.08	0.09	0.04	0.02	0.3	0.3	0.03	0.01		
Exact	0.2	0.5	1.25	3.125	0.25	0.25	0.25	0.25	1.02	
Index	4.4	4.1	8.8		9.9	9.6	10	9.8		
ML	0.198	0.45	1.1	2.99	0.24	0.22	0.26	0.28	0.99	0.078
S.D.	0.009	0.07	0.1	0.06	0.03	0.02	0.03	0.01		
LS	0.210	0.5	1.1	3.02	0.27	0.20	0.26	0.27	1.03	0.080
S.D.	0.009	0.1	0.2	0.09	0.04	0.04	0.05	0.02		

[a] R.D. denotes the reduced deviance defined as in Table 3

[b] Displayed parameters for a given method correspond to that simulation (among 101 with the given "exact" parameters) for which the average relative error achieve its median value

[c] The standard deviations (S.D.) in parameters are estimated as in Table 3

goodness of fit (reduced deviance close to one, random residuals and Durbin-Watson factor higher than 1.8) were ideally satisfied. This implies that under given conditions the accuracy of estimated parameters is that which is maximally achievable. It is likely that in the case of measured data the analysis by the ML (or LS) method yields less accurate estimates due to possible nonrandom errors (e.g. light scatter (Small et al. 1990)). In order to improve the accuracy, the expected nonrandom errors have to be mathematically modeled.

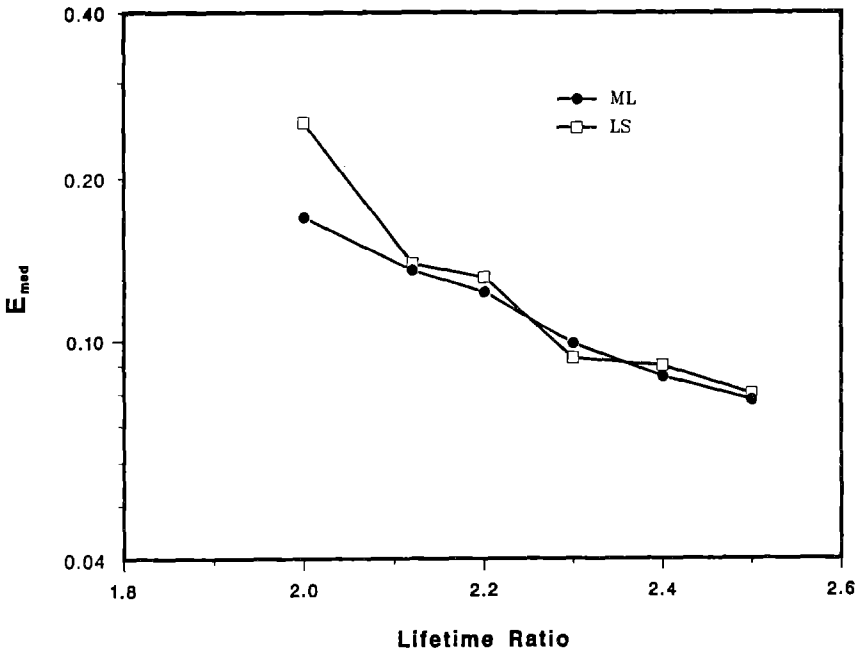

Fig. 8. The medians of the average relative errors E for the ML and LS methods as a function of lifetime ratio for the four-component decay. The IRF was the same as in Fig. 2. The parameters were: $\tau_1 = 0.2$ ns, $\tau_2, \tau_3, \tau_4 =$ variable, $f_1 = f_2 = f_3 = f_4 = 0.25$, $\delta = 0$ ns

4 Conclusions

We have demonstrated the applicability of the maximum likelihood method to the analysis of time-resolved fluoresecence decay curves. The ML method is clearly as applicable as the standard LS method. However, the ML method possesses a distinct advantage over the standard and modified LS approach since it correctly takes into account the Poisson fluctuation characteristic of single photon counting data. From the standpoint of computational demand all three considered methods are essentially equivalent.

We have shown how the properties of the maximum likelihood estimator can be used to obtain information a priori on the separability and the detectability of decay components. Specifically, we made use of the information matrix to define bounds on separability and detectability indices. These indices have not hitherto been applied to the analysis of single photon counting data. We have also introduced the likelihood ratio test as a tool for discriminating models with N and $N-1$ decay components.

To assess the ML method and to compare it to the LS method we performed a large number of simulations based on measured and synthetic instrument response functions. The main conclusions from these simulations are:

i) The ML estimates are always at least as good as those from the LS approach and, in most cases, are better.

ii) The accuracy with which the lifetimes and fractions are estimated behave similarly for the ML and LS methods as lifetime or fraction ratios are changed. The same is true when the level of the signal to noise ratio changes.

iii) The tail of the fluorescence intensity decay curves may carry considerable information on fractions and lifetimes. The ML method is more suitable than the standard or modified LS method for extracting this information (i.e., the ML method provides more accurate estimates).

iv) The lower bounds of separability and detectability indices are useful for assessing how accurately, and whether under given conditions (of signal to noise ratio, number of channels, channel width and IRF) one can resolve two components with close lifetimes. The lower bound of detectability index is specifically useful to determine, whether and how accurately one can detect a component with a small fraction.

v) The likelihood ratio test is useful in deciding whether an additional decay component is needed for "best fit". If the test implies values just above the threshold then the corresponding parameter estimates are strongly biased.

The experience we have gained from numerous simulations allow us to make some recommendations on how to approach the analysis of fluorescence intensity decay curves *assuming that they are well described by the multi-exponential model.* First, it is useful to estimate the number of exponential decay components and the corresponding parameters by using the GPL method (or other similar technique such as the method of moments). In this way one has an excellent starting point for subsequent application of the ML method, namely the expected number of components and the values of parameters for the initial guess. Then, if analysis by the ML method is performed, and if there is evidence (based on the likelihood ratio test) that two or more decay components are present, we recommend calculation of the lower bounds to separability and detectability indices for the lifetimes and fractions obtained. These calculations require a small computational effort and provide information on how resolvable and detectable the assumed components are. (If any of these indices is smaller than 2 the parameter estimates may be very inaccurate.) The next step is to do simulations based on the preliminary obtained parameters and on a *measured* instrument response function. These are computationally much more intensive, but they

should nonetheless be performed. Then, by using E_{med}, one should be able to assess the global accuracy achieved for this set of parameters. The same simulations can be used to estimate the variances for particular lifetimes or fractions and can thus provide further insight into how accurate the estimated parameters might be.

If, by this additional analysis, it becomes clear that the estimation of parameters is not satisfactorily accurate, one should consider applying another method of analysis. However, there is little hope that any general method can provide more accurate estimates of parameters, and, even if another method could do so, simulations would still be essential. The best approach may therefore be to proceed with simulations of fluorescence intensity decay data containg a greater amount of information about the desired parameters. The amount of information can be enhanced by increasing the CPC and the number of channels, and/or by adding decay curves as done for global analysis (Knutson et al. 1983). The ML analysis of these data should then yield more accurate estimates of parameters. (We have implemented a version of our computer code for the ML method in the current version of the software GLOBALS UNLIMITED (Beechem et al. 1990) for global analysis.) Depending on the results of these simulations one should be able then design and perform additional measurements and by applying the ML technique obtain more accurate estimates of both lifetimes and fractions.

One final comment. The analysis of fluorescence intensity decay profiles belongs to those complex problems which involve the art of scientific computing. With the present state of knowledge this problem *cannot* be reduced to a precise algorithm. Rather, one needs to employ a roughly defined strategy based on experience and involving the use of several approaches. The ML method described above appears to be the most accurate if the assumed decay model is correct. To assess how adequate the multiexponential decay model is, one should use methods of analysis capable of *detecting* quasi-continuous distributions of lifetimes which may correspond to a nonexponential decay model. A comprehensive set of such methods have been studied recently (Siemiarczuk et al. 1990; Small et al. 1989; Livesey and Brochon 1987; Alcala et al. 1987; Willis et al. 1990; Provencher 1982a, b)[3], but with the exception of the extended method of moments (Small et al. 1989), they involve minimization of the *Gaussian* deviance for analysis of time correlated single photon counting data. It would now be desirable to modify these methods to include the more appropriate Poisson deviance and to conform with the maximum likelihood approach. We believe that such a modification would improve the resolvability and detectability of these methods.

[3] The software package called CONTIN designed by Provencher, is a very versatile tool for analysis of possible existence of lifetime distributions, based on powerful constrained regularization method for solving the integral equations with noisy functions. Unfortunately it has not been widely used by researchers in the field of fluorescence analysis. The recent paper of Gregory and Zhu (1990) clearly shows the remarkable capabilities of CONTIN

Acknowledgements. We thank Dr. J. R. Beechem who incorporated the computer program for our maximum likelihood method implementation into "GLOBALS UNLIMITED" for global analysis. We also thank K. D. Peters for his assistance with preparation of the figures, and M. C. Stacy who efficiently kept the computers and network running properly. It is our pleasure to acknowledge very valuable referee comments. Supported by grant GM 34847 of the USPHS.

5 Appendix

Here we present some relevant numerical and computational aspects of minimization of the Poisson deviance (7), especially in comparison to minimization of the Gaussian deviances (8) and (9). Any of these deviances, $d(\boldsymbol{\theta})$, can be written in the form

$$d(\boldsymbol{\theta}) = \sum_{i=1}^{n} f_i^2(\boldsymbol{\theta}) \tag{A1}$$

where respective forms of $f_i(\boldsymbol{\theta})$ are given in Table 6.

Frequently used Gauss-Newton method of minimization (or its modifications such as Levenberg-Marquardt algorithm) can be applied to functions of the form (A1) (cf. Kennedy and Gentle 1980). The Gauss-Newton method requires the first derivatives of the functions $f_i(\boldsymbol{\theta})$ which can be written as

$$\partial f_i(\boldsymbol{\theta})/\partial\theta_j = \phi_i(\boldsymbol{\theta})\, \partial g_i(\boldsymbol{\theta})/\partial\theta_j \tag{A2}$$

with $\phi_i(\boldsymbol{\theta})$ given in the Table 6. Computationally, $\phi_i(\boldsymbol{\theta})$ for the Poisson deviance is somewhat more complex than $\phi_i(\boldsymbol{\theta})$ for the deviance (9), while $\phi_i(\boldsymbol{\theta})$ for deviance (8) is clearly computationally the simplest. However, with todays fast computers there is no reason to choose any of the Gaussian deviances on the basis of their minor computational advantage.

Other powerful methods for numerical minimization like conjugate gradient method or variable metric methods (cf. Press et al. 1986) require the first derivatives of object functions itself. In our case they are given by

$$\partial d(\boldsymbol{\theta})/\partial\theta_j = 2\sum_{i=1}^{n} \psi_i(\boldsymbol{\theta})\, \partial g_i(\boldsymbol{\theta})/\partial\theta_j \tag{A3}$$

with $\psi_i(\boldsymbol{\theta})$ presented in Table 6. From Table 6 it is clear that the first derivatives of Poisson deviance are computationally equivalent to the first derivatives of Gaussian deviance (8) and the first derivatives of deviance (9) are

Table 6. Computational differences in minimization of Poisson and Gaussian deviances

$d(\boldsymbol{\theta})$	$f_i(\boldsymbol{\theta})$[a]	$\phi_i(\boldsymbol{\theta})$	$\psi_i(\boldsymbol{\theta})$
Poisson, Eq. (7)	$[2(c_i \log u_i + g_i - c_i)]^{1/2}$	$2 f_i^{-1}(1-u_i)$	$1-u_i$
Gaussian Eq. (8)	$(c_i - g_i)\, c_i^{-1/2}$	$-c_i^{-1/2}$	$u_i^{-1}-1$
Gaussian Eq. (9)	$(c_i - g_i)\, g_i^{-1/2}$	$(g_i)^{-1/2}(1+u_i)/2$	$1-u_i^2$

[a] $u_i = c_i/g_i(\boldsymbol{\theta})$, $g_i = g_i(\boldsymbol{\theta})$, $f_i = f_i(\boldsymbol{\theta})$

negligibly more complex. Thus, in case of methods which require the first derivatives of the object function there is no computational advantage in choosing any of the two considered Gaussian deviances. The same is true for the Newton method (cf. Kennedy and Gentle 1980) where the computation of Hessian matrix (19) is required. This can be easily seen after differentiating the expressions (A3) with respect to θ_k. However, the computation of Hessian in our case requires considerable amount of cpu-time and therefore we would not recommend the use of the Newton method.

Another computationally relevant problem is the evaluation of the model function $g_i(\boldsymbol{\theta})$ and its derivatives. The discretization scheme is described in Sect. 2.3 yielded to formula (25) for $g_i(\boldsymbol{\theta})$ which can be rewritten in computationally convenient recursive form as:

$$g_i(\boldsymbol{\theta}) = \sum_{k=1}^{m} (F_i^k + f_k \varepsilon_k') + b_i, \tag{A4}$$

$$F_i^k = f_k \left[\alpha_k \sum_{j=1}^{i-1} e^{h(j-i)/\tau_k} R_j^\delta + \beta_k R_i^\delta \right]$$

$$F_{i+1}^k = F_i^k e_k^2 + f_k \beta_k (R_i^\delta e_k + R_{i+1}^\delta); \; F_1^k = f_k \beta_1 R_1^\delta \tag{A5}$$

where $e_k^2 = -e^{-h/\tau_i}$, and α_k and β_k are given by (26). In actual calculation the error ε_k' introduced by discretization is neglected.

The first derivatives of $g_i(\boldsymbol{\theta})$ can also be expressed recursively. As seen from (A4) it is only necessary to find derivatives of F_i^k. From (A5) it follows immediately

$$\partial F_i^k / \partial f_k = F_i^k / f_k, \tag{A6}$$

$$\partial F_{i+1}^k / \partial \tau_k = e_k^2 \, \partial F_i^k / \partial \tau_k + h \, e_k [(F_i^k - f_k R_i^\delta) e_k + f_k (R_i^\delta - R_{i+1}^\delta)/2] / \tau_k^2, \tag{A7}$$

$$\partial F_1^k / \partial \tau_k = -h \, f_k \, e_k \, R_1^\delta / (2 \tau_k^2), \tag{A8}$$

$$\partial F_{i+1}^k / \partial \delta = e_k^2 \, \partial F_i^k / \partial \delta + f_k \beta_k (R_i' e_k + R_{i+1}'), \; R_i' = \partial R_i^\delta / \partial \delta . \tag{A9}$$

$$\partial F_1^k / \partial \delta = f_k \beta_k R_1' . \tag{A10}$$

When $\delta \equiv 0$, then the expressions (A5–A8) can be readily computed ($R_i^0 = \langle R_i \rangle$ is replaced by measured R_i). When δ is assumed to be different from zero, the expressions (A5–A8) can be computed by approximating R_i^δ with P_i (cf. 30). In order to compute R_i' we have to express it in terms of R_i. This can be accomplished by differentiation of R_i^δ with respect to δ (cf. 23), by subsequent substitution of (28) and by Taylor expansion with respect to ε yielding: $R_i' = R((i+l)h) - R((i+l-1)h) + O(\varepsilon)$. Then, analogously as in Sect. 2.3, (29), (27) and (5) are used to obtain:

$$R_i' = P_i' - (v_{i+l+1} - v_{i+l})/h + O(h) + O(\varepsilon), \tag{A11}$$

$$P_i' = (R_{i+l+1} - R_{i+l})/h$$

Now (A9) and (A10) can be approximated by replacing R_i' with P_i' and the obtained expression computed from the measured IRF data.

References

Alcala JR, Gratton E, Prendergast FG (1987) Interpretation of fluorescence decays in proteins using continuous lifetime distributions. Biophys J 51:925–936

Ameloot M, Hendrickx H (1983) Extension for the performance of Laplace deconvolution in the analysis of fluorescence decay curves. Biophys J 44:27–38

Ameloot M, Beechem JM, Brand L (1986) Simultaneous analysis of multiple fluorescence decay curves by Laplace transforms. Deconvolution with reference or excitation profiles. Biophys Chem 23:155–171

Andre JC, Vincent LM, O'Connor D, Ware WR (1979) Applications of fast fourier transform to deconvolution in single photon counting. J Phys Chem 83:2285–2293

Awaya T (1979) A new method for curve fitting to the data with low statistics not using the χ^2-method. Nucl Instrum Methods 165:317–323

Awaya T (1980) Application of a new method to the analysis of radioactive decays. Nucl Instrum Methods 174:237–242

Bajzer Ž, Sharp JC, Sedarous SS, Prendergast FG (1990a) Padé-Laplace method for the analysis of time resolved fluorescence decay curves. Eur Biophys J 18:101–115

Bajzer Ž, Sharp JC, Therneau TM, Prendergast FG (1990b) Comparison of the maximum likelihood and least squares methods for the analysis of time-resolved fluorescence decay curves. Biophys J 57:190a

Beechem JM, Gratton E (1988) Fluorescence spectroscopy data analysis environment: A second generation Global analysis program. In: Lakowicz RJ (ed) SPIE Proceedings: Time resolved laser Spectroscopy in Biochemistry 909:70–81

Beechem JM, Gratton E, Mantulin WW (1990) GLOBALS UNLIMITED – Technical Reference Manual. Butzow JK (ed) Laboratory for Fluorescence Dynamics, University at Illinois, Department of Physics, Urbana

Bronshtein IN, Semendyayev KA (1985) Handbook of mathematics. Van Nostrand, New York

Cox DR, Hinkley DV (1974) Theoretical statistics. Chapman and Hall, London

Demas JN, Adamson AW (1971) Evaluation of photoluminescence lifetimes. J Phys Chem 75:2463–2466

Eaton DF (1990) Recommended methods for fluorescence decay analysis. Pure Appl Chem 62:1631–1648

Eisenfeld J, Ford CC (1979) A systems-theory approach to the analysis of multiexponential fluorescence decay. Biophys J 26:73–84

Gafni A, Modlin RL, Brand L (1975) Analysis of fluorescence decay curves by means of the Laplace transformation. Biophys J 15:263–280

Good HP, Kallir AJ, Wild UP (1984) Comparison of fluorescence lifetime fitting techniques. J Phys Chem 88:5435–5441

Gregory RB, Zhu Y (1990) Analysis of positron annihilation lifetime data by numerical Laplace inversion with the program CONTIN. Nucl Instrum Methods Phys Res A 290:172–182

Grinvald A, Steinberg IZ (1974) On the analysis of fluorescence decay kinetics by the method of least squares. Anal Biochem 59:583–598

Hall P, Selinger BK (1981) Better estimates of exponential decay parameters. J Phys Chem 85:2941–2946

Hall P, Selinger BK (1984) Better estimates of multiexponential decay parameters. Z Phys Chem NF 141:77–89

Haydock C, Sedarous SS, Bajzer Ž, Prendergast FG (1990) Fluorescence of variant-3 scorpion neurotoxin: multiple decay components and multiple tryptophan conformers. In: Lakowicz RF (ed) SPIE Proceedings: Time resolved laser Spectroscopy in Biochemistry II 1024:92–99

Hinde AL, Selinger BK, Nott PR (1977) On the reliability of fluorescence decay data. Aus J Chem 30:2383–2394

Isenberg I (1983) Robust estimation in pulse-fluorometry, a study of the method of moments and least squares. Biophys J 43:141–148

Isenberg I, Dyson RD (1969) The analysis of fluorescence decay by a method of moments. Biophys J 9:1337–1350

Janssens LD, Boens N, Ameloot M, De Schryver FC (1990) A systematic study of the global analysis of multiexponential fluorescence decay surfaces using reference convolution. J Phys Chem 94:3564–3576

Jezequel JY, Bouchy M, Andre JC (1982) Estimation of fast fluorescence lifetimes with single photon counting apparatus and the phase plane method. Anal Chem 54:2199–2204

Kendall MG, Stuart A (1979) The advanced theory of statistics, vol 2, 4th edn. Macmillan, New York

Kennedy WJ Jr, Gentle JE (1980) Statistical computing. Dekker, New York

Knutson JR, Beechem JM, Brand L (1983) Simultaneous analysis of multiple fluorescence decay curves: a global approach. Chem Phys Lett 102:501–507

Lehmann EL (1981) Theory of point estimation. Wiley, New York

Livesey AK, Brochon JC (1987) Analyzing the distribution of decay constants in pulse-fluorometry using the maximum entropy method. Biophys J 52:693–706

McCullagh P, Nelder JA (1983) Generalized linear models. Chapman and Hall, London

McKinnon AE, Szabo AG, Miller DR (1977) The deconvolution of photoluminiscence data. J Phys Chem 81:1564–1570

Meyer SL (1975) Data analysis for scientists and engineers. Wiley, New York

More JJ (1977) The Levenberg-Marquardt algorithm: implementation and theory. In: Watson GA (ed) Numerical analysis. Springer, New York Berlin Heidelberg

Morris AH Jr (1981) NSWC/DL library for mathematics subroutines. Naval surface weapons center. Dahlgren Vi, pp 183–184

Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7:308–313

O'Connor DV, Phillips D (1984) Time-correlated single photon counting. Academic Press, New York

O'Connor DV, Ware WR, Andre JC (1979) The deconvolution of photoluminescence data. J Phys Chem 83:1333–1343

Olsson DM, Nelson LF (1975) Nelder-Mead simplex procedure for function minimization. Technometrics 17:45–51

Periasamy N (1988) Analysis of fluorescence decay by the nonlinear least squares method. Biophys J 54:961–967

Phillips GR, Eyring M (1988) Error estimation using the sequential simplex method in nonlinear least squares data analysis. Anal Chem 60:738–741

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical recipes. Cambridge University Press, Cambridge

Provencher SW (1982a) A constrained regularization method for inverting data represented by linear algebraic or integral equations. Comp Phys Comm 27:213–227

Provencher SW (1982b) CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. Comp Phys Comm 27:229–242

Sandor T, Conroy MF, Hollenberg NK (1970) The application of the method of maximum likelihood to the analysis of tracer kinetic data. Math Biosci 9:149–159

Selinger BK, Harris CM (1983) A critical appraisal of analytical methods. NATO ASI (Adv Sci Inst) Ser A Life Sci 69:155–168

Siemiarczuk A, Wagner BD, Ware WR (1990) Comparison of the maximum entropy and exponential series methods for the recovery of distributions of lifetimes from fluorescence lifetime data. J Phys Chem 94:1661–1666

Small EW, Libertini LJ, Brown DW, Small JR (1989) Extensions of the method of moments for deconvolution of experimental data. In: Menzel ER (ed) SPIE Proceedings: Fluorescence Detection III 1054:36–53

Szabo AG, Bramal L (1983) Modulating functions – a deconvolution method. NATO ASI (Adv Sci Inst) Ser A Life Sci 69:271–283

Valeur B, Moirez J (1973) Analyse des courbes de décroissance multiexponentielles par la méthode des fonctions modulatrices – application à la fluorescence. J Chim Phys 70:500–506

Ware WR, Doemeny LJ, Nemzek TL (1973) Deconvolution of fluorescence and phosphorescence decay curves. A least-squares method. J Phys Chem 77:2038–2048

Willis KJ, Szabo AG, Zuker M, Ridgeway JM, Alpert B (1990) Fluorescence decay kinetics of the tryptophyl residues of myoglobin: Effects of heme ligation and evidence for discrete lifetime components. Biochemistry 29:5270–5275

Wong AL, Harris JM (1989) Quantitative estimation of component amplitudes in multiexponential data: Application to time-resolved fluorescence spectroscopy. Anal Chem 61:2310–2315