ELSEVIER

Biophysical Chemistry 104 (2003) 249–258

Biophysical Chemistry

www.elsevier.com/locate/bpc

Application of the maximum entropy method to absorption kinetic rate processes

Zs. Ablonczy[1], A. Lukács[1], E. Papp*

Department of Biological Physics, Eötvös University, Pázmány P. sétány 1/A, Budapest H-1117, Hungary

Received 18 December 2001; received in revised form 15 June 2002; accepted 10 November 2002

Abstract

The maximum entropy method, originally developed for astronomical image restoration, has already been successfully applied to a variety of biophysical problems. Through numerical inverse Laplace transformation, the method determines the lifetime distribution function with the largest informational entropy. Starting from a flat distribution, it results in the consistent selection of a single distribution from the numerous possible ones that correctly fit the data. In this paper, we discuss the application of the method to kinetic processes that have both rise and decay components, and test the algorithm with different signal to noise ratio generated data. It is proved that the mass conservation constraint can be taken into account by reducing the search to a lower dimensional subspace. The effect of noise on the width of lifetime distribution is studied and it is shown that an inherent entropy connected to the underlying kinetics can be separated from the noise generated entropy. The possibility of the application of the method to the photocycle kinetics of bacteriorhodopsin is also shown.

Keywords: Lifetime distribution; Inherent entropy; Bacteriorhodopsin

1. Introduction

In consequence of the complexity of the structure and function of even the smallest proteins, discrete exponentials are in many cases inadequate to describe the observed kinetics in protein related processes. The complexity and nonlinearity of these reactions necessarily cause lifetime widening that restricts the use of the discrete exponential analysis [1,2].

Another problem is that the number of discrete exponentials used for fitting the observed data has to be estimated. In order to estimate the number of the necessary exponentials, an assumption has to be made about the way the kinetics appear. In consequence, the fitting function contains kinetic model dependent information. This characteristic of the discrete analysis makes it difficult to draw appropriate conclusions about the underlying kinetics [3].

Finally, within the discrete exponential analysis the 'goodness' of a chosen kinetic model depends mainly on a minimum misfit parameter calculated from the resulting fitting function and the observed

*Corresponding author. Tel.: +36-1-372-2785; fax: +36-1-372-2757.

E-mail address: pappe@ludens.elte.hu (E. Papp).

[1] These authors contributed equally to this work.

0301-4622/03/$ - see front matter
PII: S0301-4622(02)00379-4

data. However, the larger the number of the discrete exponentials, the less is the misfit which property poses difficulty in choosing the correct realistic kinetic model [4].

The basic idea behind maximum entropy method (MEM) is that an entropy function can be defined on a set of positive numbers. If these numbers approximate the observed data, then maximizing this entropy subject to the minimum misfit constraint will give a single image from the numerous images that may correctly fit the data. It is believed that this image will not have any structural information that could be regarded as an effect of noise and no additional a priori information is necessary about the underlying structure [5].

Although the MEM was originally developed for astronomical image restoration [5,6] it has already been successfully used in a variety of fields involving time resolved fluorescence spectroscopy [2,7,8], ligand rebinding rate determination [1,9], X- and γ-ray crystallography [10,11], neutron scattering [12] and several other fields. In the case of kinetic rate analysis, the distribution of the rate coefficients and the appropriate time functions are connected through Laplace transformation. The entropy function (Shannon–Jaynes informational entropy) is defined over the rate coefficients and the misfit constraint is the χ^2-statistics which measures how well the Laplace transformed rate coefficients fit the observed data [1,2,8]. The MEM gives a way to directly (model independently) determine the probable rate constant distribution of the processes under examination. The observed features (peak number, -position or -widening) can then be interpreted and distinction be made between possible kinetic models [1].

However, in all of the above applications, the implemented MEM algorithms have only three constraints: positivity and maximum entropy of the underlying functions and minimum misfit of the noisy data. The problem where we wanted to apply MEM is more complex. In absorption kinetic measurements, the rate distribution shows not only decay but also an increasing component, and the overall integral of the rise and decay parts of the distribution must be equal in some cases. This additional restriction poses a problem which the algorithms described earlier [1,5,6,13] are not able to cope with. Our first aim is to expand the general algorithm described by Skilling and Bryan [5] to empower it to handle this new constraint. Within this analysis we make a simplifying assumption: we assume that the rise and decay parts of the distribution are separated on the lifetime scale. This restricts the possible application of the presented MEM analysis.

With the operational algorithm, we constructed a set of different signal-to-noise ratio generated data that corresponded to the above restriction condition. This set of data made it possible to check the operation of the algorithm and study the effect of noise and distribution width on the results. Finally, we checked the operation of the algorithm on bacteriorhodopsin (BR) M state (410 nm) absorption kinetic data. All the tests gave a satisfactory result.

In this paper, we report the result of the extension of MEM for absorption kinetic analysis. The application for the analysis of the BR photocycle is under way and the result of this analysis will be published elsewhere.

2. Maximum entropy method

In the simplest possible case, the time course of an absorption kinetic rate process can be represented by the following function:

$$A(t)=f(\mathrm{e}^{-t/\tau_2}-\mathrm{e}^{-t/\tau_1}) \tag{1}$$

This function describes a rate process of a single rise lifetime τ_1, decay lifetime τ_2 and amplitude f. If the rate process becomes more complex both the rise and the decay will be described by distributions:

$$A(t)=\int_0^\infty f(\tau)\mathrm{e}^{-t/\tau}\mathrm{dlog}(\tau) - \int_0^\infty f'(\tau)\mathrm{e}^{-t/\tau}\mathrm{dlog}(\tau) \tag{2}$$

Here $f(\tau)$ is the decay and $f'(\tau)$ is the rise distribution coefficients (amplitudes). The distributions are defined on a logarithmic time scale

since the rate processes may stretch several orders of magnitude in the lifetime scale. The numerical approximation of this integral also requires a time constant space with logarithmically uniformly spaced intervals. In this approximation the discretized distributions at a given t_j time are the following:

$$A_j = -\sum_{i=1}^{n_1} f_i e^{\frac{t_j}{\tau_i}} + \sum_{i=n_1+1}^{n} f_i e^{\frac{t_j}{\tau_i}} \quad (3)$$

f_i are a set of positive numbers, the discretized lifetime distribution coefficients of an absorption kinetic rate process. They represent two continuous smooth functions, the rise $(i=1\ldots n_1)$ and the decay $(i=n_1+1\ldots n)$ processes. The f_i amplitudes and the resulting time function are connected through a Laplace transformation. (Note that any sign is not included in the distribution since the f_i must be positive numbers in order to get a simple entropy function definition.) We have to emphasize here that accepting Eq. (3) a strong limitation is introduced on a possible time function $A(t)$: we assume that the underlying kinetics allows the separation of the rise and decay parts on the lifetime scale. For a more general case this restriction in Eq. (3) has to be lifted. However, in this paper we apply the MEM analysis mostly for generated data where the separation is fulfilled.

In any real observation the data are noisy. This means that there cannot be a complete agreement between the calculated smooth time function (Eq. (3)) and the obtained data. $f_{1\ldots n}$ should satisfy a condition of minimum misfit. The misfit can be defined as the ratio of the averaged squared difference of the measured and calculated data and the square of the standard deviation for the jth data point:

$$C = \frac{1}{m}\sum_{j=1}^{m} \frac{(D_j - A_j)^2}{\sigma_j^2} \quad (4)$$

Here $D_{1\ldots m}$ are the data observed at time points $t_{1\ldots m}$, respectively and σ_j is the standard deviation for the jth data point. $C=1$ is obviously a good fit criterion in which case the average difference of the measured and calculated data is equal to the average noise. Decreasing C below 1 results in a fit to the noise on the underlying data. For experimental data σ_j are usually not known. In that case instead of the $C=1$ criterion, iterations are finished where least-square deviations do not change essentially with further iterations.

The distributions built up with the resulting $f_{1\ldots n}$ numbers must also correspond to maximum entropy. Entropy, however, can be defined in several ways. In other applications of MEM it was reported that several different definitions of entropy could almost equally well be used [5,13]. We mention here two definitions. The Shannon informational entropy [14–16] is given by the following relation:

$$S = -\sum_{i=1}^{n} p_i \ln(p_i), \qquad \text{where } p_i = \frac{f_i}{\sum_{i=1}^{n} f_i} \quad (5)$$

Skilling and Bryan [5] used another definition:

$$S = -\sum_{i=1}^{n} f_i \left[\ln(f_i/A) - 1\right] \quad (6)$$

where A is a user-defined value. This definition has some advantage in the realization of the MEM algorithm (see later).

Although Eq. (6) was applicable to our problem, it turned out that we could not give any physical meaning to A and the result depended on it. We decided to use the strict form given by Eq. (5), where S depends only on f_i and there is no additional parameter.

The maximum entropy criterion requires to select a solution (f_i) which maximizes S subject to $C \leqslant C_{\text{aim}}$ where $C_{\text{aim}} \approx 1$ as a good fit criterion. If a Lagrangian multiplier α is set up the solution will lie at the extremum of $Q = \alpha S - C$ for some appropriate value of α. The problem is solved by an iterative algorithm. Starting with a uniform $(f_i = \text{constant})$ distribution a solution is approached after some iterations varying the α parameter. However, the space in which one has to find the solution is n-dimensional. Instead of using this huge space (n may be several hundred) Skilling

and Bryan [5] suggested the use of a small subspace for the search of the maximum of Q. The very effective algorithm of Skilling and Bryan has two essential elements: an entropy metric is introduced in the lifetime distribution (amplitude space) and a low dimensional subspace is used for iterations. Three sophisticated search directions—defining the subspace—can capture the structure of the maximum entropy problem:

$$e_1(i)=\frac{\partial S}{\partial f_i} \quad e_2(i)=\frac{\partial C}{\partial f_i}$$
$$e_3(i)=\sum_{j=1}^{n}\frac{\partial^2 C}{\partial f_i \partial f_j}\left(\frac{(\nabla S)_i}{|\nabla S|}-\frac{(\nabla C)_i}{|\nabla C|}\right) \tag{7}$$

The S and C functions are approximated by quadratic forms around the last iterated point (f_i^{last}) in the amplitude space:

$$S'=S_0+\sum_i S_i\delta f_i+\frac{1}{2}\sum_{ij}S_{ij}\delta f_i\delta f_j$$
$$\text{where } S_i=\frac{\partial S}{\partial f_i} \quad S_{ij}=\frac{\partial^2 S}{\partial f_i \partial f_j} \tag{8}$$

(and a corresponding expansion for C).

As the subspace defined by Eq. (7) is not orthogonal, a metric tensor has to be used. However, this metric tensor can be orthogonalized making the problem simpler. S' and C' are transformed with the metric tensor. In the three-dimensional orthonormal subspace (with x_μ coordinates, $\mu=1, 2, 3$) it is possible to make another orthogonal transformation by diagonalizing $S_{\mu\nu}$ or $C_{\mu\nu}$. In this subspace, Q' is maximized at a given parameter value α chosen such that the quadratic approximation shall still be valid (α-chop, [5]). After that x_μ is transformed back to the (f_i) space giving the next iteration. In this work the α-chop was realized by an algorithm which compared directly the exact values for S and C and the approximate values given by the quadratic approximation. The relative deviations were always less than 10^{-3}.

The use of an entropy metric was proposed by Skilling and Bryan [5] before reducing the search to the subspace. One advantage of the entropy metric which introduces the following scale transformation (around each iteration):

$$\delta z_i=\frac{1}{\sqrt{f_i}}\delta f_i \tag{9}$$

is that the second derivative of the entropy given by Eq. (6) becomes the negative of the unit tensor:

$$\frac{\partial^2 S}{\partial z_i \partial z_j}=-\delta_{ij} \tag{10}$$

and transforms in the same way as the metric tensor. In that case $C_{\mu\nu}$ can also be diagonalized in the subspace and the extremum problem becomes simpler. Although we used the entropy metric, this simplification does not occur for the Shannon entropy (as the second derivative of Eq. (5) is not so simple) and $S_{\mu\nu}$ and $C_{\mu\nu}$ cannot be diagonalized at the same time. Usually we diagonalized $S_{\mu\nu}$ in the subspace but sometimes we met the problem that one eigenvalue of $S_{\mu\nu}$ turned out to be small positive number (saddle point). To avoid this problem—though the iteration worked well—an approximative second derivative was used for S_{ij}. The first derivative of the Shannon entropy is given by:

$$\frac{\partial S}{\partial f_i}=-(\ln f_i-a/sn)/sn$$
$$\text{where } \mathrm{sn}=\sum_i f_i \quad a=\sum_i f_i \ln f_i \tag{11}$$

Assuming that sn and a are slowly changing between subsequent iterations, the second derivative of S will be given approximately by:

$$\frac{\partial^2 S}{\partial f_i \partial f_j}\approx -\frac{1}{sn}\frac{1}{f_i}\delta_{ij} \tag{12}$$

This form transforms to a unit tensor in the entropy metric. For the quadratic approximation, usually Eqs. (11) and (12) were used.

Having found the maximum solution (x^μ) of Q' in the subspace for the given iteration, it has to be transformed back to the n-dimensional amplitude

space:

$$f_i^{\text{new}} = f_i^{\text{old}} + \delta f_i = f_i^{\text{old}} + \sqrt{f_i} \sum_{\mu} e_\mu(i) x^\mu \tag{13}$$

(Here the entropy metric was taken into account.)

In absorption kinetic problems, an additional constraint may arise. If the absorption $A(t)$ is related to an intermediate rate product which appears and disappears during the process (e.g. the M intermediate in the BR photocycle) or $A(t)$ is the concentration of an intermediate, then $A = 0$ for $t = 0$ and $t \to \infty$. While the second condition is satisfied by Eq. (3), the first will give the following constraint:

$$-\sum_{i=1}^{n_1} f_i + \sum_{i=n_1+1}^{n} f_i = 0 \tag{14}$$

We found that the MEM algorithm by itself can not satisfy this constraint with enough precision. It is interesting to examine the condition given by Eq. (14). If we define a vector (or rather a direction) in the n-dimensional amplitude space with the following components:

$$ie(i) = \begin{cases} -1 & \text{if } i \leqslant n_1 \\ \\ +1 & \text{if } i > n_1 \end{cases} \tag{15}$$

then Eq. (14) can be written as a scalar product in the n-dimensional space:

$$ie \cdot \boldsymbol{f} = 0 \tag{16}$$

It means that the solution of the problem should be included in a $(n-1)$-dimensional hyperplane orthogonal to the vector ***ie***.

The iteration starts with a flat distribution which satisfies Eq. (14), but as the search for minimum misfit and maximum entropy is iterated this condition is soon spoiled. However, if in each successive iteration, the condition is successfully maintained, then it remains to be valid throughout. Combining Eqs. (13), (14) and (16) this criterion can be expressed by the following equation:

$$\begin{aligned} 0 &= -\sum_{i=1}^{n_1} \delta f_i + \sum_{i=n_1+1}^{n} \delta f_i = \sum_{i=1}^{n} ie(i) \delta f_i \\ &= \sum_{i=1}^{n} \sum_{\mu=1}^{3} ie(i) \sqrt{f_i} e_\mu(i) x^\mu \end{aligned} \tag{17}$$

Therefore, this condition is maintained if the $\mathbf{x}^\mu$ vector is chosen from a plane in the subspace orthogonal to a vector, whose components are:

$$eg_\mu = \sum_{i=1}^{n} ie(i) \sqrt{f_i} e_\mu(i) \tag{18}$$

since after the proper substitution Eq. (17) is equivalent to:

$$\sum_{\mu=1}^{3} eg_\mu x^\mu = 0 \tag{19}$$

Restricting the search directions to the plane of the subspace orthogonal to the ***eg*** vector automatically satisfies the restriction condition throughout the algorithm. It can also be seen (Eq. (18)) that the ***eg*** vector in the subspace is the projection of the normal vector, ***ie***, to the hyperplane in the n-dimensional search space. In practice, an Euler rotation was made in the subspace, which turned one of the coordinate axes (x^3) into the ***eg*** direction and the search directions were restricted to the (x^1, x^2) plane to fulfill Eq. (19).

3. Results and discussion

In order to test the algorithm described above, we generated a set of gaussian data satisfying the zero integral distribution condition with different peak widths and random noise. The lifetime data set was generated with 200 logarithmically uniform time steps from 10^{-3} to 10^{3} ms. The first half of this data set was the rise and the second the decay distribution. The gaussians were normalized with centres at 0.042 ms (rise) and 4.2 ms (decay), respectively. Four increasing width Gauss distributions were generated ($10^{0.075}$, $10^{0.15}$, $10^{0.3}$, $10^{0.45}$). With the different width gaussians it may be possible to simulate some effect of the gaussian widening that may occur in real protein

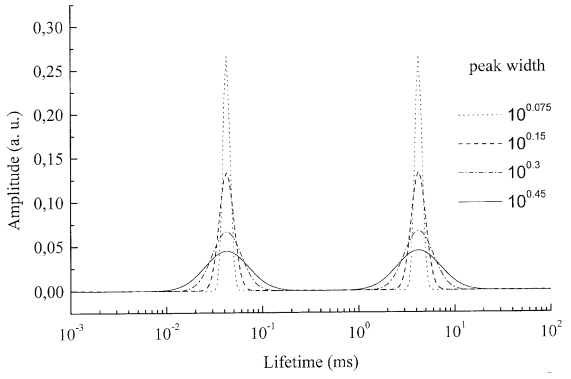

Fig. 1. The four generated lifetime distributions.

environments when proteins interact with their neighbourhood or as a result of their own dynamic fluctuations. The lifetime distributions were then Laplace transformed to the appropriate time functions and to each of them four different standard deviation gaussian noises were added ($10^{0.001}$, $10^{0.005}$, $10^{0.01}$, $10^{0.05}$). Some noise is always present in any real measurements and also causes some widening in the distribution. Our aim was to obtain some information on the separability of these two distribution-widening effects. Fig. 1 shows the different lifetime distributions while Fig. 2 indicates the effect of the added noise on the time functions. (Note that the lifetime distributions in Fig. 1 satisfy Eq. (3).)

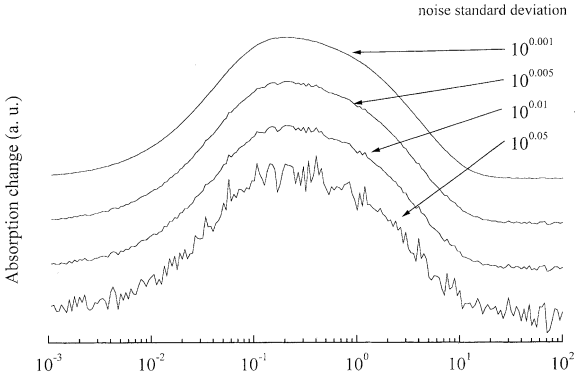

Fig. 2. The time functions of the distribution with peak widths $10^{0.15}$ with the different amounts of added noise.

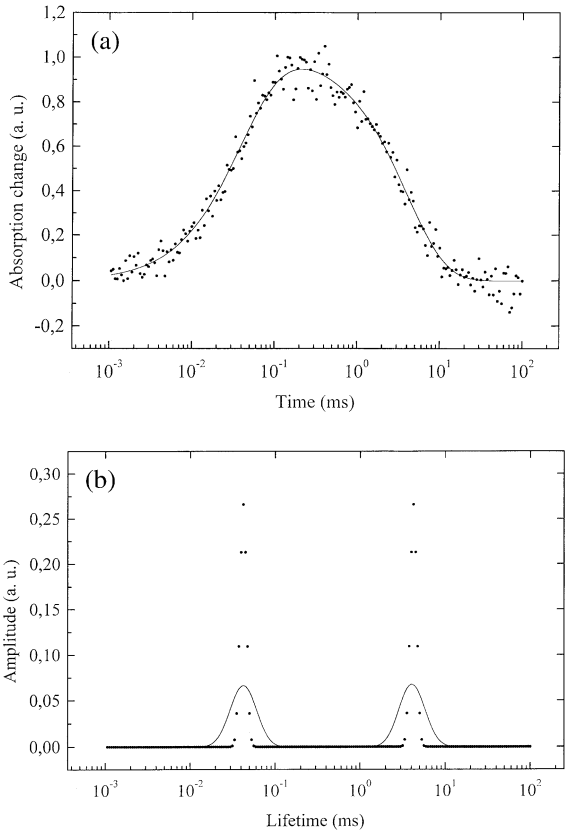

Fig. 3. (a) Generated and resulting time functions compared. Circles correspond to the generated data the continuous line is the MEM fit. The peak width of the generating gaussian was $10^{0.075}$ the standard deviation of the added noise $10^{0.05}$. (b) Generated and resulting lifetime distributions compared. Circles correspond to generated data, the continuous line is the MEM fit.

The generated absorption data were analyzed by MEM. The results are shown in Fig. 3a and b and Fig. 4a and b for two characteristic and limiting cases. Fig. 3a shows the fit to a noisy absorption data where the initial, generating lifetime distribution is narrow. There is a strong widening effect of the noise on the lifetime distribution (Fig. 3b). In the other case the starting lifetime distribution is wide but the noise level is low. Fig. 4a shows the MEM fit to the absorption data and Fig. 4b shows the starting lifetime distribution and the MEM result. The effect of noise here is almost negligible and it can be seen that the MEM result

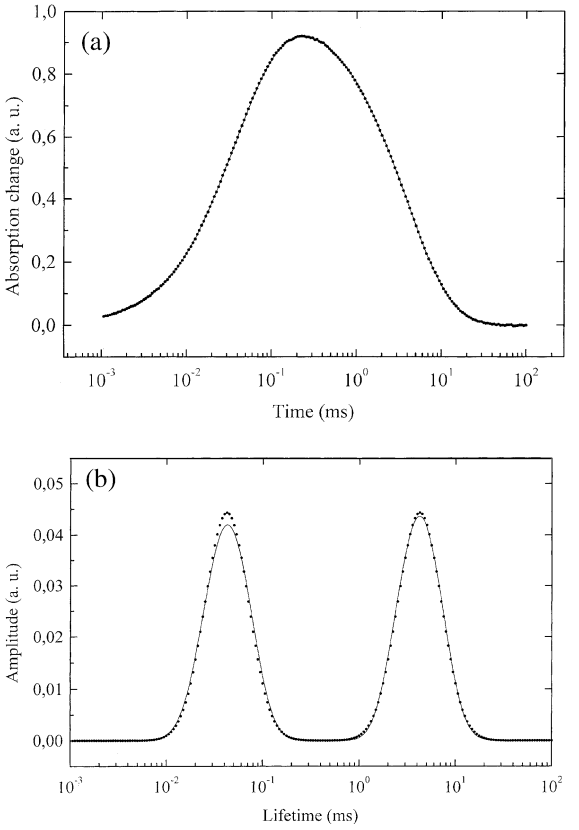

Fig. 4. (a) Generated and resulting time functions compared. Circles correspond to the generated data the continuous line is the MEM fit. The peak width of the generating gaussian was $10^{0.45}$ the standard deviation of the added noise $10^{0.001}$. (b) Generated and resulting lifetime distributions compared. Circles correspond to generated data, the continuous line is the MEM fit.

is approximately gaussian, i.e. the MEM retains the general shape of the starting lifetime distribution. Some deviation from the gaussian shape can be explained by the fact that gaussian is only a quadratic model of the MEM result [13].

The result of the MEM analysis of the generated data is summarized in Fig. 5. From the starting lifetime distribution (Fig. 1) we can calculate an 'intrinsic' entropy. The noise will increase this intrinsic entropy and in the absorption data (Fig. 2) they appear together. With generated data, in contrast with real-time measurement, they can be separated. Naturally, the entropy increases with the increase of the added noise. The MEM iteration starts from a flat distribution. The initial flat distribution has the largest possible informational entropy. It depends only on the number of the divisions on the logarithmic lifetime scale ($S=\ln(n)$ in our case for $n=200$ and $S=5.3$). The simplest kinetics can be described by two single exponentials (one rise and one decay). Its entropy is $\ln(\sqrt{2})=0.35$. The expected information entropies must lie in between depending on the peak width of the distribution. The effect of noise on the informational entropy content is complex. The wider the distribution the smaller the increase in the entropy caused by the noise. If a distribution is narrow the same amount of added noise increases the entropy much stronger than if it is wider. However, there is still some increase even with the widest distribution. This means that it is easier to detect the effect of noise if the distribution is narrower and more difficult when the distribution is widened. This observation may give a tool for at least some basic quantitative analysis of the widening of the protein absorption kinetics due to the different environments. Although in itself the information entropy content is not enough to characterize the heterogeneity of the environment [2] recording the data with different amount of inherent noise (using different amount of averaging) may overcome this difficulty [17,18].

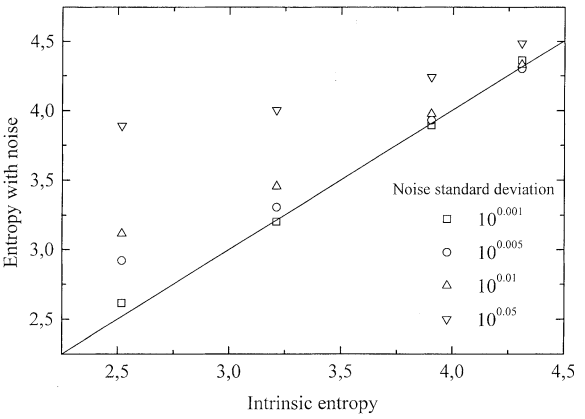

Fig. 5. The dependence of the entropy on the noise. Straight line: entropy without noise calculated from the amplitude distributions (Fig. 1). Symbols: MEM results of the noisy data.

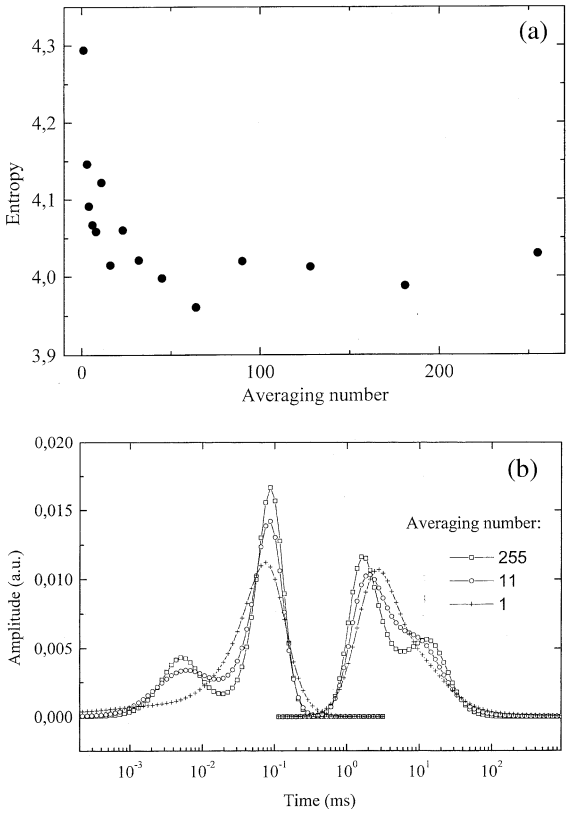

Fig. 6. (a) The dependence of the entropy on the averaging number. (b) The lifetime distributions at various numbers of averaging.

As a test to the algorithm we tried its operation on BR M state (410 nm) absorption kinetic data. At this wavelength the rise and decay are separated and Eq. (3) is applicable. (The data were recorded at 25 °C and 7 mJ exciting laser energy at 532 nm wavelength, purple membrane in 100 mM phosphate buffer at pH 9.) Absorption data were recorded with increasing number of averaging and analyzed by MEM. Fig. 6a shows the entropy in function of the averaging number in the range of 1–256. The entropy decreases with the number of averaging (i.e. with decreasing noise) and it levels off at approximately 4. This behavior is similar to the generated data result (Fig. 5) for the case when the intrinsic entropy is rather high. We believe that this high intrinsic entropy is related to the BR, it reflects protein dynamics and/or some protein heterogeneity. A thorough analysis is needed to answer this question. Fig. 6b shows the lifetime distributions for three different averaging numbers. There are four broad, nearly gaussian peaks which are connected with the rise and decay of the M intermediate state.

The above view is also supported by the following considerations. If we describe the photocycle of the BR with a first order kinetic model containing definite number of intermediates then the time dependence of the intermediate concentrations can be calculated with the help of eigenvalues and eigenvectors of the matrix (K matrix) determined by the rate constants [19]. Taking a kinetic model [20] most accepted in the literature:

$$K \underset{k_2}{\overset{k_1}{\rightleftarrows}} L \underset{k_4}{\overset{k_3}{\rightleftarrows}} M_1 \underset{k_6}{\overset{k_5}{\rightleftarrows}} M_2 \underset{k_8}{\overset{k_7}{\rightleftarrows}} N \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} O \xrightarrow{k_{11}} \mathrm{BR} \quad \text{and} \quad N \xrightarrow{k_{12}} \mathrm{BR}$$

the 12 rate constants can be determined by a fit to the experimental absorption change taken at different wavelengths [20]. Using the K-matrix approach the time dependence of the concentrations of the intermediate states can be given by a sum of exponentials:

$$C_i(t) = \sum_{j=1}^{n} c_{ij} \exp(k_j^* t) \tag{20}$$

Here n is the number of the intermediate states and k_j^* are the eigenvalues of the K-matrix (apparent rate constants) and the c_{ij} parameters are determined by the initial conditions. This is the form which can be analyzed by MEM. (A remark has to be made here: The (k_i) set uniquely determines the k_j^* eigenvalues but the opposite is not true. This makes the comparison of the MEM results and a given kinetic model more difficult. The determination of concentrations from absorption kinetic measurements, without using some specific kinetic model, is not a straightforward problem. Only the M-state is an easier case where the absorption change at 410 nm is approximately proportional to the concentration of the M ($M_1 + M_2$) intermediate.) As the concentrations of the intermediate states can not be determined directly from absorption kinetic measurements a MEM

analysis should start at the beginning. Absorption kinetic measurements at different wavelengths give absorption changes during the photocycle functionally similar to Eq. (20) but not necessarily positive values (therefore Eq. (14) is not satisfied). Analyzing these absorption changes with MEM could yield the k_j^* eigenvalues and the amplitudes without any assumption for the kinetics and this result can be compared with different kinetic models. For concentrations the rise and decay parts probably are separated in some cases (e.g. for the M state) and the assumption made in Eq. (3) is relevant. But for the general case of the absorption changes this is not valid and the restriction given by Eq. (3) has to be removed. This generalization of the method is under work and the result will be published elsewhere.

Using the rate constants given in Ref. [20] for pH 9 and 25 °C, Eq. (20) gives four significant exponentials for the M ($M = M_1 + M_2$) intermediate. If we add a little noise to this curve and analyze it by MEM we get the result shown in Fig. 7a and b. As can be seen MEM gives similar discrete lifetimes for that case as the starting kinetic model.

As a conclusion we can say that if the concentration results from discrete lifetime distribution MEM is able to determine the discrete apparent rate constants. However, fitting to a measured absorption change at 410 nm MEM gives a broad (nearly gaussian) distribution (Fig. 6b). The lifetimes and amplitudes of the distribution correlate well with the discrete lifetime amplitudes derived from the kinetic model.

Skilling and Bryan [5] introduced a test parameter to check whether a real entropy maximum solution resulted. In that case the test parameter should be close to 0. Our algorithm does not fulfill this criterion, as we get larger and varying values for the test parameter. We believe that this is connected with the constraint given by Eq. (14) and that the solution should lie in the lower dimension hyperplane. Lifting this constraint improves the test parameter but the resulting rate distribution is differing from the generated one even at low noise level, though the fit to the time function is good.

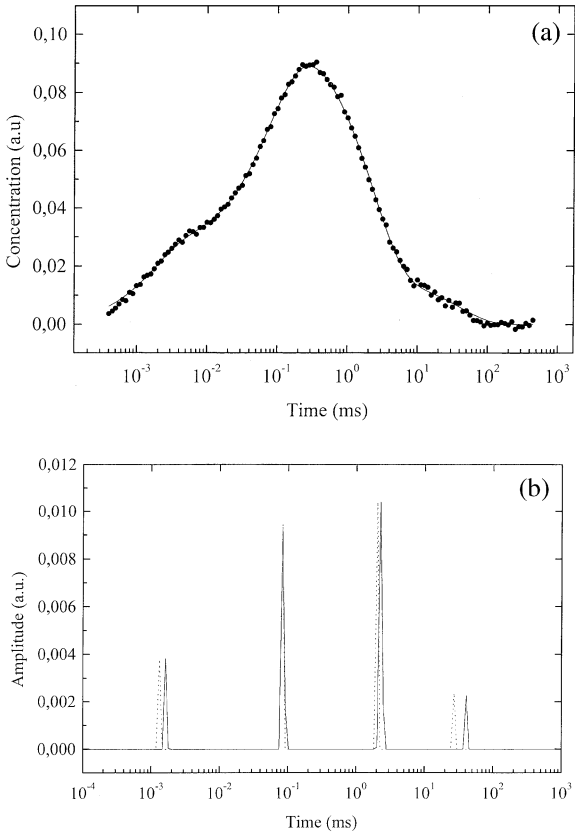

Fig. 7. (a) MEM fit (continuous line) to the concentration determined by the kinetic model (circles). (b) Lifetime distributions for the M state concentration given in (a). (Dotted line: starting kinetic model; continuous line: MEM fit.)

In conclusion, the above results prove that the MEM can be applied to absorption kinetic analysis. On generated data intrinsic and noise contribution to the entropy can be separated, noise contributes essentially to the entropy only in that cases when the rate distribution is narrow. In the special case analyzed here by MEM for BR photocycle, the broadening in rate distribution seems to be an intrinsic property related to the protein dynamics or heterogeneity.

Acknowledgments

We thank Dr Zs. Dancsházy and Dr Gy. Váró at the Biophysics Institute of the Hungarian Acad-

emy of Sciences for providing us with PM preparations and for the valuable discussions, Dr F. Deák for his help with the generation of the gaussian noise and Dr P. Gnädig for his valuable consultation.

References

[1] P.J. Steinbach, K. Chu, H. Frauenfelder, et al., Determination of rate distributions from kinetic experiments, Biophys. J. 61 (1992) 235–245.
[2] R. Swaminathan, N. Periasamy, Analysis of fluorescence decay by the maximum entropy method: influence of noise and analysis parameters on the width of the distribution of lifetimes, Proc. Indian Acad. Sci. (Chem. Sci.) 108 (1996) 39–49.
[3] E. Papp, V.H. Ha, Local analysis of the M state kinetics of bacteriorhodopsin, Biophys. Chem. 57 (1996) 155–161.
[4] J.F. Nagle, L. Zimányi, J.K. Lányi, Testing BR photocycle kinetics, Biophys. J. 68 (1995) 1490–1499.
[5] J. Skilling, R.K. Bryan, Maximum entropy image reconstruction: general algorithm, Mon. Not. R. Astron. Soc. 211 (1984) 111–124.
[6] S.F. Gull, G.J. Daniell, Image reconstruction from incomplete and noisy data, Nature 272 (1978) 686–690.
[7] A.K. Livesey, J.C. Brochon, Analyzing the distribution of decay constants in pulse-fluorometry using the maximum entropy method, Biophys. J. 52 (1987) 693–706.
[8] R. Swaminathan, G. Krisnamoorthy, N. Periasamy, Similarity of fluorescence lifetime distributions for single tryptophan proteins in the random coil state, Biophys. J. 67 (1994) 2013–2023.
[9] P.J. Steinbach, Two-dimensional distributions of activation enthalpy and entropy from kinetics by the maximum entropy method, Biophys. J. 70 (1996) 1521–1528.
[10] A.C. Fabian, R. Willingale, J.P. Pye, S.S. Murray, G. Fabbiano, The X-ray structure and mass of the Cassiopeia A supernova remnant, Mon. Not. R. Astron. Soc. 193 (1980) 175–188.
[11] J. Skilling, A.W. Strong, K. Bennett, Maximum entropy image processing in gamma-ray astronomy, Mon. Not. R. Astron. Soc. 187 (1979) 145–152.
[12] D.S. Sivia, Bayesian inductive inference maximum entropy and neutron scattering, Los Alamos Science, Summer 1990 (1990) 180–206.
[13] R. Narayan, R. Nityananda, Maximum entropy image restoration in astronomy, Ann. Rev. Astron. Astrophys. 24 (1986) 171–203.
[14] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[15] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 623–656.
[16] B.R. Frieden, Restoring with maximum likelihood and maximum entropy, J. Opt. Soc. Am. 62 (1972) 511–518.
[17] J.R. Alcala, E. Gratton, F.G. Prendergast, Interpretation of fluorescence decays in proteins using continuous lifetime distributions, Biophys. J. 59 (1987) 925–936.
[18] E. Bismuto, I. Sirangelo, G. Irace, Conformational dynamics of unfolded peptides as a function of chain length: a frequency domain fluorescence approach, Arch. Biochem. Biophys. 291 (1991) 38–42.
[19] J.J. Tyson, The Belousov–Zhabotinskii reaction, in: S. Levin (Ed.), Lecture notes in Biomathematics 10, Man. Berlin, Springer, Heidelberg, New York, 1976.
[20] K. Ludmann, Cs. Gergely, Gy. Váró, Kinetic and thermodynamic study of the bacteriorhodopsin photocycle over a wide pH range, Biophys. J. 75 (1998) 3110–3119.