Global and target analysis of time-resolved spectra

Ivo H.M. van Stokkum,

Department of Physics and Astronomy, Faculty of Sciences, Vrije Universiteit, Amsterdam, The Netherlands, ivo@nat.vu.nl,

over the past fifteen years in collaboration with:

Kate Mullen, Hans Spoelder († 1.4.2003), Emiel Corten († 7.7.1998), Zsófia Derzsi, Radmila Pribić, Fons Ullings, Frans Groen, Henri Bal, *Physics Applied Computer Science*,

Aad van den Bos, *Signal Processing, Applied Physics,*

Wouter Hoff, Johnny Hendriks, Michael van der Horst, Wouter Laan, Jasper van Thor, Remco Kort, Sergey Yeremenko, Jechiam Gural, Andrea Haker, Thomas Gensch, Klaas Hellingwerf, *Microbiology*

Taco Scherer, Xavier Lauteslager, George Wurpel, Mattijs Koeberg, John van Ramesdonk, Jan Verhoeven, Fred Brouwer, *Organic Chemistry,*

René Williams, Irina Petkova, Grzegorz Balkowski, Wybren Jan Buma, *Molecular Photonic Materials*

Cees van Walree, *Physical Organic Chemistry*,

Petra van den Berg, Ton Visser, *Biochemistry,*

Richard Lozier, Aihua Xie, *Biophysics, Biochemistry,*

Helen Linsdell, Geoff Abbott, Jonathan Hadden, *Biochemistry,*

Mike Jones, Des Heyes, Neil Hunter, *Biochemistry*,

Tomáš Polívka, *Chemical Physics,*

Delmar Larsen, Lucas Beekman, Marloes Groot, Frank van Mourik, Marion van Brederode, Bas Gobets, Manolis Papagiannakis, Mikas Vengris, Janne Ihalainen, Matthieu Visser, Michael Bloemendal, Bart van Haeringen, Paul van Kan, Andrius Baltuška, Olaf Larsen, Anjali Pandit, Erwin Peterman, Foske Kleima, Jante Salverda, Brenda Groen, Raoul Frese, Ronald Visschers, Chantal de Wit, Claudiu Gradinaru, Stefania Lampoura, Frank de Weerd, Markus Wendling, Miguel Palacios, Magdalena Gauden, Natalia Pawlowicz, Rudi Berera, Cosimo Bonetti, Alexander Doust, Luuk van Wilderen, Marta Germano, Brent Krueger, Krzysztof Gibasiewicz, John Kennis, Jan Dekker, Herbert van Amerongen, Rienk van Grondelle, *Biophysics*.

Summary

In (bio)physical/(bio)chemical research the response of a complex system to an externally applied perturbation is often studied. In most cases the obtained response is of a complexity that prohibits direct interpretation of the obtained data. In the case of time-resolved spectroscopy, measurements at multiple wavelengths can often be used to disentangle complex kinetics. These time-resolved spectra can be considered as an example of multiway data. In these lectures, the methodology for global and target analysis of time-resolved spectra is explained, both from a practical, and from a fundamental/mathematical point of view. To fully extract the information from the overwhelming amount of data a model-based analysis is mandatory. This analysis is based upon assumptions regarding the measurement process and upon a physicochemical model for the complex system. The model is composed of building blocks representing scientific knowledge and assumptions, such as the instrument response function, the components of the system connected in a kinetic scheme, and additional properties of the components (e.g. anisotropy). The use of a combined model for the kinetics and for the spectra of the components results in a more powerful spectro-temporal model. The model parameters, like rate constants and spectra, can be estimated from the data, thus providing a concise description of the complex system dynamics. This spectro-temporal modelling approach is illustrated with several case studies.

Keywords: global analysis, multiway data, photoactive yellow protein, spectrotemporal model, target analysis, time-resolved spectroscopy

Abbreviations: BR, bacteriorhodopsin; CT, charge transfer; DADS, Decay Associated Difference Spectra; DAS, Decay Associated Spectra; EADS, Evolution Associated Difference Spectra; ESA, Excited State Absorption; ESI, Excited State Intermediate; FWHM, Full Width at Half Maximum; GSA, Ground State Absorption; GSB, Ground State Bleach; GSI, Ground State Intermediate; I_0, first photocycle intermediate of PYP; I_1, second photocycle intermediate of PYP; IRF, instrument response function; MA, magic angle ; mOD, milli optical density; NLLS, nonlinear least squares; pCA, p-coumaric acid; PYP, photoactive yellow protein; PSE, Problem Solving Environment; SADS, Species Associated Difference Spectra; SAS, Species Associated Spectra; SE, Stimulated Emission; SVD, Singular Value Decomposition

1 Introduction

Time-resolved spectroscopy is a widely used tool in photophysics, photochemistry and photobiology to investigate the dynamic properties of complex systems [1,2]. Examples of such systems are chromophore-protein complexes essential for photosynthesis and photodetection, which are important model systems in bioenergetics. Many chromophore-protein complexes traverse a photocycle, e.g. the proton pump bacteriorhodopsin (BR) [3,4] and the photodetector photoactive yellow protein (PYP) [5,6]. A key question in these chromophore-protein complexes is how the chromophore dynamics are modified by the protein [7, 8]. This question can be adressed by time-resolved spectroscopy in combination with global and target analysis. Here *global* refers to a simultaneous analysis of all measurements, whereas *target* refers to the applicability of a particular target model. Fig. 1 depicts the pivotal role that global and target analysis play. The complicated data can
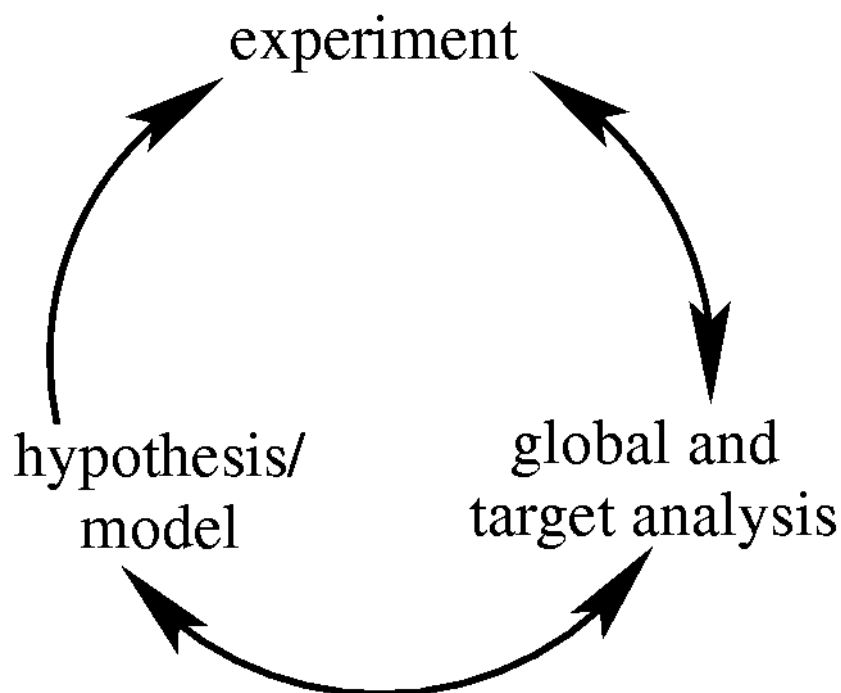
Fig.1. Graph depicting interrelations and position of global and target analysis in the scientific loop.

only be interpreted after thorough analysis (therefore the left backward arrow is missing), whereby the analysis can suggest a new model or hypothesis. The analysis can also help pinpoint systematic experimental errors, and inspire improved experimentation.

The analysis will benefit from a Problem Solving Environment (PSE), which is a computer system that provides all the computational facilities necessary to solve a class of problems [9]. Fig. 2 shows a flow chart of a prototype, with two loops. The upper parameter estimation loop
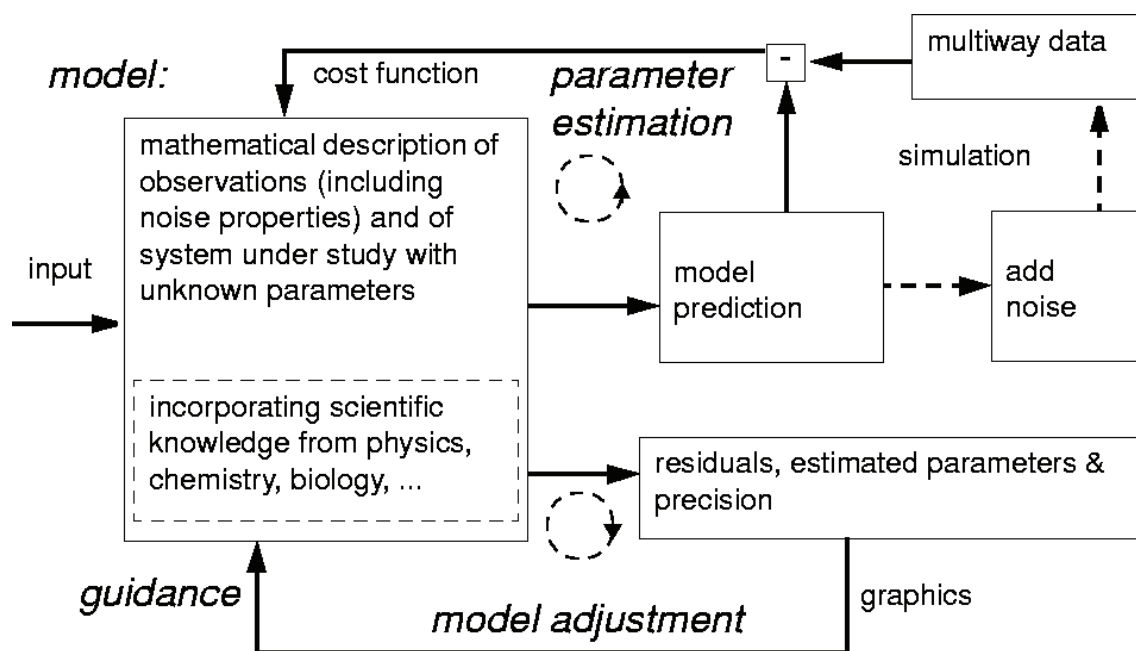
Fig.2. Flow chart of prototype Problem Solving Environment. Using the multiway data and the specified model the parameters are estimated in the upper loop. The bottom model adjustment loop is traversed many times, which is the main motivation for interactivity of the PSE. Further explanation in text.

is well known from fitting models to data. However, it is the bottom model adjustment loop which is traversed many times, and which some experimenters refer to as "the hell of fitting". Phenomena occuring on a variety of time scales ranging from femtoseconds to seconds can be studied. The input to the system usually consists of a short pulse of high energy which is absorbed and triggers a series of reactions. These reactions are often accompanied by changes in the UV, visible or IR spectroscopic properties of the system which can be measured. The output of the system is thus a collection of measurements of a spectroscopic property like absorption or emission as a function of time and wavelength, which is called a time-resolved spectrum.

A time-resolved spectrum is the most well-known example of two-way data. These data are a collection of measurements in two dimensions (ways). The first dimension is the independent experimental spectral variable: wavelength λ or wavenumber $\bar{\nu}$, or magnetic field strength B, etc. The second dimension is an independent experimental variable to monitor spectral change: time t after excitation, temperature T, polarization dependence, acidity pH or pD, excitation wavelength, or quencher concentration [Q]. Adding a third dimension results in

three-way data, of which the combination of time and wavelength with temperature or with quencher concentration are the most well known. In this paper we concentrate on time-resolved spectra, however many of the methods are also applicable for other types of multi-way data [10,11].

To unravel the processes underlying the observable spectroscopic changes, which result in overwhelming amounts of data, a model-based analysis of the measurements is mandatory. From an analysis perspective two problems can be distinguished: (a) when a parameterized model of the observations is available, the parameters have to be estimated in a statistically sound way, (b) when only a class of models is known, in addition also the "best" model needs to be determined.

Previously several methodological reviews have been written on global and target analysis by Beechem and co-workers [12, 13, 14], Ameloot, Boens et al. [15], Holzwarth [16] and Dioumaev [17]. These reviews demonstrate the importance of simultaneous (global) analysis of multiple decay traces. The combination of global analysis with testing of a photophysical or photochemical model is often called target analysis [12,16]. Reviews discussing global analysis in combination with experimental techniques are [1,18]. In addition, the bacteriorhodopsin photocycle has been a model system also for global and target analysis [19,20,21,22,23]. Recently five-way data (wavelength, time, temperature, polarization dependence, acidity) from this photocycle were subjected to a comprehensive target analysis [24]. In this paper an overview of the global and target analysis methodology will be given, emphasizing the need for modelling of both kinetics and spectra.

The word *model* is used here in two different ways. On the one hand a model for the observations is formulated in mathematical-statistical terms. Thereby the measurements are described stochastically, since taking into account the noise properties is essential for precise parameter estimation [25,26]. On the other hand, the experimentalist is studying a complex system, generating a huge amount of observations. The goal of the experiment is to arrive at a simplified description of the system and estimate the essential physicochemical parameters with the help of a parameterized model. Most often such a model consists of a kinetic scheme containing transitions between states, which is also called a compartmental model. In this case the word *model* is used as a simplified description, and not in the above statistical sense.

In §2 the assumptions regarding both the measurement process and the physicochemical model for the complex system which together lead to a model for the observations are

discussed. The most important equations for global and target analysis are presented, and explained. The subsections of §2 describe aspects of particular building blocks for the final physicochemical model. These building blocks are tailored for the experiments to be modelled. Thus a dynamic experiment usually requires a kinetic model, most often a compartmental model (§2.4.3). Taking into account the measurement conditions requires introduction of the instrument response function (IRF, §2.1.1). Extra building blocks are needed to describe e.g. temperature dependence (§2.4.3.6) or anisotropy properties (§2.7.5). Still this is only a description of the temporal aspects of the data. Spectral model assumptions have to be added on top of this. This combination of kinetic and spectral model assumptions, discussed in §2.7, is most promising for the unravelling of complex systems. It enables estimation of crucial parameters like branching ratios, which can only be estimated because of the spectral model assumptions. Throughout §2 examples are described to illustrate the methods. The most relevant aspects of the parameter estimation are explained in §3. In §4 an in depth case study of ultrafast dynamics in PYP will be presented.

These lecture notes are based upon a review [27], which is the preferred reference source.

2 Model for the observations

2.1 Measurement process

A time-resolved spectrum is a collection of measurements done at different (distinct) times and wavelengths. Three measurement sequences can be distinguished:

(a) Measurements can be done simultaneously at a great number of wavelengths and at a certain time delay with respect to the exciting pulse. This is called a time-gated spectrum. A collection of such time-gated spectra at different time delays constitutes a time-resolved spectrum. With pump-probe spectroscopy, a time-gated spectrum is susceptible to baseline fluctuations.

(b) Alternatively, at a particular wavelength a decay trace is measured as a function of time with respect to the exciting pulse. Again a collection of such decay traces measured at different wavelengths constitutes a time-resolved spectrum.

(c) Detection of decay traces at a great number of wavelengths simultaneously, providing high resoltion in both dimensions, is possible with a (synchroscan) streak camera in combination with a spectrograph [28,29].

The three types of measurements require different preprocessing (e.g., baseline correction) and differ in their noise statistics. The resolution of the measurements is determined by a number of instrumental characteristics and by the stochastics of the measurements. Time resolution is limited by both the width of the exciting laser pulse and the width of the detector response. Wavelength resolution is determined by the characteristics of the spectrograph used. Below we discuss in some detail aspects of the measurement process.

2.1.1 The instrument response function

Usually the system is excited by a short laser pulse of a certain energy. The convolution of the shape of this exciting pulse and the detector response is called the instrument response function (IRF). The IRF limits the fastest response observable in the experiment. With pump-probe spectroscopy the IRF is given by the convolution of pump and probe pulses. With Gaussian shaped pump and probe pulses, the convolution of the two will again be Gaussian shaped, but with an increased width. Ideally the IRF width should be shorter than the shortest time scale under study. This is impossible when studying ultrafast phenomena. On a (sub)picosecond time scale, the shape of the IRF as well as its timing precision become important.

Ideally the IRF should be measured once and for all with infinite precision, to avoid complications in the further analysis. In practice the problem can be tackled in different ways, depending upon the experimental technique. With fluorescence decay trace measurements there are three options: (a) the instrument response can be measured at the excitation wavelength (which is different from the emission wavelength) allowing for a wavelength-dependent time-shift parameter which needs to be estimated [30]. (b) the instrument response can be measured indirectly by adding a reference compound whose kinetic properties are known (reference convolution method, [31,32]). (c) a parameterized description of the instrument response can be included in the model function, leading to the necessity to estimate these extra parameters. (§2.4.2)

A further complication with multiwavelength dispersed (sub)picosecond measurements is the dependence of the IRF upon the detection wavelength. This wavelength-dependent group velocity delay (dispersion) can be described by a polynomial function for the IRF location parameter (μ in Eq. 3, §2.4.2.1), of which the parameters must be estimated (from the data, or from a separate measurement, e.g. of the cross-phase modulation [33]).

2.1.2 Stochastics

The stochastics of spectroscopic measurements originate from photon properties. Single photon timing fluorescence measurements constitute a counting process which is Poisson distributed, where the variance is equal to the mean and all observations are independent. In contrast, the stochastic properties of the other types of measurements are much more uncertain. The observations are assumed to contain additive normally distributed noise. In general these observations are also assumed to be statistically independent, which seems justified because the measurements are done sequentially. There is one exception: with time-gated spectra a whole spectrum is observed simultaneously, and in principle the responses measured at different wavelengths could be statistically dependent with (unknown) covariance matrix Σ (independent of time) [34]. With independent measurements there are several possible cases for the usually unknown variance σ^2 which may in principle depend upon time t and wavelength λ. (a) constant variance $\sigma^2(t, \lambda) = \sigma^2$ which is the most common assumption. (b) wavelength-dependent variance $\sigma^2(t, \lambda) = \sigma^2(\lambda)$ which is appropriate with difference absorption measurements. In general, $\sigma^2(\lambda)$ needs to be estimated as well. This procedure is called *iteratively reweighted least squares* [35]. (c) in addition to wavelength-dependence there may be time dependence $\sigma^2(t, \lambda)$, because the time interval of the measurement, and thus the signal to noise ratio, may increase with time [36]. (d) For large numbers of counts the above mentioned Poisson distributed single photon timing measurements are usually well approximated by a normal distribution with the variance equal to the mean.

2.2 Model assumptions

2.2.1 Homogeneity

A classical problem in describing reaction dynamics is (in)homogeneity [37, 38]. The common assumption is that the properties of the system studied are homogeneous, which means that a discrete set of parameters describes the whole system. The observed dynamics of the ensemble can be ascribed to the dynamics of each individual member of that ensemble. In some cases there are indications that subpopulations are present [36,39] and it is assumed that such a system can be represented by a weighted average of homogeneous subsystems. When

many subsystems are present this can be described by a model with distributed parameters. A frequent discussion is the possibility to distinguish between models with discrete parameters and models with distributed parameters [42,43]. Biophysical knowledge of e.g. a distribution of protein conformations necessitates the use of models with distributed parameters. However, nearly all models used in practice lump parameters into a discrete set. An extra complication is that the two types of model can only be distinguished experimentally with very high signal to noise ratios, or when measurements are done over many orders of time, or as a function of temperature, or as a function of excitation wavelength.

Unless noted otherwise, we will assume that we are dealing with a homogeneous system which can be modelled with discrete parameters.

2.2.2 Separability

The spectroscopic properties of a mixture of components are a superposition of the spectroscopic properties of the components weighted by their concentration. With absorption this is known as the Beer-Lambert law. Strictly speaking, this law is an approximation which only is valid with small absorbance changes. Likewise, with fluorescence it is assumed that the sample is optically thin and self absorption is negligible [44]. Thus the noise-free, time-resolved spectrum $\psi(t, \lambda)$ is a superposition of the contributions of the n_{comp} different components:

$$\psi(t, \lambda) = \sum_{l=1}^{n_{comp}} c_l(t)\varepsilon_l(\lambda) \tag{1}$$

where $c_l(t)$ and $\varepsilon_l(\lambda)$ denote, respectively, the concentration and spectrum of component l. Typical values for the number of components which can be studied with time-resolved spectroscopy are $1 \leq n_{comp} \leq 10$, whereas both the number of different wavelengths and the number of different time instants can vary from n_{comp} to thousands.

Note that according to Eq. 1 a separability of time and wavelength properties is possible. However, with ultrafast measurements there is a caveat: the properties of the detector system are in general wavelength-dependent on a (sub)picosecond time scale (§2.1.1), thus with wavelength-dependent parameters $\theta(\lambda)$ the model for the concentration reads $c_l(t, \theta(\lambda))$. Regarding Eq. 1 we note that the quantity which will be estimated is the product $c_l\varepsilon_l$ which in itself is insufficient for the determination of the absolute values of c_l and ε_l. Thus when we

have e.g. a kinetic model for c_l and no additional information, we can only identify the parameters which determine the shapes of c_l and ε_l. With a detailed kinetic model sometimes the relative concentrations of the components can be estimated, and thus also the relative amplitudes of their spectra. We will return to this indeterminacy in §2.6.2.

2.3 Inverse problem

Measurement of ψ poses the inverse problem: recovery of the spectroscopic and kinetic (dynamic) properties of the components from the data. In practice various problems can arise: firstly, the number of components present in the system may be unknown. Secondly, in general neither the concentration profiles $c_l(t)$ nor the spectra $\varepsilon_l(\lambda)$ are known. However the experimentalist usually has a priori knowledge about which shapes of concentration profiles or spectra are realistic. This amounts to common statements regarding continuity, non-negativity, unimodality (single peak), etc. A large amount of research is based upon such physical constraints in the self-modelling of two-way data [45, 46]. Self-modelling was applied to the bacteriorhodopsin photocycle [47,48,49].

In many cases more knowledge is available in the form of a parameterized model for $c_l(t)$, termed a kinetic model (§2.4), or for $\varepsilon_l(\lambda)$, termed a spectral model (§2.5). Still, there may be several candidate kinetic models available, and dependencies in a kinetic model may render parameters unidentifiable. In the following two sections we present often well-known models in detail. In §2.6 a fundamental identiability problem is discussed. Finally in §2.7 the more powerful spectrotemporal models are introduced, which provide a solution to this problem.

2.4 Kinetic models

A first distinction to be made is the order of the kinetics. In case the concentrations are described by linear differential equations we are dealing with first-order kinetics. The solution of a system of linear differential equations is given by a sum of exponential decays convolved with the IRF (§2.1.1). When the differential equations contain a product of concentrations terms we are dealing with second-order kinetics [50, 51]. Multiphoton processes taking place during the short exciting laser pulse can approximately be described by direct population of a multiply excited state, which subsequently obeys first-order kinetics. In the following we will restrict ourselves to first-order kinetics, but many aspects of these methods are also applicable with more complex kinetics.

2.4.1 Global analysis

Without a priori knowledge about a detailed kinetic model the first step is to fit the data with a sufficient number of exponential decays and their amplitudes [52], which constitute the Decay Associated Spectra (DAS) [12,13,53,54,55]. Note that this number can be larger than the number of spectrally different components present. Subsequently, the DAS can be fitted with a spectral model [34]. DAS are most common with fluorescence or absorption spectroscopy. With *difference* absorption spectroscopy the amplitudes associated with exponential decays are termed Decay Associated Difference Spectra (DADS) [16]. When the IRF width is negligible the model reads $c_l(t) = \exp(-k_l t)$ with decay rate parameter k_l. Otherwise, as explained in §2.1.1, with ultrafast measurements the exponential decay has to be convolved with the IRF. Thus in global analysis the model function reads:

$$\psi(t, \lambda) = \sum_{l=1}^{n_{\text{comp}}} (\exp(-k_l t) \oplus i(t)) DAS_l(\lambda) \tag{2}$$

where $\oplus$ indicates convolution of the exponential decay with the IRF $i(t)$.

2.4.2 Convolution of exponential decay with IRF

When the IRF is smooth and can be described with an analytical function, the convolution of this IRF with an exponential decay often can be done analytically as well. In this section we present the most important applications, starting with the Gaussian shaped IRF (§2.4.2.1). Note that an IRF can also be a linear combination of simple IRF's (§2.4.2.4).

2.4.2.1 Gaussian shaped IRF

Often the IRF $i(t)$ can be well-described by a Gaussian (§2.1.1) with two parameters for the location (mean) μ and the full width at half maximum (FWHM) Δ:

$$i(t) = \frac{1}{\tilde{\Delta}\sqrt{2\pi}} \exp(-\log(2)(2(t-\mu)/\Delta)^2) \tag{3}$$

where $\tilde{\Delta} = \Delta/(2\sqrt{2\log(2)})$. The convolution of this IRF with an exponential decay results in an analytical expression which facilitates the estimation of the IRF parameters μ and Δ (which is often necessary):

$$c(t;k, \mu, \Delta) = \exp(-kt) \oplus i(t) =$$

$$= \frac{1}{2}\exp(-kt)\exp\left(k\left(\mu + \frac{k\tilde{\Delta}^2}{2}\right)\right)\left\{1 + \mathrm{erf}\left(\frac{t-(\mu + k\tilde{\Delta}^2)}{\sqrt{2}\tilde{\Delta}}\right)\right\} \quad (4)$$

where $\oplus$ indicates convolution. Note that with dispersion μ can be described by a polynomial:

$$\mu(\lambda;\mu_0, \mu_1, ..., \mu_p) = \mu_0 + \sum_{l=1}^{n_{\mathrm{disp}}} \mu_l\left(\frac{\lambda-\lambda_c}{100}\right)^l \quad (5)$$

where n_{disp} is the order of the dispersion, λ_c is a center wavelength for the polynomial, and the division by 100 is convenient for the scale of the dispersion parameters μ_l.

A complication arises with fluorescence measurements by a synchroscan streak camera [28, §2.1c] because of additional contributions to Eq. 4 due to long lived components [56]. This will be treated in §2.4.2.3.

It is noteworthy that Eq. 4 also appears in the context of chromatography, where an ideal Gaussian elution profile is convolved with an exponential decay that is caused by e.g. dead volume disturbances [57]. Consequently, global analysis methods can also be used with wavelength-resolved chromatography measurements (hyphenated measurements) [58] (unpublished results).

2.4.2.2 Measured IRF

With single photon timing data the IRF usually is a complicated function that cannot easily be described with analytical functions. The problems involved in measuring the IRF were discussed in §2.1.1. Assuming an IRF has been measured sufficiently precisely, then a wavelength-dependent delay parameter μ is often used to account for the wavelength dependence of the instrument response:

$$c(t;k, \mu) = \int_0^t \exp(-ks)i(t-s-\mu)ds \quad (6)$$

where now $i(t)$ represents the measured IRF. Note that the stochastic character of the measured IRF is neglected. A convolution algorithm for Eq. 6, that takes into account the decay within a time bin which is relevant with short lifetimes, is described in [59].

2.4.2.3 IRF of synchroscan streak camera

Here we derive the expression used for describing the contribution of an exponentially decaying component to fluorescence recorded with a synchroscan streak camera. Apart from time and wavelength calibration, the CCD image has to be corrected [28] for curvature (the wavelength-dependence of the IRF location due to the instrumental geometry), dark current, shading (sensitivity of the CCD pixels), and wavelength sensitivity. Furthermore it might require tilt correction. Clearly, a lot of preprocessing is needed, and when there is no correction for instrumental curvature, this has to be taken into account in the global analysis as dispersion, which is suboptimal.

The width of $i(t)$ and thus the temporal resolution is limited by the point spread function of the imaging device and by the temporal spread resulting from the monochromator. When using the fastest time base of ~200 ps the temporal resolution is typically 3 ps FWHM (which corresponds to 15 pixels on the CCD). With an 0.8 or 2.2 ns time base the temporal resolution is typically 8 or 20 ps FWHM, which corresponds to about 1% of the full time range, indicating that it is dominated by the point spread function of the imaging device.

The periodicity of the synchroscan results in detection of the remaining fluorescence after multiples of half the period T of the laser (typically about 13 ns). Thus for lifetimes longer than 1 ns a sum of exponential decays resulting from forward and backward sweeps has to be added to Eq. 4:

$$\sum_{n=0}^{\infty} \exp(-nkT)(\exp(-k(t-\mu+T)) + \exp(-k(T/2-t-\mu))) =$$

$$\frac{(\exp(-k(t-\mu+T)) + \exp(-k(T/2-t-\mu)))}{1-\exp(-kT)} \tag{7}$$

Note that we assume here that time zero of the time base corresponds to the zero crossing of the sweep, and that the IRF shape is no longer important after T/2. Adding Eq. 7 to Eq. 4 the full concentration profile reads:

$$c(t, k, \mu, \Delta, T) \;=\; \frac{1}{2}\exp(-kt)\exp\!\left(k\!\left(\mu + \frac{k\tilde{\Delta}^2}{2}\right)\right)\!\left\{1 + \mathrm{erf}\!\left(\frac{t - (\mu + k\tilde{\Delta}^2)}{\sqrt{2}\tilde{\Delta}}\right)\right\} +$$

$$\frac{(\exp(-k(t - \mu + T)) + \exp(-k(T/2 - t - \mu)))}{1 - \exp(-kT)} \tag{8}$$

With dilute samples the Raman scattering of water complicates the measurement. It can be modelled as an extra component contributing with a concentration time course identical to the instrument response $i(t)$:

$$c_{\mathrm{pulse}}(t, \mu, \Delta) \;=\; i(t, \mu, \Delta) \tag{9}$$

The adequacy of the Gaussian shaped IRF is shown in a fit of experimentally measured scatter (Fig. 3).

2.4.2.4 Composite Gaussian shaped IRF

Although the Gaussian approximation of the IRF is most often satisfactory, sometimes a more complicated IRF is needed. With fast processes (like ps lifetimes, or scatter) the IRF is sampled well, which means that the measured response will closely resemble the IRF. Thus, slight deviations from the assumed Gaussian shape will cause inadequacy of the fit. Origins of these deviations may be time jitter caused by incorrect synchronization, or non-Gaussian shaped leakage of current in the CCD of the detector. With slow processes slight deviations would be smeared out, and practically invisible. To deal with this problem a second Gaussian with location μ and width Δ_2, and relative amplitude a_2 can be added:

$$c_{\mathrm{pulse},2}(t, \mu, \Delta, a_2, \Delta_2) \;=\; i(t, \mu, \tau) + a_2 i(t, \mu, \Delta_2) \tag{10}$$

Consequently, the convolution is a linear combination of two responses described by Eq. 8:

$$c_2(t, k, \mu, \Delta, T, a_2, \Delta_2) \;=\; c(t, k, \mu, \Delta, T) + a_2 c(t, k, \mu, \Delta_2, T) \tag{11}$$

A further complication can arise due to a small reflection (relative amplitude a_2) which arrives after a delay μ_2 relative to the exciting laser pulse (Fig. 3). The full concentration profile can then be modelled as:

$$c_2(t, k, \mu, \Delta, T, a_2, \mu_2) \;=\; c(t, k, \mu, \Delta, T) + a_2 c(t, k, \mu + \mu_2, \Delta, T) \tag{12}$$

or of the scatter only (see Fig. 3):

$$c_{\text{pulse},2}(t, \mu, \Delta, a_2, \mu_2) = i(t, \mu, \Delta) + a_2 i(t, \mu + \mu_2, \Delta) \qquad (13)$$
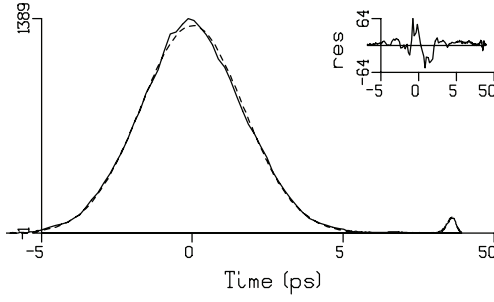

Fig.3. IRF of streak scope fitted with a Gaussian. Estimated FWHM $\Delta = 4$ ps, note a small (7%) reflection after 26 ps.

Further complications can arise with combined deviations from the assumed Gaussian shape. When calibrating the time base an etalon is used, resulting in a train of N pulses:

$$c_{\text{pulse,cal}}(t, \mu, \Delta, a_2, \mu_2, \ldots, a_N, \mu_N) = i(t, \mu, \Delta) + \sum_{n=2}^{N} a_n i(t, \mu + \mu_n, \Delta) \qquad (14)$$

A prototype PSE has been developed at the VU over the past fifteen years [61]. This program contains various numbered model options, referred to here as timmodel#. The Gaussian shaped IRF is used in timmodel 1,14,16,22. Relevant options are lstreak (Eq. 8), ldoublegaussian (Eq. 10, Eq. 11), lsecondresponse (Eq. 12), lpulse (Eq. 9), lsecondpulse (Eq. 13), lmultipulse (Eq. 14).

2.4.2.5 Non-Gaussian shaped IRF

An exponentially decaying IRF can result from an integrating detector. It arises in e.g. Electron Spin Resonance spectroscopy [60]. (timmodel 4):

$$i(t, \mu, \tau) = \exp(-(t - \mu)/\tau) \qquad t \geq \mu \qquad (15)$$

where τ is the integration time of the detector. The convolution with an exponential decay results in

$$c(t, k, \mu, \tau) = \exp(-k(t - \mu)) - \exp(-(t - \mu)/\tau) \qquad (16)$$

Alternatively, a gamma-shaped IRF can sometimes be used (timmodel 6):

$$i(t, \mu, \tau) = (t - \mu)^2 \exp(-(t - \mu)/\tau) \qquad (17)$$

The applicability of the use of an analytically described instrument response in iterative reconvolution depends upon the availability of a suitable analytical description. A versatile candidate function is (a linear combination of) Gamma-functions

$a((t - \mu)/(\tau\gamma))^\gamma \exp(\gamma - (t - \mu)/\tau)$ with parameters a, γ, μ, τ for amplitude, shape, location, width. Convolution of this function with an exponential decay results in the confluent hypergeometric function [62].

2.4.3 Compartmental models

When a priori knowledge about a detailed kinetic model is available, a linear time-invariant compartmental model [63,64] can be used. Because in contrast to global analysis a specific kinetic model is tested, this is often termed *target* analysis [12,16,65,66]. The *target* here is to describe the real concentrations of the components. Note that the global analysis is equivalent to fitting with a kinetic model comprised of a number of non-interacting, parallelly decaying compartments. An important question is whether all unknown kinetic parameters can be estimated from the data. To answer it, a first step is detection of structural (un)identifiability [16,67], which is caused by incomplete information on the system. When different compartmental schemes result in the same model output, the system is structurally unidentifiable. But even a structurally identifiable model may be numerically unidentifiable [64].

Transitions between compartments are described by microscopic rate constants which constitute the off-diagonal elements of the transfer matrix K. The diagonal elements of K contain the total decay rates of each compartment. The concentrations of each compartment are described by a vector $c(t) = \begin{bmatrix} c_1(t) & \dots & c_{n_{\text{comp}}}(t) \end{bmatrix}^T$. Thus, a linear compartmental model with n_{comp} compartments is described by a differential equation for these concentrations:

$$\frac{d}{dt}c(t) = Kc(t) + j(t) \tag{18}$$

where the input to the system is described by a vector $j(t) = i(t)\begin{bmatrix} 1 & x_2 & \dots & x_{n_{\text{comp}}} \end{bmatrix}^T$, with $i(t)$ the IRF (§2.1.1) and x_l representing a possible extra input to compartment l. Eq. 18 can be solved analytically, which is important for both insight into the problem and for computational speed.

We assume that all eigenvalues of the transfer matrix K are different, and that $c(-\infty) = 0$. The solution of Eq. 18 is then given by $c(t) = e^{Kt} \oplus j(t)$ where $\oplus$ indicates convolution.

For a diagonal K-matrix ($K = \text{diag}(-k_1, \dots, -k_{n_{comp}})$) with all inputs x_l equal to 1, the concentration matrix C consists of elements $[C_I]_{pq} = c_q^I(t_p, k_q) = \exp(-k_q t_p) \oplus i(t)$. The sub- or superscript I indicates that this is Model I which is comprised of independently

decaying compartments, and also is called the parallel model.

For the evaluation of the exponential of a non-diagonal K matrix we use the eigenvector-eigenvalue decomposition $K = U\Lambda U^{-1}$. With $\Lambda = \mathrm{diag}(-k_1, \ldots, -k_{n_{comp}})$ we have $e^{Kt} = Ue^{\Lambda t}U^{-1}$ and

$$e^{Kt} \oplus j(t) = U\mathrm{diag}\left(U^{-1}\begin{bmatrix}1 & x_2 & \ldots & x_{n_{comp}}\end{bmatrix}^T\right)\begin{bmatrix}e^{-k_1 t} \oplus i(t) & \ldots & e^{-k_{n_{comp}} t} \oplus i(t)\end{bmatrix}^T \equiv A_{II}^T C_I^T \quad (19)$$

Thus the solution of the general compartmental model is a linear combination of the c_l^I and therefore a transformation of C_I (derived from the eigenvalues of the transfer matrix K) for which we can write

$$C_{II} = C_I A_{II} \quad\quad (20)$$

with $A_{II}^T = U\mathrm{diag}\left(U^{-1}\begin{bmatrix}1 & x_2 & \ldots & x_{n_{comp}}\end{bmatrix}^T\right)$. Note that a compartmental model is closely related to the state space representation in mathematical systems theory, with the vector of concentrations of compartments being analogous to the state vector [68].

2.4.3.1 Simulation of a simple two-compartment model

To illustrate the previous section, a simple two-compartment model has been simulated, using realistic parameters [34]. In this model the first compartment, representing component 1, is excited. Component 1 irreversibly decays with rate k_1 thereby forming component 2 with fractional yield Φ_{12}. Component 2, which is represented by the second compartment, decays with rate k_2, which is smaller than k_1. This kinetic scheme is depicted in the right inset of Fig. 4, and the concentrations are shown in Fig. 4D. For this scheme the transfer matrix K is of lower triangular form:

$$K_{II} = \begin{bmatrix} -k_1 & 0 \\ \Phi_{12}k_1 & -k_2 \end{bmatrix} \quad\quad (21)$$

The subscript II indicates that this is Model II, which is abbreviated $1 \to 2$. Recall that Model I, which was defined in §2.4.3, consists of independently decaying compartments (which is abbreviated $1|2$), and is depicted in the left inset of Fig. 4. We now have for the matrix A_{II} of Eq. 20:

$$A_{II} = \begin{bmatrix} 1 & -\alpha \\ 0 & \alpha \end{bmatrix} \quad\quad A_{II}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & \alpha^{-1} \end{bmatrix} \quad\quad (22)$$
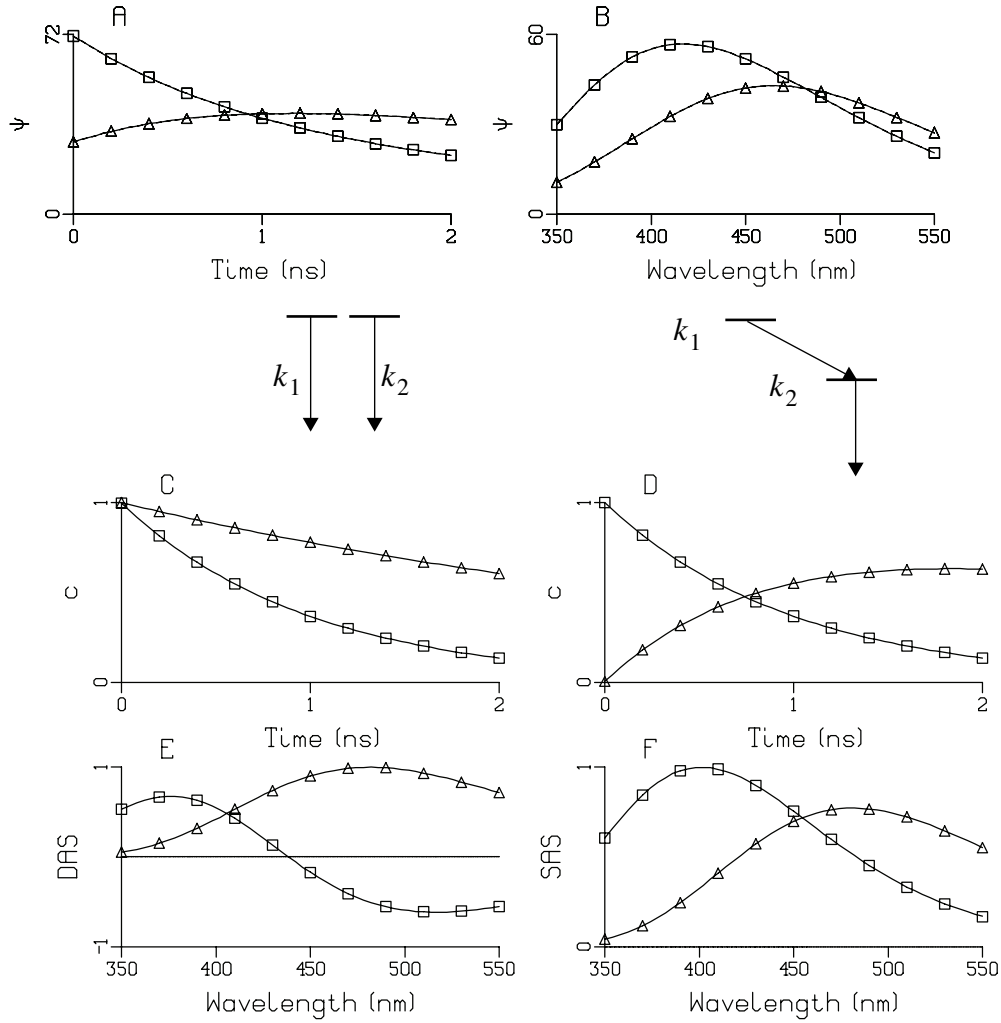
Fig.4. Global analysis of simulated data from a two-compartment model with kinetic scheme $1 \rightarrow 2$ (right inset). The first component (indicated by squares in panel D,F) decays in 1 ns, thereby forming the second component (indicated by triangles, life time 4 ns). (A) Decay traces at 400 and 500 nm (indicated by squares and triangles). (B) Time-gated spectra at 0.4 ns (squares) and 1.6 ns (triangles). (C,E) $c(t)$ and estimated DAS using the incorrect parallel scheme $1|2$ (left inset). (D,F) $c(t)$ and estimated SAS using the correct sequential scheme. with $\alpha = \Phi_{12} k_1 / (k_1 - k_2)$. A time-resolved (fluorescence or absorption) spectrum $\psi(t, \lambda)$ is simulated by a superposition of the concentrations of the components multiplied by their differing spectra (Fig. 4F) according to Eq. 1 (to improve readability we suppress here the time respectively wavelength-dependence of $c_i(t)$, $\varepsilon_i(\lambda)$)

$$\psi = c_1^{II} \varepsilon_1^{II} + c_2^{II} \varepsilon_2^{II} \tag{23}$$

Typical traces and spectra are depicted in Fig. 4A,B.

Now these simulated data can be fitted in two different ways, using Model I or II. When using Model II, with the correct kinetic scheme $1 \to 2$, and with $k_1 > k_2$, the estimated amplitudes associated with the concentrations c_l^{II} are called Species Associated Spectra (SAS) [12,53,55]. The estimated SAS in Fig. 4F are indistinguishable from the simulated component spectra ε_l^{II}. By contrast, when these data are analysed using the incorrect kinetic scheme Model I with two independent decays c_l^I the Decay Associated Spectra (DAS) ε_l^I depicted in Fig. 4E are estimated. Using Eq. 20 and Eq. 22 we find $c_1^{II} = c_1^I$, $\varepsilon_1^{II} = \varepsilon_1^I + \varepsilon_2^I$, $c_2^{II} = \alpha(c_2^I - c_1^I)$ and $\varepsilon_2^{II} = \varepsilon_2^I / \alpha$, which means that the difference between the two models is that the concentration of the formed component is proportional to the difference between the two decays, whereas the spectrum of the precursor is the sum of the two DAS (cf. Fig. 4E and F). The other way around, the DAS of the precursor is a linear combination of the SAS, $\varepsilon_1^I = \varepsilon_1^{II} - \alpha\varepsilon_2^{II}$ ($= \varepsilon_1^{II} - \varepsilon_2^I$, cf. Fig. 4E and F). Therefore it can contain negative amplitudes when the kinetics obey model II and the SAS overlap. Observation of negative amplitudes in a DAS indicates the presence of an (excited state) reaction [12], like in model II. This Model II has been successfully applied with bridged electron donor-acceptor systems where after excitation the charge-transfer state exhibits conformational dynamics, which can be observed by a red shift of the emission spectrum [34, 69]. This case study will be presented in §2.6.4.

2.4.3.2 The unbranched, unidirectional model

Apart from the global analysis with independent decays $(1|2|\ldots|n_{comp})$, the simplest kinetic scheme is the unbranched, unidirectional model $(1 \to 2 \to \ldots \to n_{comp})$. These models are also termed parallel and sequential, and correspond to the generalization of the models I and II of the previous section. In the sequential model back-reactions are ignored on the assumption that the energy losses are large enough that the reverse reaction rates are negligible. Note the assumption that there are no losses in the chain $1 \to 2 \to \ldots \to n_{comp}$. The compartmental model can be solved to yield [19]:

$$c_l(t) = \sum_{j=1}^{l} b_{jl}\exp(-k_j t) \oplus i(t) \tag{24}$$

where k_j is the decay rate of compartment j and the amplitudes b_{jl} of the (convolved) exponential decays are defined by $b_{11} = 1$ and for $j \leq l$:

$$b_{jl} = \prod_{m=1}^{l-1} k_m \Big/ \prod_{\substack{n=1 \\ n \neq j}}^{l} (k_n - k_j) \tag{25}$$

In particular, for $j < l$, $b_{jl} = b_{j,l-1}k_{l-1} / (k_l - k_j)$.

Of course, hybrids of the generalized models I and II, containing a mixture of parallelly and sequentially decaying compartments, can also easily be solved.

2.4.3.3 Multiexponential decays

Multiexponential decay of a component can originate from inhomogeneity (§2.2.1) or from equilibria. With equilibria, the number of compartments involved determines the degree of multiexponentiality. Take as an extreme case a model with just one emitting component P which is reversibly coupled to a number of other non-emissive states $h\nu \leftarrow P \leftrightarrow Q \leftrightarrow R$. When this three-compartmental system is excited, the decay of P shows a three-exponential decay. Since P is the only fluorescing component, its concentration profile is directly observed which offers possibilities for determining (relations between) the other rate constants in this intricately coupled system. Note that in this example only one component is observed, but three compartments are needed to describe its kinetics. The alternative kinetic scheme, $Q \leftrightarrow P \leftrightarrow R$, where the emissive state P is connected to non-emissive states Q and R, also results in three-exponential decay of P. Thus when only P-emission is observed, the system is structurally unidentifiable. An example is the case of the bacterial photosynthetic reaction centre, for which the emission from the excited primary donor P* is assumed to be in equilibrium with non-emitting relaxed radical pair states [71]. Such radical pair states are directly observed in transient absorption. In an elaborate target analysis of transient absorption from photosystem I reaction centres [72] many different kinetic schemes were considered. The final scheme describing the energy transfer and charge separation kinetics also contained an equilibrium between the excited primary donor and a radical pair. In turn, the primary donor is in equilibrium with two energy transferring antenna pigment pools. In this way equilibria describe the four exponential decay of the two antenna pools, primary donor and radical pair state. These equilibria quantitatively describe the free energy differences in this complex system.

2.4.3.4 Case study: analysing a P decay trace*

The aim of this section is to show how a single trace can be analysed using one or two exponentials, or a distribution. Beekman et al. [43] studied the kinetics of primary electron transfer in the bacterial reaction centre monitoring the difference absorption at 918 nm from RC only mutants, after exciting at 800 nm, using the pump-probe technique for optimal time resolution. The signal shown in this section has been inverted; it represents mainly the decay of stimulated emission from P*. The description of this kinetics requires more than just one exponential. Possible explanations for this multiexponentiality have been discussed in §2.4.3.3. Here we present an in depth analysis of one of the decay traces measured. The analysis of just the decay part using Provencher's programs DISCRETE [40] and CONTIN [41] is shown in Fig. 5 (these automated programs cannot deal with an IRF, effectively
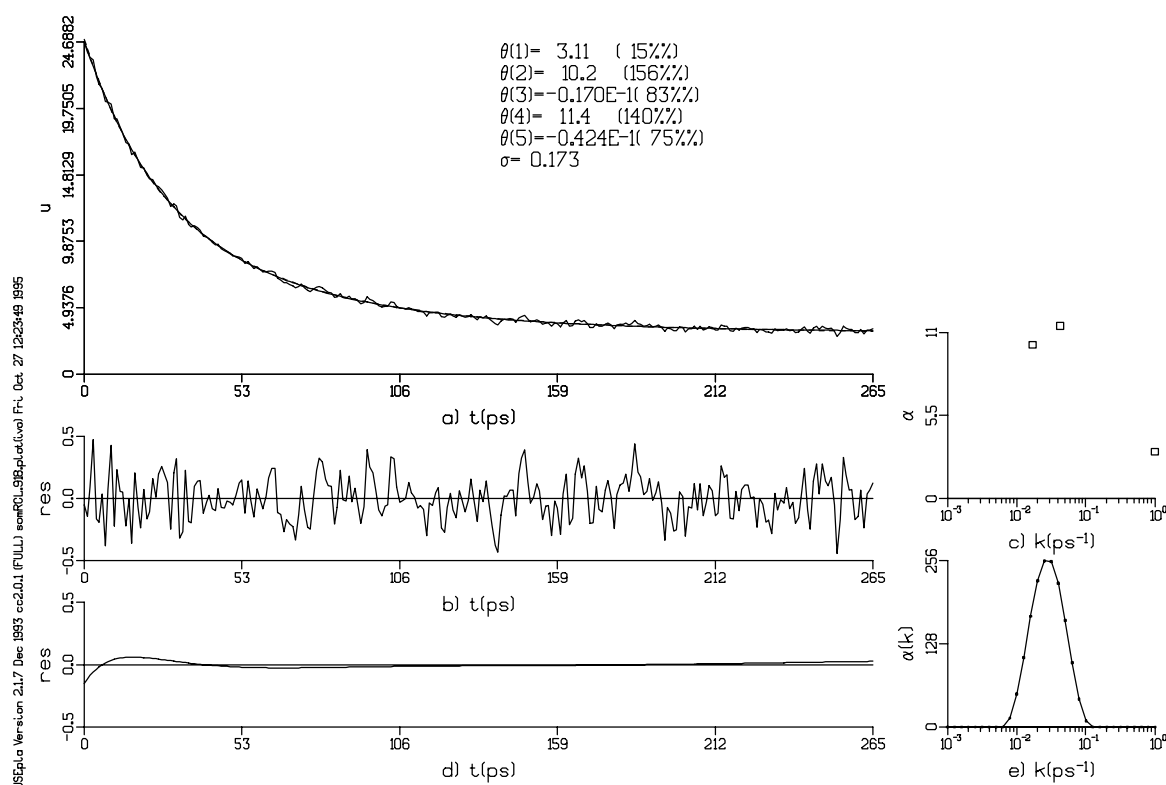
Fig.5. Fit of decay (part of) trace (a) with discrete sum of exponential decays (b,c), and with a continuous distribution of exponential decays (d,e). The residuals of the DISCRETE fit are shown in b and the difference between the residuals from a continuous and discrete fit are depicted in d. Amplitudes α and rate k of the exponential plus "offset" (square at far right) fit are shown in c, and the distribution $\alpha(k)$ is depicted in e.

assuming a δ-function IRF). Clearly an "offset" is present, which is attributed to a (charge separated) state which lives much longer than 300 ps, the time scale of the experiment. When allowing for an "offset" the DISCRETE program decided that two exponentials were sufficient to fit the data (Fig. 5c). CONTIN favoured an unimodal, smooth distribution (Fig. 5e), with hardly any difference in the residuals (Fig. 5d). Of course we have to take into account the IRF, which we model as a Gaussian (§2.4.2.1). We compare fits of the complete decay with one and two exponentials according to Eq. 4 (plus the "offset") in Fig. 6. Note the
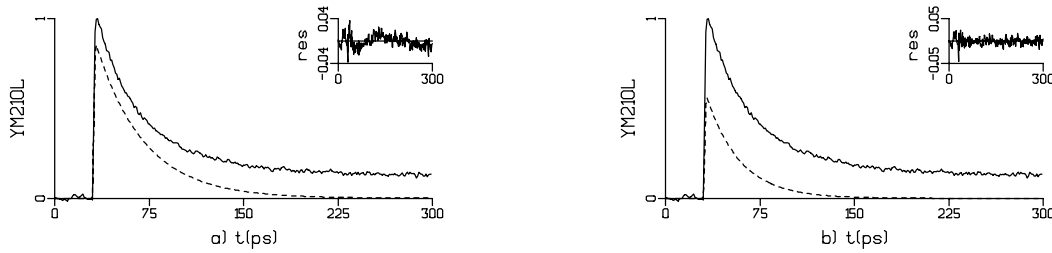
Fig.6. Fits of the complete decay with (a) one and (b) two exponentials (the first shown as dashed line) according to (4) (plus the "offset"). Note the difference in the residuals (insets).

trends in the residuals with the monoexponential fit (inset of Fig. 6a).

With inhomogeneous kinetics (§2.2.1) a decay trace can also be fitted with a parameterized distribution, for instance a Gaussian on the (natural) $\log(k)$ scale with location k_0 and width σ [43]:

$$\int_{-\infty}^{\infty} \exp(-(\log(k) - \log(k_0))^2/(2\sigma^2))(\exp(-kt) \oplus i(t))d\log(k) \tag{26}$$

In view of the Gaussian-looking distribution found by CONTIN we are not surprised to find again a good fit (Fig.7.a). Note that it is essential here to fit the whole decay including the rise, because the distribution changes with a shift of time "zero" (which in this case is the location μ of the instrument response. The advantage of using a distribution instead of two exponentials is that one parameter less is needed (principle of parsimony) (distribution: location k_0, width σ, and amplitude makes three parameters; two exponentials plus two amplitudes makes four parameters). Note the resemblance between the CONTIN distribution (Fig.5.e) and the right hand side of Fig.7.

Fig.7. Fit of decay trace with distribution of exponentials convolved with instrument

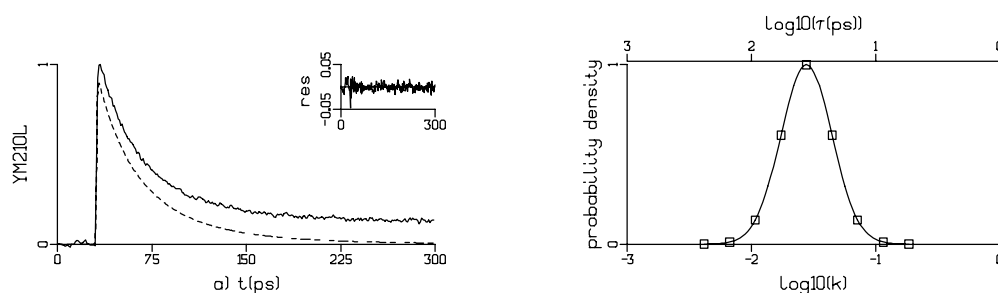response (Eq. 26, dashed line). Inset of (a) shows residuals. At the right hand side the

fitted distribution is depicted, with the squares placed at k_0 (top), $k_0 \pm \sigma$, $k_0 \pm 2\sigma$, ...

Alternatively, a non-parametric distribution of lifetimes (or decay rates) can be estimated (e.g.

Fig. 5e). Fits of single traces with distributions of lifetimes [41] have been reviewed in [42],

whereas lifetime density maps of time-resolved spectra have been presented in [72, 73].

Subsequent target analysis can be performed on the kinetics obtained from the lifetime density

maps [72].

2.4.3.5 "Invisible" compartments

A compartment can be spectrally invisible, e.g. because the species represented by the

compartment does not emit or absorb light. In difference absorption spectroscopy,

compartments may possess indistinguishable spectral properties (giving rise to spectrally

silent transitions [74]). Or there may exist linear relations between spectra of the

compartments. Then the number of spectrally and temporally different components whose

properties can be estimated will be less than the number of compartments. The *C*-matrix of

these components can be obtained by postmultiplication in Eq. 20 of C_I by a modified A_{II}.

Alternatively, this can be described by a spectrotemporal model using spectral equalities or

constraints (see §2.7.2).

2.4.3.6 Measurements at different temperatures

The temperature dependence of microscopic rate constants can be described with a model

containing thermodynamic parameters. Measurements at a greater number of different

temperatures than the number of unknown thermodynamic parameters will provide extra

information. This offers an opportunity to identify and estimate both forward and backward microscopic rate constants [20, 21], thus enabling estimation of free energy differences. In order to fit such measurements globally, a target analysis using a detailed compartmental model is mandatory. The temperature dependence of a microscopic rate constant k can be described in three different ways. The first is an Eyring relationship [20, 24, 37]

$$\ln(K^{\#}) = \frac{\Delta S^{\#}}{R} - \frac{\Delta H^{\#}}{RT} \tag{27}$$

where $K^{\#} = (kh)/(k_B T)$ is the activation equilibrium constant and $\Delta S^{\#}$, $\Delta H^{\#}$ are the entropy, and enthalpy changes of activation. k_B, h, R are the Boltzmann, Planck and gas constant, respectively. Alternatively, an Arrhenius relationship

$$k = k_0 e^{-E_A / RT} \tag{28}$$

has been used, e.g. in modelling the bacteriorhodopsin photocycle [21].

To explain the anomalous temperature dependence of the recovery rate in the photoactive yellow protein photocycle Van Brederode et al. [75] introduced a heat capacity change of activation parameter $\Delta C_p^{\#}$:

$$\ln(K^{\#}) = \frac{\Delta S^{\#}(T_0)}{R} - \frac{\Delta H^{\#}(T_0)}{RT} - \frac{\Delta C_p^{\#}}{R}\left(1 - \frac{T_0}{T} + \ln\left(\frac{T_0}{T}\right)\right) \tag{29}$$

where T_0 is a reference temperature.

2.5 Spectral models

When the spectral resolution of two-way spectral data is high, and a parameterized spectral model is available, analysis with a spectral model can be appropriate. In particular when the time resolution is low, or when systematic errors like time jitter of time-gated spectra (§2.1a) are present, global analysis using a spectral model is appropriate [69].

Spectral models are more phenomenological than kinetic models. In general they require more parameters than kinetic models. With difference absorption spectroscopy the ground state spectrum can be included in the spectral model. Analogous to global analysis with single exponential decays, two-way spectral data can be analysed globally. Without a priori knowledge about detailed component spectra the first step is to fit the data with a sufficient number of band shapes and their amplitudes, the band amplitude curves (BAC). With time-resolved spectra, subsequently the BAC can be fitted with a kinetic model [69]. With steady-

state spectra measured as a function of temperature, subsequently the BAC can be fitted with a thermodynamic model [76]. When there are linear dependencies in the BAC's, and the number of bands used is larger than the number of spectrally and temporally different components, a combination of (some of the) band shapes may lead to component spectra. This is analogous to the target analysis with compartmental models. Now the targets are the spectra of the real components.

2.5.1 Spectral shapes

The absorption spectrum for a homogenously broadened (no inhomogeneity) transition can be well-described by a Lorentzian bandshape, whilst for an inhomogeneously broadened system (which is the case in most proteins and glasses), the corresponding spectrum is well-described by a Gaussian [38,77]. Often times, the measured spectrum (both absorption and fluorescence) of an ensemble of chromophores can be deconvolved into a superposition of such bands, with specified spectral bandshapes (e.g. Gaussian, Lorentzian, Voigt, skewed Gaussian, Pearson, ...). Thus the shape of a spectrum of a component is assumed to be a superposition of such standard bandshapes. Here we present model functions for the (skewed) Gaussian shape. The shape of a charge transfer fluorescence emission spectrum is often well-described by a Gaussian in the energy domain [78]:

$f(\bar{\nu})/\bar{\nu}^3 = f_{max}\exp(-\ln 2[2(\bar{\nu} - \bar{\nu}_{max})/\Delta\bar{\nu}]^2)$ where $\bar{\nu} = \lambda^{-1}$ denotes the wavenumber and $f(\bar{\nu})$ is the converted fluorescence emission spectrum: $f(\bar{\nu}) = \lambda^2\varepsilon(\lambda)$ [79]:

$$\varepsilon(\bar{\nu}) = \bar{\nu}^5 f_{max}\exp(-\ln 2[2(\bar{\nu} - \bar{\nu}_{max})/\Delta\bar{\nu}]^2) \tag{30}$$

Even better fits can be achieved when an extra skewness parameter is introduced [80,81]. Thus we arrive at the model function [34]

$$\varepsilon(\bar{\nu}) = \bar{\nu}^5 f_{max}\exp(-\ln 2[\ln(1 + 2b(\bar{\nu} - \bar{\nu}_{max})/\Delta\bar{\nu})/b]^2) \tag{31}$$

Note that with skewness parameter $b = 0$ the exponent in Eq. 31 reduces to the Gaussian of Eq. 30 (since $\lim_{b \to 0} (\ln(1 + bx))/b = x$). The maximum of Eq. 31 in the wavelength domain is given by the numerical solution of the nonlinear equation $(d\varepsilon(\bar{\nu}))/(d\bar{\nu}) = 0$. The Full Width at Half Maximum is given by $\Delta\bar{\nu}_{1/2} = \Delta\bar{\nu}\sinh(b)/b$.

Analogously the shape of a charge transfer absorption spectrum is often well-described by a Gaussian in the energy domain [78]: $f(\bar{\nu})/\bar{\nu} = \varepsilon_{max}\exp(-\ln 2[2(\bar{\nu} - \bar{\nu}_{max})/\Delta\bar{\nu}]^2)$. When the vibronic coupling, quantified as a Huang-Rhys factor, between the chromophore and its

vibrations is smaller than the magnitude of phonon-chromophore coupling, the resulting absorption spectrum can be well-described as a Gaussian [82]. Even when the two are comparable, the introduction of a skewness degree of freedom to the Gaussian can satisfactorily account for increased vibronic coupling. This leads to the model function:

$$\varepsilon(\bar{\nu}) = \bar{\nu}\varepsilon_{\max}\exp(-\ln 2[\ln(1 + 2b(\bar{\nu} - \bar{\nu}_{\max})/\Delta\bar{\nu})/b]^2) \tag{32}$$

In modelling an absorption difference spectrum the ground state spectrum ε_0 is subtracted from a linear combination of $\varepsilon(\bar{\nu})$ as in Eq. 32:

$$\varepsilon(\bar{\nu};(\bar{\nu}_{\max}, \Delta\bar{\nu}, b)_1, ..., (\bar{\nu}_{\max}, \Delta\bar{\nu}, b)_M, a_1, ..., a_M) =$$

$$\sum_{m=1}^{M} a_m\bar{\nu}\exp(-\ln 2[\ln(1 + 2b_m(\bar{\nu} - \bar{\nu}_{\max,m})/\Delta\bar{\nu}_m)/b_m]^2) - \varepsilon_0 \tag{33}$$

Note that analogous to the convolution with the measured instrument response Eq. 6, the stochastic character of this measured ground state spectrum is neglected.

2.6 Model for the observations in matrix notation

In many cases the data can be collected in a matrix Ψ of dimensions $m \times n$, where m and n are, respectively, the number of different time instants and wavelengths. The matrix element ψ_{ij} then contains the measurement at time instant t_i and wavelength λ_j. Using the matrix notation greatly simplifies the description of the model for the observations and allows the use of matrix decomposition techniques (§2.6.1). Assuming additive noise (§2.1.2), the basic model which describes the time evolution of spectra is the following:

$$\underline{\psi}_{t_i\lambda_j} = \sum_{l=1}^{n_{\text{comp}}} c_{lt_i}\varepsilon_{l\lambda_j} + \underline{\xi}_{t_i\lambda_j} \tag{34}$$

c_{lt_i} denotes the concentration of component l at time t_i, $\varepsilon_{l\lambda_j}$ denotes the contribution of component l at wavelength λ_j, and $\underline{\xi}_{t_i\lambda_j}$ denotes a normally distributed stochastic disturbance with zero mean (the underlining indicates that a variable is stochastic). The c_{lt_i} and $\varepsilon_{l\lambda_j}$ are collated in the matrices C and E, of dimension $m \times n_{\text{comp}}$ and $n \times n_{\text{comp}}$, respectively. The columns of C are the concentration profiles of the components, whereas the columns of E are the component spectra. Note that possible systematic errors are not included in this description (see §2.1). When there is a wavelength-dependent time delay we have

$$\underline{\psi}_{\lambda_j} = C\varepsilon_{\lambda_j} + \underline{\xi}_{\lambda_j} \tag{35}$$

and else we can write

$$\underline{\Psi} = CE^T + \underline{\Xi} \tag{36}$$

Matrix Ξ represents the noise and is, like Ψ, $m \times n$. ψ_{λ_j} and ξ_{λ_j} are the λ_j^{th} column of, respectively, Ψ and Ξ. ε_{λ_j} is the transpose of the λ_j^{th} row of E.

2.6.1 Estimation of the number of components

When there are no linear dependencies between the component spectra, and thus no linear relationships between the columns of E, the matrix E is of full rank. Analogously, when there are no linear dependencies between the concentrations of the components, and thus no linear relationships between the columns of C, the matrix C is also of full rank. This rank is equal to the number of components. Consequently, when $\text{rank}(C) = \text{rank}(E) = n_{\text{comp}}$ and with noise-free data ($\Xi = 0$), we have $\text{rank}(\Psi) = \text{rank}(CE^T) = n_{\text{comp}}$. Thus with experimental data of which we do not know the number of components, we can estimate this number by estimating the rank of $\underline{\Psi}$, using the Singular Value Decomposition [83,84,85,86,87,88,89] of $\underline{\Psi}$

$$\underline{\Psi} = \underline{U}\,\underline{S}\,\underline{W}^T \tag{37}$$

where U and W are orthogonal matrices, respectively $m \times m$ and $n \times n$, whose columns contain the left and right singular vectors. S is an $m \times n$ matrix which is zero except for its diagonal, which contains the singular values. With n_{comp} components and noise-free data we have exactly n_{comp} significant singular values: $s_1 \geq s_2 \geq \ldots \geq s_{n_{\text{comp}}} > s_{n_{\text{comp}}+1} = \ldots = 0$. The addition of the noise Ξ perturbs the SVD of the noise-free Ψ in two respects. Firstly, the singular values are changed. According to Corollary 8.3.2. of [83] an upper bound for this perturbation is given by the largest singular value of Ξ, say $s_{1,\Xi}$. Thus the perturbation is negligible when $s_l - s_{l+1} \gg s_{1,\Xi}$, $l = 1, \ldots, n_{\text{comp}}$. Secondly, the singular vector pairs (u_l, w_l), (u_{l+1}, w_{l+1}) are perturbed. Theorem 8.3.5. of [83] states that the amount of the perturbation depends upon the isolation of the relevant singular values. Thus when $s_l - s_{l+1} \approx s_{1,\Xi}$ the noise can greatly alter the singular vector pairs.

When Eq. 36 is applicable, the number of components can be determined from the number of singular vector pairs and accompanying singular values significantly different from noise. The

consequences of a wavelength-dependent time delay upon the rank are unclear, and need to be considered case by case.

2.6.2 Equivalence of spectral or kinetic models

Let us assume in the following that we have successfully determined n_{comp}. Starting from Eq. 36 we here distinguish two different kinds of parametrizations:

The concentrations are described by a kinetic model, which depends upon the parameters θ

$$\underline{\Psi} = C(\theta)E^T + \underline{\Xi} \tag{38}$$

The spectra are described by a parametric model, which depends upon the parameters θ

$$\underline{\Psi} = CE^T(\theta) + \underline{\Xi} \tag{39}$$

Subsequently, the estimated matrices E from Eq. 38 and C from Eq. 39 can be fitted with, respectively, a spectral and a kinetic model.

Because of the linearity of the model function the decomposition of Ψ into the product of two matrices is problematic, which is to a differring extent recognized by many authors (e.g. [12, 19, 20, 53, 88, 90, 91, 92]). Let A be an invertible matrix then:

$$CE^T = CAA^{-1}E^T = (CA)(EA^{-T})^T \tag{40}$$

Because we are dealing with a model function CE^T the *linear combinations of spectral or kinetic models will produce the identical residual matrix* $Z = \Psi - CE^T$. Thus with a linear kinetic model (like the compartmental model) the minimum of the least squares criterion is independent of the details of the model, as are the estimated lifetimes. From Eq. 40 we conclude that the difference between two compartmental models lies in their spectral parameters. A priori knowledge about E (for instance non-negativity of the $\varepsilon_l(\lambda)$) offers us the possibility to choose between alternative models. With a spectral model an analogous problem exists, and a priori knowledge must also guide model selection. For instance, it may be known that estimated concentrations need to be non-negative, or strictly negative in case of bleaching. Summarizing: when no a priori information about the correct model is present, two steps must be distinguished in the analysis of time-resolved spectra: fitting the data and finding the "best" model. A common procedure is the following: exponential decays are assumed to be present and the data are fitted with a sufficient number of decays n_{decay}, so that the residuals appear satisfactory. This is usually termed "global analysis" (§2.4.1). Judging goodness of fit is strongly problem dependent. Suppose the Decay Associated

Spectra are all different (which is often the case), and the SVD analysis is in accordance with the presence of n_{decay} spectrally and temporally independent components. Note that in general the DAS do not correspond to real spectra; this is only the case when a component decays without interconversions to other (spectrally active) components. Then, to find the best compartmental model (§2.4.3), several models with $n_{comp} = n_{decay}$ components can be tried, comparing them by the plausibility of what is then termed their Species Associated Spectra (SAS). Here a spectral model can be of great help. This is usually termed "target analysis". Furthermore thermodynamic considerations can be helpful in case the dependence on temperature or pH has been measured (§2.4.3.6).

2.6.2.1 Revisiting the unbranched, unidirectional model

When a sequential scheme with increasing lifetimes (§2.4.3.2) is used, and it is not yet known whether this is the correct kinetic scheme, the estimated spectra are called Evolution Associated Spectra (EAS) or Evolution Associated Difference Spectra (EADS).   In the following we will not differentiate between spectra (S) and difference spectra (DS). The DAS are interpreted as loss or gain of emission or absorption with a certain lifetime, whereas the EAS represent the spectral evolution, e.g. the third EAS rises with the second lifetime and decays with the third lifetime. With simple systems the interpretation of the DAS or EAS can be straightforward. E.g. when the sequential scheme with increasing lifetimes represents the correct physicochemical picture, the EAS correspond to true SAS characterizing the intermediate states. Combining Eq. 20 and Eq. 24 we find

$$C_{II} = C_I B \tag{41}$$

where B is an upper triangular matrix containing the elements defined in Eq. 25. In turn, this can be combined with Eq. 40, to derive the relation between DAS and EAS:

$$C_{II} \cdot EAS^T = C_I B \cdot EAS^T = C_I \cdot DAS^T \tag{42}$$

therefore

$$DAS = EAS \cdot B^T \tag{43}$$

Note that now B^T is a lower triangular matrix, expressing that the i-th DAS is a linear combination of the i-th and following EAS. An example of these relations was already discussed in §2.4.3.1. However, apart from the simple two-compartment scheme presented

there, the inverse relations are less obvious. E.g. the first DAS is a linear combination of all

EAS, because the first lifetime affects all subsequent components. The other way around:

$$EAS = DAS \cdot B^{-T} \tag{44}$$

Again, B^{-T} is a lower triangular matrix, expressing that the i-th EAS is a linear combination

of the i-th and following DAS. Since a triangular matrix can easily be inverted, the coefficients

b_{jl}^{-1} and b_{lj}^{-T} of B^{-1} and B^{-T} are found to be $b_{1l}^{-1} = 1 = b_{l1}^{-T}$ and for $j \leq l$:

$$b_{jl}^{-1} = \prod_{n=1}^{j-1} \frac{(k_n - k_l)}{k_n} = b_{lj}^{-T} \tag{45}$$

In words, the first EAS, which is the time zero spectrum, is the sum of all DAS. Furthermore,

the last EAS and the last DAS are proportional because of the sequential scheme with

increasing lifetimes. Thus when the DAS and lifetimes have been estimated, the EAS

resulting from any desirable sequential kinetic scheme can easily be calculated using Eq. 44

and Eq. 45.

2.6.3 Projecting the data upon singular vectors

In [86,88] applications of SVD in time-resolved spectroscopy are discussed. SVD is also

instrumental in self-modelling of the BR photocycle [47,48,49]. Next to the use of SVD for

rank estimation (§2.6.1), an important application is data reduction and noise suppression.

Assuming the noise is small, the Singular Value Decomposition results can be used to project

the data upon the first n_{comp} singular vectors. Projecting upon the first n_{comp} right singular

vectors the kinetic model Eq. 38 becomes

$$\underline{\Psi} W_{n_{comp}} = C(\theta) E^T W_{n_{comp}} + \underline{\Xi} W_{n_{comp}} \tag{46}$$

Analogously, projecting upon the first n_{comp} left singular vectors the spectral model Eq. 39

becomes

$$U_{n_{comp}}^T \underline{\Psi} = U_{n_{comp}}^T C E^T(\theta) + U_{n_{comp}}^T \underline{\Xi} \tag{47}$$

These projections reduce the dimensionality of the nonlinear least-squares fits [34,51,93,94]

thus saving computational resources. A drawback of the projection is that the noise can easily

perturb the projected data (§2.6.1), resulting in loss of information. This can easily be

demonstrated with simulated data, e.g. [89]. Furthermore, it becomes much more difficult to

calculate summary statistics (§3). These drawbacks can easily be avoided by using the variable projection algorithm [95], discussed more fully in §3, by which the nonlinear least-squares fit of the full, unprojected data becomes feasible.

2.6.4 Case study: conformational change studied via emission

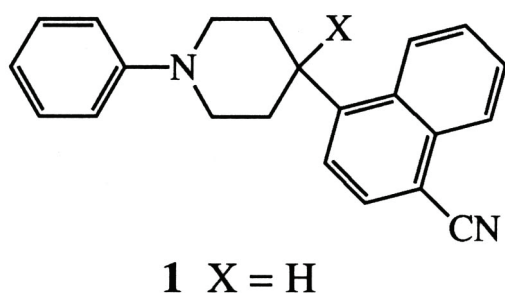After excitation at 308 nm of a bridged electron donor-acceptor (DA) system (Fig. 8), a

1 X = H

Fig.8. Bridged electron donor-acceptor (DA) system consisting of a photoexcitable cyanonaphthalene electron acceptor, an aromatic amine electron donor, and a semiflexible ring system linking the two.

charge-transfer (CT) state can be formed. Depending upon the solvent polarity, this CT state can exhibit conformational dynamics, since the charges attract each other in a nonpolar solvent, and if there is enough rotational freedom the bridged system will fold. The energetics and kinetic scheme are depicted in Fig. 9, whereas molecular models of the extented and compact CT state are shown in Fig. 10. This process is termed harpooning, since the excited
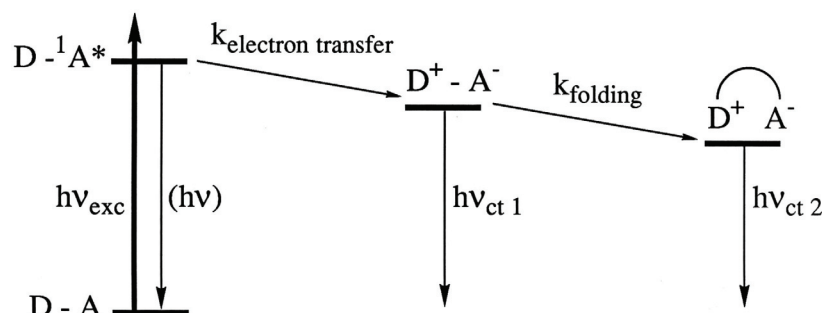
Fig.9. Proposed reaction scheme for harpooning. Left the excitation of the acceptor with possible fluorescence of a locally excited $D - {}^{1}A^{*}$ state. $h\nu_{ct\,1}$ and $h\nu_{ct\,2}$ correspond to fluorescence of the extended and folded conformation respectively

acceptor will first accept an electron from the donor, thereby creating the CT state. In turn, the opposite charges attract each other, as if the donor first throws an electron harpoon, and then attracts the acceptor. This will be accompanied by a red shift of the emission spectrum. An excellent methodology for the study of these conformational dynamics is time-resolved fluorescence spectroscopy in combination with global and target analysis, establishing the

Fig.10. Molecular models of extended and folded CT-states of compound **1** depicted in Fig. 8 derived using molecular mechanics with AM1/UHF atomic charges.

time scales, the number of different conformers, their dipole moments, energetics, etc. [34, 69, 70].

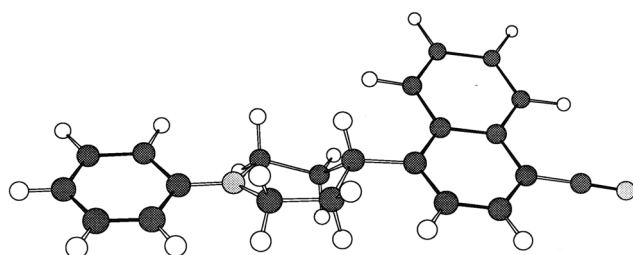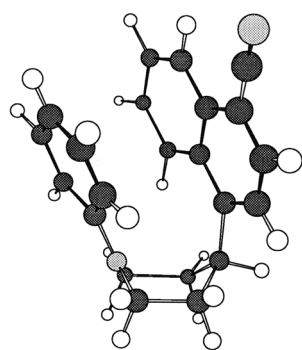Emission was measured on a nanosecond time scale after excitation at 308 nm of the acceptor of compound **1** in the apolar solvent cyclohexane. The IRF width is typically 11 ns, precluding the observation of ultrafast processes like solvation or the actual charge transfer. Therefore emission from an extended ECT state will appear "instantaneously" with this time resolution. In the following we will analyse two different types of emission measurements: traces with acceptable time resolution measured up to 40 ns at 11 wavelengths (Fig. 11) and time-gated spectra with very good wavelength resolution measured up to 80 ns at intervals of 5 ns. (Fig. 12 and following). The IRF of the trace measurements depicted as a black line straddling time zero in Fig. 11C, displays a small tail. Global analysis revealed that these data could be satisfactorily described by three lifetimes: 0.76 ns (blue), 7.8 ns (black) and 10.8 ns (red). Panels A and B depict the excellent fit of two representative traces, and the residuals shown in the insets show little structure. The contributions of the three decays to the fits are also shown as dashed lines with the respective colors. The decay at 420 nm (Fig. 11B) shows clear evidence of a rise of this emission with 7.8 ns. The DAS (Fig. 11D) show that this rise is present for all wavelengths above 380 nm. Note however that there are no negative amplitudes for the fastest lifetime (0.76 ns, blue). Singular Value Decomposition of the matrix of residuals revealed no outstanding singular values (Fig. 11G). The first left and right
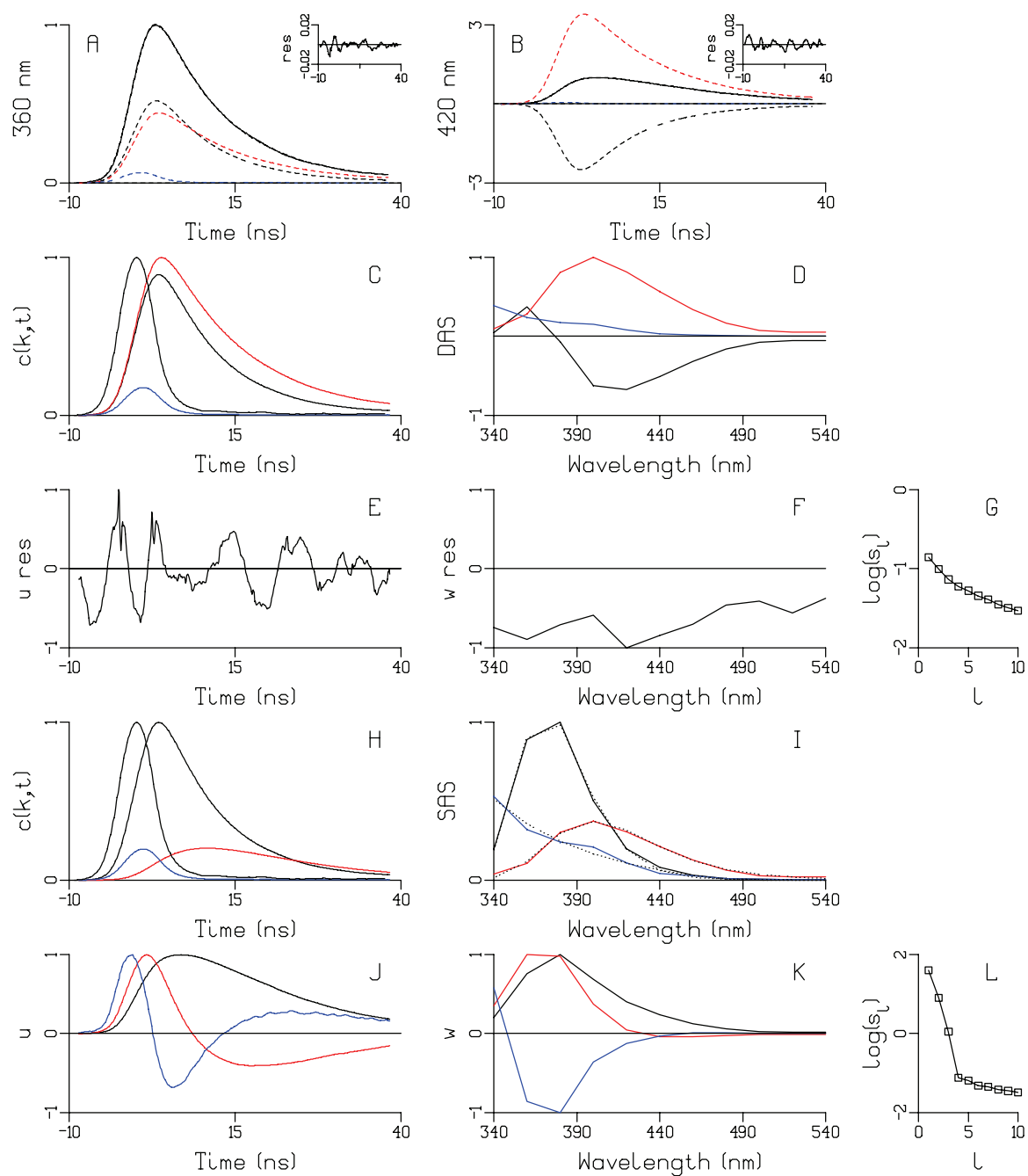
Fig.11. Global (C,D) and target (H,I) analysis of traces measured up to 40 ns. (A,B) fit using three lifetimes: 0.76 ns (blue), 7.8 ns (black) and 10.8 ns (red). Insets show residuals. (C) Exponential decays convolved with measured IRF (black line straddling time zero, with a small tail). (D) DAS. SVD of the residual matrix: (E,F): first left and right singular vector, (G,L): Singular values on logarithmic scale. (H) In contrast to C, the red curve now results from a rise with 7.8 ns, and decay with 10.8 ns. (I) SAS. SVD of the data matrix: (J,K): first three (order black, red, blue) left and right singular vectors.

singular vectors (Fig. 11E,F), which represent the main contribution to the residuals, show no obvious structure in the wavelength dimension, but oscillations along the time dimension. These might arise from errors in the measured IRF, from time jitter of the laser, or from fluctuations in the laser intensity.

SVD of the matrix of data revealed three outstanding singular values (Fig. 11L), indicating that three spectrally and temporally different components are present in these data. The first three left and right singular vectors (Fig. 11J,K) are practically noise-free. Note that the singular values Fig. 11G are somewhat larger than the fourth and following singular values of Fig. 11L, which is due to the small systematic errors indicated in Fig. 11E.

Analogous to the analysis of the simulated two component model in Fig. 4, the data could also be fit using a target model whereby the longest lived emission (10.8 ns, red) grows in with 7.8 ns. The fast (0.76 ns) emission is assumed to represent a parallel decay. Schematically:

$$1 \rightarrow 2 \rightarrow \qquad 3 \rightarrow$$
$$7.8\text{ns} \quad 10.8\text{ns} \qquad 0.76\text{ns}$$

$$(48)$$

This results in the concentration profiles and SAS shown in Fig. 11H,I. Recall that this target analysis results in exactly the same residuals. Note the analogy between the black and red components in Fig. 11H,I and the first and second component in Fig. 4D,F, and also between the black and red components in Fig. 11A,B and the first and second component in Fig. 4C,E. Subsequently, the satisfactory SAS in Fig. 11I can be well fit using skewed Gaussians (Eq. 31), as indicated by the dotted black lines. We conclude that the black and red components in Fig. 11H,I represent the extended and folded CT-states (ECT and CCT) of compound **1**. The blue component is attributed to local emission of the $D - {}^{1}A^{*}$ state (cf. Fig. 9).

When fitting the time-gated spectra, the first lifetime was fixed at 0.77 ns, because of the 5 ns time interval. Using three lifetimes, the results are shown in Fig. 12. Note that here the residuals are unsatisfactory (Fig. 12C). SVD of the matrix of residuals revealed four outstanding singular values (Fig. 12G). The first left and right singular vectors (Fig. 12E,F), which represent the main contribution to the residuals, show large oscillations along the time dimension, which are largest around the maximum of the data. These might arise from time jitter of the laser, or from fluctuations in the laser intensity. Note that the red and blue DAS in Fig. 12D are similar to the black and red DAS in Fig. 11D. Furthermore, the long lifetimes of 8.5 ns (red) and 11.3 ns (blue) are close to the 7.8 ns and 10.8 ns from the trace measurements. The black EAS in Fig. 12B of the 0.77 ns component is similar to the red ECT SAS of Fig.

11I. The red EAS in Fig. 12B resembles more a sum of the ECT and CCT SAS, cf. the black and red SAS in Fig. 11I. Thus the agreement between the two analyses can be considered half-way satisfactory. Clearly, the quality of the fit of the time-gated spectra leaves much room for improvement. When using a spectral model, with skewed Gaussians, the amplitude fluctuations of the laser intensity can be absorbed in the amplitudes of the component concentrations. Therefore, the rms error of the fit decreases from 480 to 292. This is clearly visible when comparing Fig. 12C and Fig. 13D. The estimated spectra from Fig. 13B are very similar to the SAS from Fig. 11I. SVD of the matrix of residuals revealed three outstanding singular values (Fig. 13G). The first right singular vector (Fig. 13F), shows systematic shortcomings of the spectral shape model. In particular, the fine structure at about 404 and 436
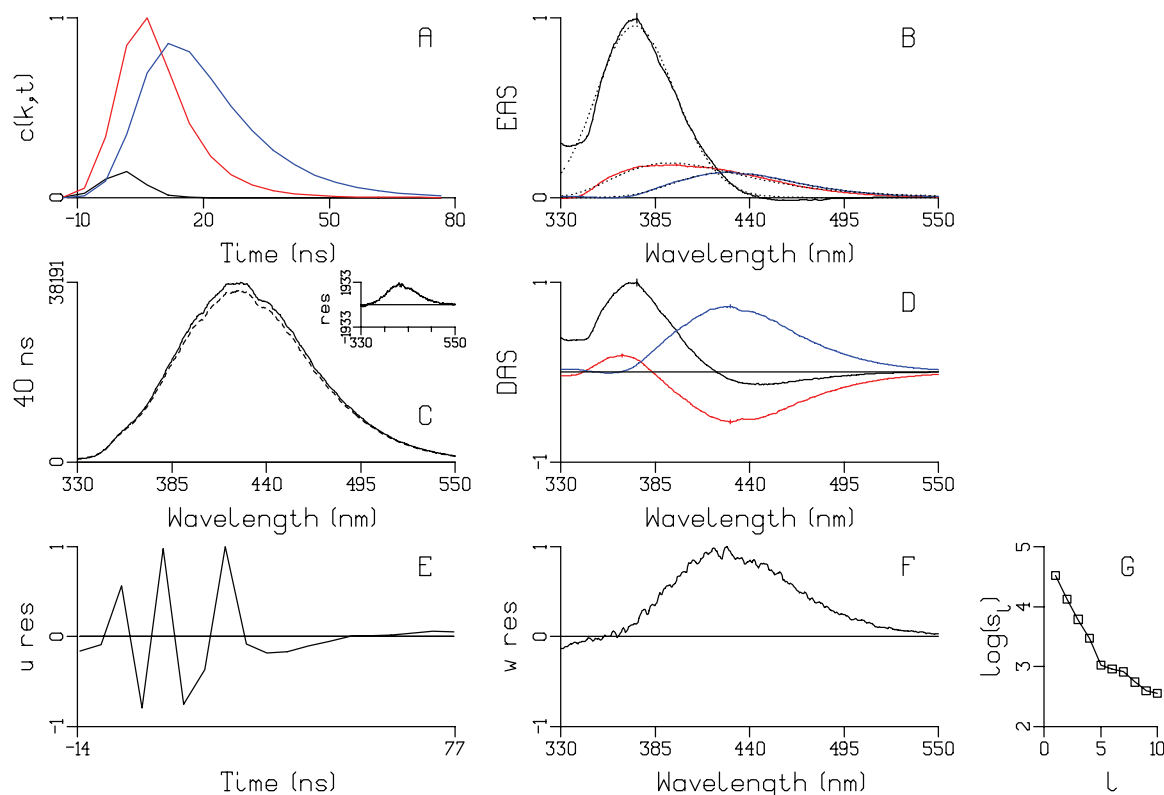
Fig.12. Global (D) and evolutionary (A,B) analysis of spectra measured up to 80 ns. (A) evolutionary model, with increasing lifetimes. FWHM of Gaussian IRF 11.2 ns. Lifetimes: 0.77 ns (black, fixed), 8.5 ns (red) 11.3 ns (blue). (B) EAS, dotted black lines. indicate fit of EAS using skewed Gaussians (Eq. 31). (C) Lack of fit with 40 ns spectrum, inset shows residuals. (D) DAS. SVD of the residual matrix: (E,F): first left and right singular vector, (G): Singular values on logarithmic scale. Rms error 480.

nm points at vibrational fine structure of the emission spectrum. This corresponds to an energy of $10^7/404 - 10^7/436 \approx 1800 \text{cm}^{-1}$. This type of systematic error is proportional to the signal, and can to some extent be corrected by estimation of a multiplicative bias β. Thus for each wavelength λ_j we estimate β_{λ_j} from a linear regression between the fit and the residuals z_{λ_j} (the circumflex ^ indicates the estimate):

$$z_{\lambda_j}(t) = \psi_{\lambda_j}(t) - \hat{\psi}_{\lambda_j}(t) = \hat{\psi}_{\lambda_j}(t)(\beta_{\lambda_j} - 1) \tag{49}$$

where the fit is given by

Fig.13. Global analysis (A,B) of time-gated spectra using a spectral model. (A) amplitudes of skewed Gaussians from (B). Red ECT, blue CCT, black local emission. Since the product $c(t)\varepsilon(\theta, \lambda)$ is estimated, the extremum of this product is calculated per component. Since red has the largest extremum, it is 1 in both A and B. Next the blue extremum is 0.88^2 times the red extremum, thus it is 0.88 in both A and B. Likewise, the black extremum is 0.38^2 times the red extremum. (C,D) Lack of fit with 15 and 40 ns spectrum, insets show residuals. SVD of the residual matrix: (E,F): first left and right singular vector, (G): Singular values on logarithmic scale. Rms error 292.

$$\hat{\psi}_{\lambda_j}(t) = \sum_{l=1}^{n_{\text{comp}}} \hat{c}_l(t)\varepsilon_l(\hat{\theta}, \lambda_j) \tag{50}$$

and where the $\hat{c}_l(t)$ and $\hat{\theta}$ are estimated from the global analysis. When the estimate β_{λ_j} resulting from Eq. 49 differs significantly from 1, it can be concluded that it represents multiplicative bias. Now the spectral model can be augmented with these estimates:

$$\hat{\psi}_{\lambda_j}(t) = \sum_{l=1}^{n_{\text{comp}}} \hat{c}_l(t)\varepsilon_l(\hat{\theta}, \lambda_j)\hat{\beta}_{\lambda_j} \tag{51}$$

and in a refined global analysis the spectral model function $\varepsilon_l(\theta, \lambda_j)\hat{\beta}_{\lambda_j}$ can be used. Note that this bias can correct for sharp features in the spectral response of the detection system, but not for a smooth change in the sensitivity, which will be absorbed in the spectral parameters. When in a certain wavelength range the response is predominantly due to a single component, e.g. the local emission, this procedure can correct for systematic deviations (like vibrational
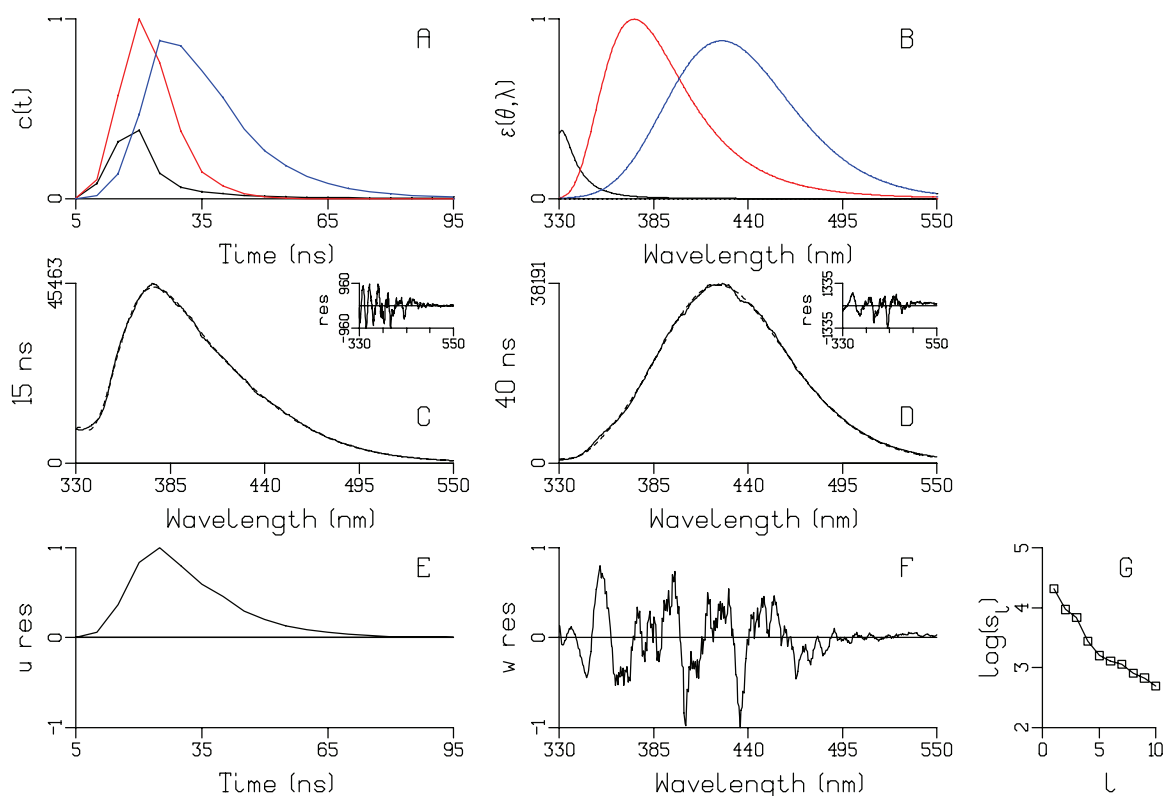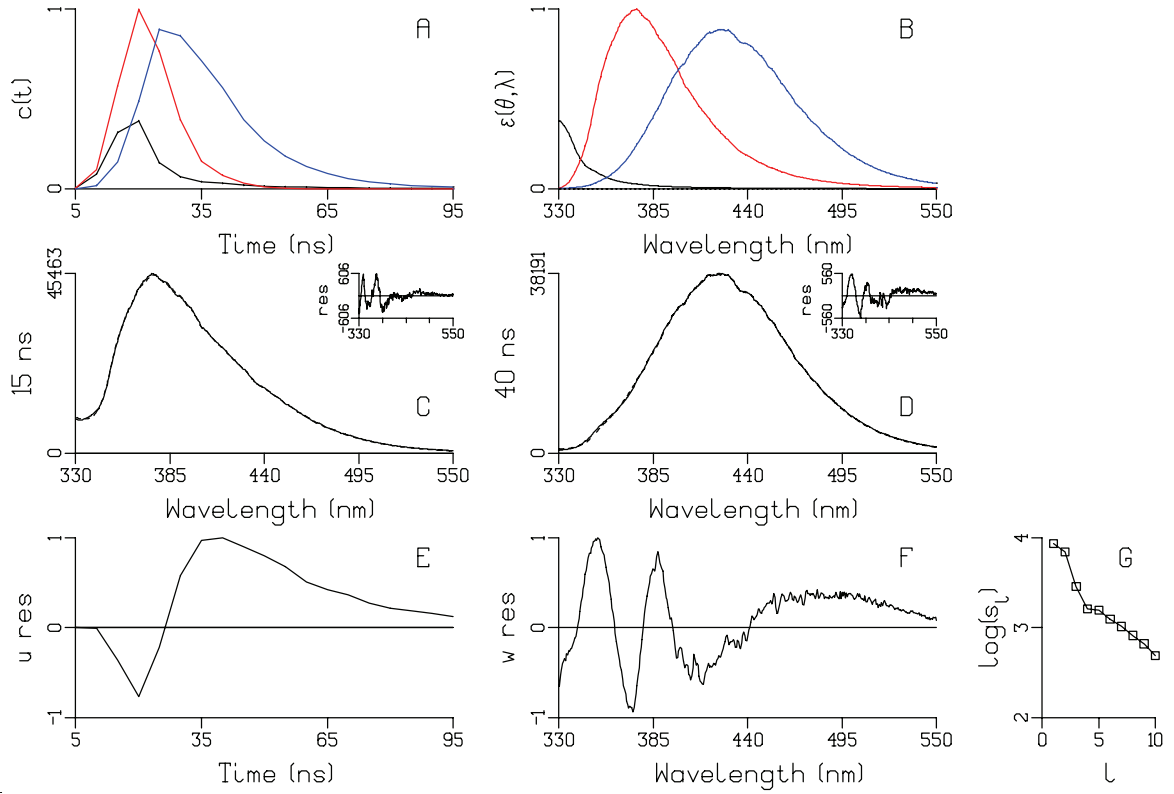

Fig.14. Global analysis (A,B) of spectra using a spectral model with bias correction. (A) amplitudes of skewed Gaussians from (B). Red ECT, blue CCT, black local emission. (C,D) Lack of fit with 15 and 40 ns spectrum, insets show residuals. SVD of the residual matrix: (E,F): first left and right singular vector, (G): Singular values on logarithmic scale. Rms error 144.

fine structure, or some other shortcoming of the model function) from the assumed skewed Gaussian shape.

Fig. 14 shows the results of this refined analysis. Note that the bias results in a reduction of the rms error from 292 to 144. However, again SVD of the matrix of residuals revealed at least two outstanding singular values (Fig. 14G). The first left and right singular vectors (Fig. 14E,F), now indicate that a long lived red shifted emission is present, as well as some
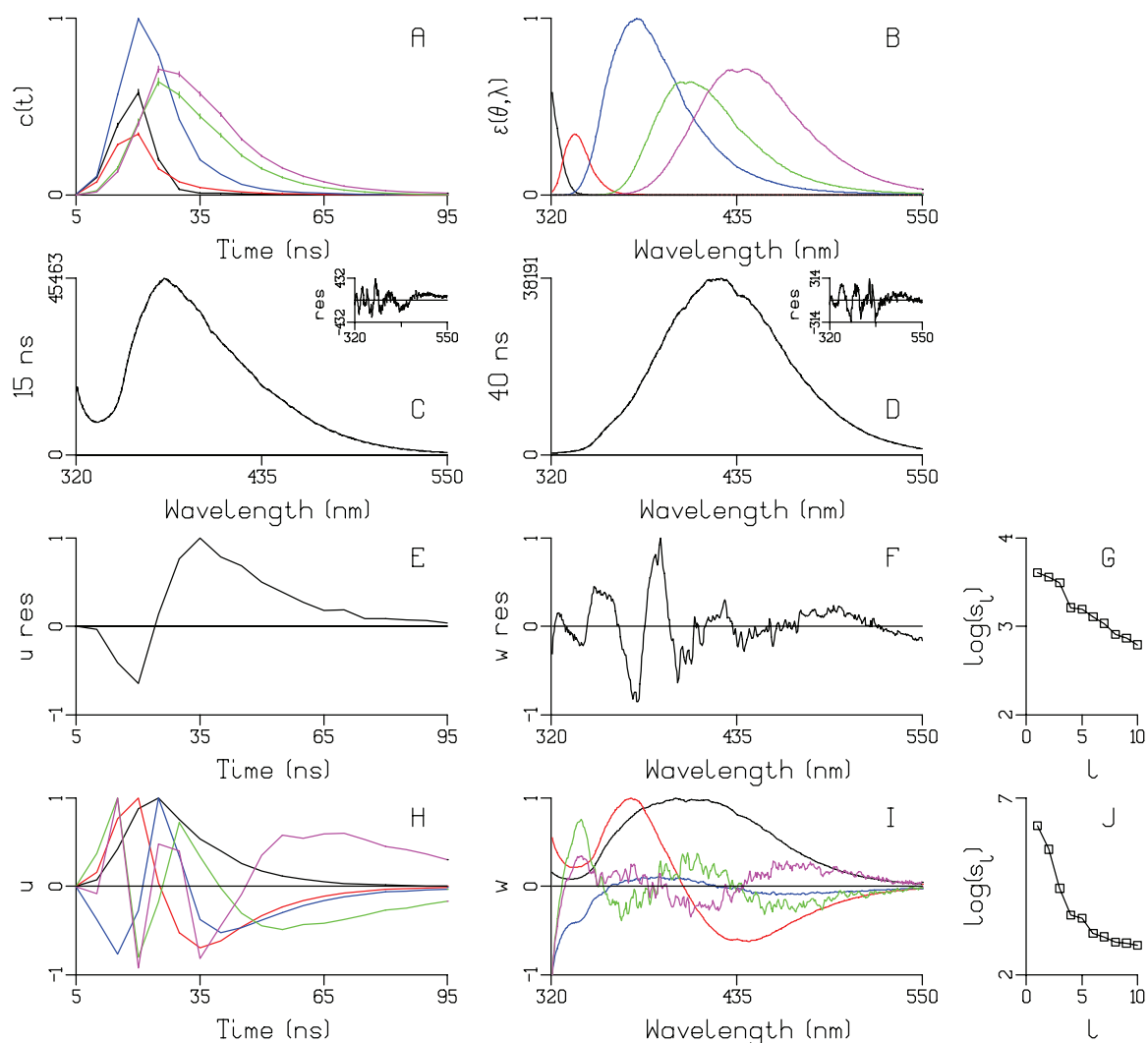
Fig.15. Global analysis (A,B) of spectra using a five component spectral model with bias correction. (A) amplitudes of skewed Gaussians from (B). Black scatter, red local emission, blue ECT, green CCT1, magenta CCT2. (C,D) Lack of fit with 15 and 40 ns spectrum, insets show residuals. SVD of the residual matrix: (E,F): first left and right singular vector, (G,J): Singular values on logarithmic scale. SVD of the data matrix: (H,I): first five (order black, red, blue, green, magenta) left and right singular vectors. Rms error 84.

component at the blue edge. Thus a five component fit was attempted. This results in a further reduction of the rms error from 144 to 84. Still some structure remains in the singular vectors of the residual matrix Fig. 15E,F, albeit much less than in Fig. 14E,F. Especially the blue edge is now fitted very well by a tail attributable to scatter from the exciting laser light (black) and local emission (red). Note that the SVD of the data Fig. 15H,I,J matrix also indicates five components. The late spectra are now dominated by a more red shifted (thus more relaxed, resulting in lower emission energy) CCT state. The estimated amplitude profiles of ECT, CCT1 and CCT2 from Fig. 15A were then subjected to a kinetic analysis, thereby down-weighting the early time points which are very susceptible to the systematic error induced by the laser fluctuations. The target model used contained three compartments, with the ECT (red in Fig. 16A, component 3 in Fig. 16B,D) decaying in 5.9 ns, thereby giving rise to both CCT1 (component 4 in Fig. 16B,D) and CCT2 (component 5 in Fig. 16B,D). This rise is clearly visible in the Decay Amplitudes (DA) of Fig. 16D. The CCT lifetimes are 9.4 ns (blue) and 19 ns (green). Note that CCT1 decays dominantly with 9.4 ns (component 4 in Fig. 16B) although a small decay with 5.9 ns also contributes. About one third of the CCT2 decay is with 19 ns, as deduced from the relative green and blue amplitudes of component 5 in Fig. 16B. From Fig. 15A,B it is estimated that about 45% of ECT decays to CCT1, and 55% to CCT2. Combining this with the amplitudes from Fig. 16B we estimate that about 7% of the ECT states decays through CCT2 with 19 ns, 2% decays through CCT1 with 5.9 ns, and the large majority, 91% decays with 9.4 ns through CCT1 or CCT2. Schematically:

$$
\begin{array}{cccc}
43\% & 2\% & 48\% & 7\% \\
3 \to 4 \to & 3 \to 4' \to & 3 \to 5 \to & 3 \to 5' \to \\
5.9\text{ns} \ \ 9.4\text{ns} & 5.9\text{ns} \ 5.9\text{ns} & 5.9\text{ns} \ \ 9.4\text{ns} & 5.9\text{ns} \ \ 19\text{ns}
\end{array}
\tag{52}
$$
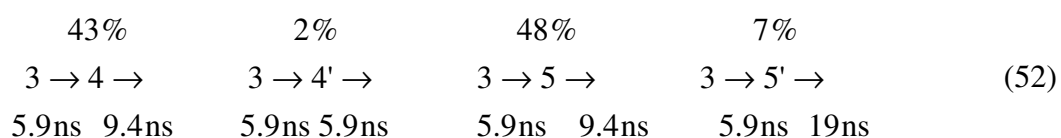
Thus from this elaborate refined analysis of these time-gated spectra it is concluded that the ECT folds (in ~6 ns) to CCT, and that the simple picture of Fig. 10 is quite a good approximation. However, there are strong indications that the CCT emission derives from different conformers with different lifetimes, ranging from 6 to 19 ns. Thus some heterogeneity of the CCT conformers appears to be present.

Concerning the data analysis, we conclude that over a limited time range of 40 ns the simple $ECT \to CCT$ model describes the trace measurements well (Fig. 11), i.e. up to the level of the systematic errors in the data. Apart from the local emission, this is analogous to the simulation of Fig. 4. The time-gated spectrum suffered heavily from time jitter and laser intensity
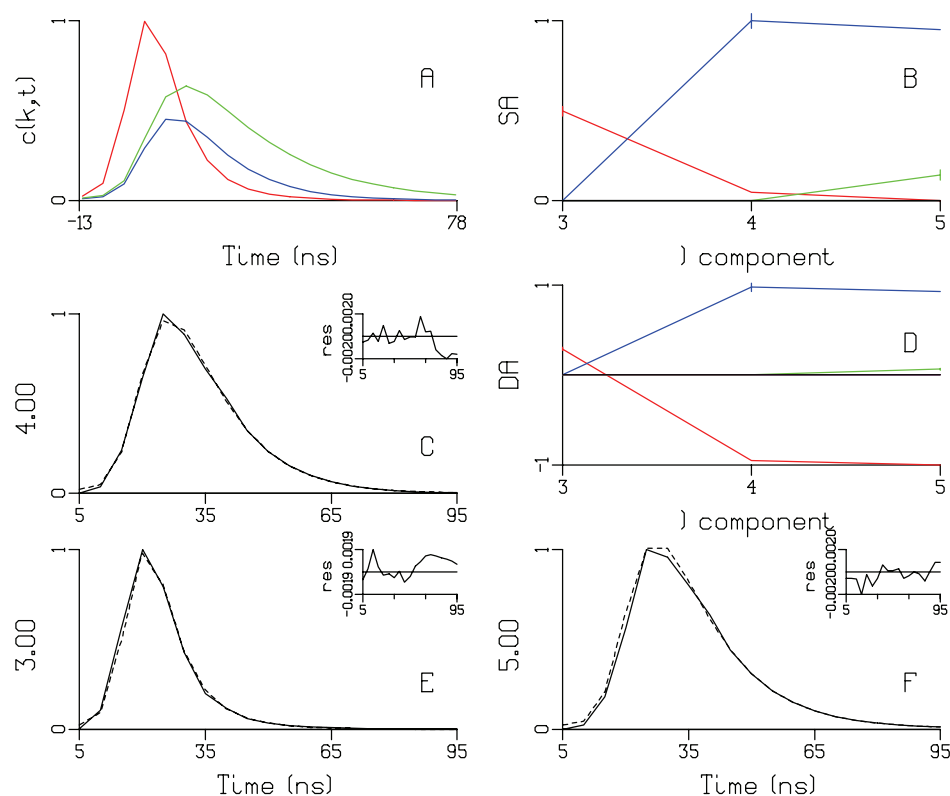
Fig.16. Global (D) and target (A,B) analysis of amplitude profiles of ECT (component 3, panel E), CCT1 (component 4, panel C) and CCT2 (component 5, panel F) from Fig. 15A. Insets in panels C,E,F depict weighted residuals, disemphasizing early times before 40 ns. Key: red 5.9 ns, blue 9.4 ns, green 19 ns. (B) Species Amplitudes (SA), (D) Decay Amplitudes (DA)

fluctuations. In order to deal with these systematic errors a spectral model is most suitable, in particular when augmented with a multiplicative bias. The rms error decreased from 480 with a three component kinetic model, via 292 with a three component spectral model, 144 with a three component spectral model and bias augmentation, to 84 with a five component spectral model with bias augmentation. Thus, even five components could be resolved (Fig. 15) and a subsequent kinetic analysis of the ECT and CCT component traces indicated heterogeneity of the CCT conformers. In this iterative modelling process, depicted in the model adjustment loop of Fig. 2, careful inspection of the residuals, as summarized by the SVD of the residuals was crucial.

2.7 Spectrotemporal models

In the ideal case all a priori knowledge is used for a model-based fit of the data. This leads to a spectrotemporal model [69]. Three cases can be distinguished: (1) a parameterized model for both kinetics and spectral shapes is available, (2) in addition to the kinetic model, limited spectral knowledge is available, in the form of spectral equalities or constraints (e.g. the spectrum is assumed to be zero in a certain wavelength range) and (3) in addition to the model for the spectral shapes, limited band amplitude curve knowledge is available, e.g. a certain amplitude is assumed to be zero in a certain time range.

When the separability of time and wavelength properties, Eq. 1, is not applicable, sometimes a spectrotemporal model can be used [96,97]. E.g., in describing solvation the time dependence of the spectral parameter $\bar{\nu}_{max}$ from Eq. 31 or Eq. 32 can be described as $\bar{\nu}_{max}(t) = \bar{\nu}_{max}(\infty) + (\bar{\nu}_{max}(0) - \bar{\nu}_{max}(\infty))\exp(-t/\tau)$ where τ represents the characteristic solvation time.

2.7.1 Spectral shape model and kinetic model

The most straightforward spectrotemporal model includes a kinetic model (e.g. a compartmental model from §2.4.3), a spectral shape model (e.g. from §2.5.1) and in general also some amplitude parameters. In case both the kinetic and the spectral model are of a global type, these are the amplitudes of each combination of exponential decay and bandshape. With a more detailed model they can be limited to one or more scaling parameters. It is shown in [98] that with such a detailed model, the parameters of a multicomponent model can be estimated with higher precision. In the case of zero spectral overlap this precision is equal to the precision of a single component model.

In Eq. 53 both the concentrations and the spectra are described by a model, which depends upon the parameters θ. Assuming first-order kinetics, a matrix of amplitude parameters A describes the concentrations of the components in terms of a superposition of simple decays which are collated in the matrix $C(\theta)$.

$$\underline{\Psi} = C(\theta)AE^T(\theta) + \underline{\Xi} \tag{53}$$

When applied to the simulated data of §2.4.3.1, the matrix $C(\theta)$ contains two exponential decays (parameters k_l) , whereas the matrix $E(\theta)$ consists of two skewed Gaussian shapes (Eq. 31, parameters $(\bar{\nu}_{max}, \Delta\bar{\nu}, b)_l$). The matrix of amplitude parameters A to be estimated

will be proportional to the A_{II} from Eq. 22.

The vector representation [25] of the matrix Ψ is given by

$$\text{vec}(\underline{\Psi}) = (E(\theta) \otimes C(\theta))\text{vec}(A) + \text{vec}(\underline{\Xi}) \tag{54}$$

where $\otimes$ denotes the Kronecker product [25]. In case a particular kinetic model (with concentrations $c_i(\theta)$) is tested we put

$$\text{vec}(\underline{\Psi}) = \sum_{i=1}^{n_{\text{comp}}} \text{vec}(c_i(\theta)\varepsilon_i^T(\theta))a_i + \text{vec}(\underline{\Xi}) \tag{55}$$

When applied to the simulated data of §2.4.3.1, this is equivalent to the target analysis with the $c_i(\theta)$ from Model II. Instead of the n_{comp}^2 parameters of the A matrix now only n_{comp} amplitude parameters need to be estimated. It was demonstrated in [98] that this improves the precision of the parameters θ in Eq. 55 relative to Eq. 54.

2.7.2 A priori spectral knowledge and kinetic model

The second type of spectrotemporal model includes a kinetic model (e.g. a compartmental model from §2.4.3), and a spectral model which incorporates limited a priori spectral knowledge, e.g. a spectrum is assumed to be zero in a certain wavelength range. Thus in that wavelength range a component with zero spectrum does not contribute. An extreme case is the emission from P* (from the bacterial reaction centre) which is assumed to be in equilibrium with non-emitting relaxed radical pair states (see §2.4.3.3) [71].

With difference absorption measurements the situation becomes even more complicated, because the ground state bleach spectrum is needed as well, and in the case of a photocycle an extra parameter for the fraction cycling has to be introduced (§2.7.4).

The kinetic model of Eq. 38 needs to be modified in order to incorporate the a priori spectral knowledge

$$\underline{\Psi} = C(\theta)\tilde{E}^T + \underline{\Xi} \tag{56}$$

where $\tilde{E}^T$ contains less unknown parameters than $n \times n_{\text{comp}}$. E.g. in a certain wavelength range for certain components some of the elements of $\tilde{E}$ are equal to zero, or some linear relationships exist. Examples will be discussed in §2.7.4 and in the case study (§4).

2.7.3 A priori band amplitude knowledge and spectral model

The third type of spectro"temporal" (between quotes because the other dimension can be time, pH, temperature, excitation wavelength, etc.) model includes a spectral shape model (e.g. from §2.5.1) and a BAC model which incorporates a priori knowledge, e.g. that in a certain time (or pH, or temperature, ...) range the band amplitude is assumed to be zero. Thus in that range a component with zero amplitude does not contribute.

The spectral model of Eq. 39 needs to be modified in order to incorporate the a priori band amplitude knowledge

$$\underline{\Psi} = \tilde{C}E^T(\theta) + \underline{\Xi} \tag{57}$$

where $\tilde{C}$ contains less unknowns than $m \times n_{comp}$. E.g. in a certain range for certain components some of the elements of $\tilde{C}$ are equal to zero, or some linear relationships exist. This type of model was applied in studying the oligomerization of photosynthetic antenna peptides [99]. From a series of detergent-dependent absorption spectra an intermediate spectrum could be resolved with the help of the constraints that this spectrum did not contribute at the extreme detergent concentrations.

2.7.4 Spectrotemporal model for a photocycle

With difference absorption measurements from a photocycle the situation becomes even more complicated, because the ground state bleach spectrum is needed as well, and an extra parameter for the fraction cycling has to be introduced. The noise-free, time-resolved difference absorption ΔA is a superposition of the contributions of the n_{comp} different components (analogous to Eq. 1):

$$\Delta A(t, \lambda) = \sum_{l=1}^{n_{comp}} c_l(t)\Delta\varepsilon_l(\lambda) \tag{58}$$

where $c_l(t)$ and $\Delta\varepsilon_l(\lambda)$ denote, respectively, the concentration and SADS of component l. By definition $\Delta\varepsilon_l(\lambda) = \varepsilon_l(\lambda) - \varepsilon_0(\lambda)$, where $\varepsilon_0(\lambda)$ is the ground state bleach spectrum. Regarding Eq. 58 we note that the quantity which will be estimated is the product $c_l \cdot \Delta\varepsilon_l$ which in itself is insufficient for the determination of the absolute values of c_l and $\Delta\varepsilon_l$. Since in the photocycle no states are lost the relative concentrations of the components can be estimated, and thus also the relative amplitudes of their difference spectra. Here we take

$c_1(0) \equiv 1$, and thus all concentrations are relative to the concentration of the first photocycle state in the model.

In matrix notation Eq. 58 reads:

$$\Delta A = C(E - \varepsilon_0)^T = CE^T - C1\varepsilon_0^T \tag{59}$$

where the $m \times n$ matrix ΔA denotes the time-resolved difference absorption, measured at m time instants t_i, and n wavelengths λ_j. The columns of the matrices C and E, of dimension $m \times n_{\text{comp}}$ and $n \times n_{\text{comp}}$, respectively, contain the concentration profiles and SAS of the components. The matrix-vector product $C1$ is a vector containing the sum of the concentrations of the photocycling intermediates, which is equal to the ground state depletion. This sum decreases monotonically from one at time zero to zero at the end of the photocycle. When the ground state spectrum of the sample before excitation $\tilde{\varepsilon}_0$ has been measured on exactly the same setup (which is not always feasible in the case of ultrafast measurements), the model can be extended with the fraction cycling parameter fc. We can then substitute for the bleach spectrum $\varepsilon_0 = fc \cdot \tilde{\varepsilon}_0$. Using the vector representation of a matrix and the Kronecker product [25] ($\otimes$), Eq. 59 can then be rewritten:

$$\text{vec}(\Delta A) = (I_n \otimes C)\text{vec}(E^T) - (\tilde{\varepsilon}_0 \otimes C1)fc \tag{60}$$

When we use a kinetic model $C(\theta)$ we can express Eq. 60 as

$$\text{vec}(\Delta A) = \left[I_n \otimes C(\theta) \; -\tilde{\varepsilon}_0 \otimes C(\theta)1 \right] \begin{bmatrix} \text{vec}(E^T) \\ fc \end{bmatrix} \tag{61}$$

Since the last column $-\tilde{\varepsilon}_0 \otimes C(\theta)1$ is a linear combination of all the other columns, an extra assumption is necessary to remove this dependence. With bacteriorhodopsin (BR) the M state(s) are assumed not to absorb above ~540 nm [24], thus removing these elements from $\text{vec}(E^T)$ (arriving at $\text{vec}(\tilde{E}^T)$) and deleting the accompanying columns in $I_n \otimes C(\theta)$. With photoactive yellow protein [6, 101,102] the pB state(s) are assumed not to absorb above ~430 nm. Both the M and pB states possess an absorption maximum to the blue of the ground state, and occur on a ms time scale during the photocycle.

2.7.5 Anisotropy models

Measurement and subsequent modelling of polarization dependence offer an opportunity to resolve components provided their anisotropies differ. Commonly measurements are made at

three angles relative to the polarization of the excitation: parallel, perpendicular and magic angle. At the magic angle there is no anisotropy effect present; from the two other measurements the anisotropies can be estimated. In an associative model, an anisotropy decay function $r_i(t)$ is associated with each component i. Alternatively, a single anisotropy $r(t)$ can be applied to all components or to all decays, which is called a nonassociative model. The models discussed below are all associative.

2.7.5.1 Emission anisotropy

In order to include parallel and perpendicular data the kinetic model for the magic angle data from Eq. 38 is extended by multiplying the concentration of each component i by $1 + \gamma r_i(t)$, where γ equals 0, 2, -1 for magic angle, parallel, and perpendicular data, respectively [13]. The full model for the experimental traces MA(t), VV(t), and VH(t) then reads:

$$\begin{bmatrix} MA(t, \lambda) \\ VV(t, \lambda) \\ VH(t, \lambda) \end{bmatrix} = \left(\sum_{l=1}^{n_{comp}} c_l(t)\varepsilon_l(\lambda) \begin{bmatrix} 1 \\ 1 + 2r_l(t) \\ 1 - r_l(t) \end{bmatrix} \right) \oplus i(t) \tag{62}$$

When Raman scattering is present, it can easily be included in the model. It has the time profile of the IRF and possesses an anisotropy r_{RS}, a spectrum $\varepsilon_{RS}(\lambda)$, and contributes a term $\varepsilon_{RS}(\lambda)(1 + \gamma r_{RS})i(t)$ to Eq. 62. Ideally r_{RS} equals 0.4.

Preferentially the measured data are modelled in Eq. 62. Alternatively, an anisotropy signal can be calculated from the parallel and perpendicular data

$$r(t, \lambda) = \frac{VV(t, \lambda) - VH(t, \lambda)}{VV(t, \lambda) + 2VH(t, \lambda)} \tag{63}$$

Compared to the measured data the calculated $r(t, \lambda)$ will be much more noise sensitive, in particular when the signals are small.

2.7.5.2 Difference absorption anisotropy

With difference absorption the situation is much more complicated. An excited state component possesses three spectral contributions: excited state absorption (ESA), ground state bleaching (GSB) and stimulated emission (SE), whereas a component in the ground state possesses two spectral contributions: absorption (GSA), and bleaching (GSB). In principle each contribution has its own anisotropy. Thus the total contribution of an excited state

component reads (omitting the subscript l for clarity):

$$(c(t)\{ESA(\lambda)(1 + \gamma r_{\mathrm{ESA}}(t)) + SE(\lambda)(1 + \gamma r_{\mathrm{SE}}(t)) + GSB(\lambda)(1 + \gamma r_{\mathrm{GSB}}(t))\}) \oplus i(t) \quad (64)$$

whereas a ground state component contributes:

$$(c(t)\{GSA(\lambda)(1 + \gamma r_{GSA}(t)) + GSB(\lambda)(1 + \gamma r_{\mathrm{GSB}}(t))\}) \oplus i(t) \qquad (65)$$

An important question is, under which conditions can these different contributions be resolved. Consider the following case: a ground state component possesses two contributions: a bleach with maximum anisotropy $r_{\mathrm{GSB}} = 0.4$, and absorption. Suppose the bleach is not present in part of the measured wavelength range, then from that part r_{GSA} can be estimated from

$$\begin{bmatrix} \Delta A(t, \lambda) \\ \Delta A_{//}(t, \lambda) \\ \Delta A_{\perp}(t, \lambda) \end{bmatrix} = \left(c(t) GSA(\lambda) \begin{bmatrix} 1 \\ 1 + 2 r_{\mathrm{GSA}}(t) \\ 1 - r_{\mathrm{GSA}}(t) \end{bmatrix} \right) \oplus i(t) \qquad (66)$$

When the thus estimated r_{GSA} differs from r_{GSB}, the two contributions can be resolved over the full wavelength range. In conclusion: contributions with different anisotropy can in principle be resolved when their anisotropy is known a priori, or when a contribution with unknown anisotropy appears isolated in part of the measurement range.

A coherent coupling artefact, which is often present in ultrafast experiments [33, 103], can be included in the model. Usually its time profile can be approximated by the IRF. It possesses an anisotropy r_{CA}, a spectrum $\varepsilon_{\mathrm{CA}}(\lambda)$, and contributes a term $\varepsilon_{\mathrm{CA}}(\lambda)(1 + \gamma r_{\mathrm{CA}})i(t)$ to Eq. 66. Target analysis of ultrafast difference absorption spectra was instrumental in resolving structural heterogeneity of the 2-aminopurine chromophore [100]. The different excited states possessed widely differing anisotropies, which were also wavelength-dependent. About 70% of the population showed a high anisotropy ($r \approx 0.35$) below 440 nm, and a low anisotropy ($r \approx 0.1$) above 440 nm. The remaining 30% of the population showed an opposite anisotropy ($r \approx -0.2$). The anisotropy information was limited to the first 50 ps, due to the rotational correlation time $\tau \approx 25\,\mathrm{ps}$ of the chromophore, $r_i(t) = r_{0i}\exp(-t/\tau)$.

2.7.5.3 Case study: anisotropy at one wavelength

Natalie Banerji (Geneva) measured the difference absorption anisotropy of the PMDA - PMB charge transfer complex at 610 nm after excitation at 400 nm. Fig. 17 summarizes the results
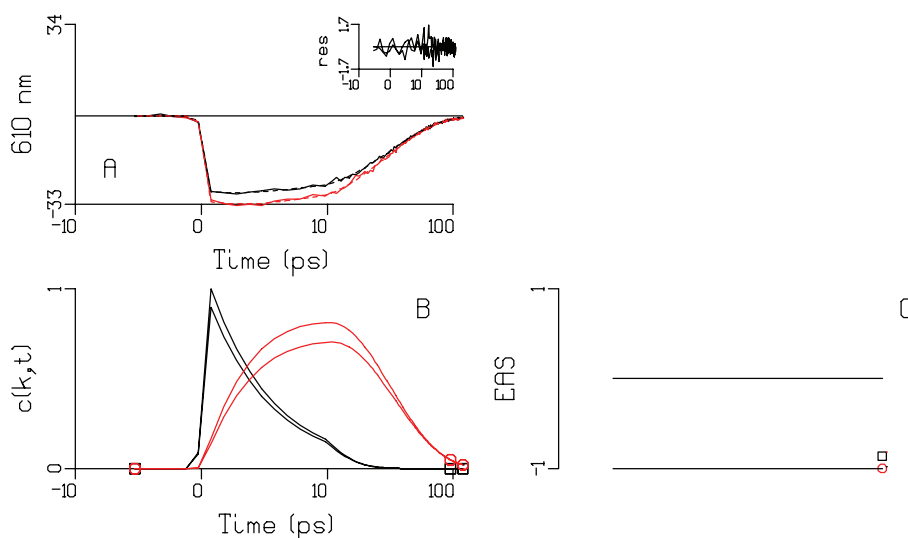
Fig.17. Target analysis of (A) difference absorption of PMDA - PMB at 610 nm detected parallelly (black) or perpendicularly (red). (B) evolutionary model, with increasing lifetimes. FWHM of Gaussian IRF 0.4 ns. Lifetimes: 5 ps (black), 28 ps (red). (B) "EAS": amplitudes of components at 610 nm.

from a target analysis. A sequential kinetic scheme with two lifetimes (5 and 28 ps, Fig. 17B) was used, and the estimated anisotropy of the two components was -0.036 and -0.061 (the outer curves in Fig. 17A,B correspond to the perpendicular detection). These anisotropies decayed with a rotational correlation time $\tau \approx 40 ps$ of the CT complex, $r_i(t) = r_{0i}\exp(-t/\tau)$. Note that although the data are quite noisy, a lot of information can be extracted from this target analysis.

2.7.5.4 Anisotropy model for the BR photocycle

In §2.7.4 a general photocycle model was introduced. Here we extend this model with an anisotropy model for a special case, the BR photocycle. In order to include the parallel and perpendicular data we extend the model for the magic angle photocycle data from Eq. 60 by multiplying the concentration of each component *i* (ground state bleaching (GSB) or photocycle intermediates) by $1 + \gamma r_i(t)$, where γ equals 2, -1 for parallel and perpendicular data, respectively. The full model then reads [24]:

$$
\begin{bmatrix} \Delta A(t, \lambda) \\ \Delta A_{//}(t, \lambda) \\ \Delta A_{\perp}(t, \lambda) \end{bmatrix} = \sum_{l=1}^{n_{\mathrm{comp}}} c_l(t)\varepsilon_l(\lambda) \begin{bmatrix} 1 \\ 1 + 2r_l(t) \\ 1 - r_l(t) \end{bmatrix} - fc \cdot \tilde{\varepsilon}_0 \begin{bmatrix} 1 \\ 1 + 2r_{\mathrm{GSB}}(t) \\ 1 - r_{\mathrm{GSB}}(t) \end{bmatrix} \sum_{l=1}^{n_{\mathrm{comp}}} c_l(t) \qquad (67)
$$

For the time dependence of the anisotropy an exponential decay model can be used:

$$
r_i(t) = r_{0i}\exp(-k_{\mathrm{tumb}}t) \qquad (68)
$$

in which k_{tumb} is the rate of (membrane) tumbling, and r_{0i} is the anisotropy at time zero of component i. An alternative method for target analysis of BR photocycle anisotropy data has been developed by Borucki and Heyn [104,105]. Based upon mild assumptions they exploited the anisotropy dimension to estimate the SAS and concentration profiles.

Lozier et al. [21] measured five-way data of the BR photocycle: wavelength from 380–700 nm, time from 1μs – 0.3s, temperature from 5–35 ˚C, under four solvent conditions (pH 5, pH 7, pH 9, and pD 7) and under three polarization conditions (magic angle, parallel and perpendicular). A spectrotemporal model was applied to these data [24], which consisted of five parts: (i) a compartmental scheme (§2.4.3) for the MA concentrations with fully reversible transitions between the photocycle intermediates (Fig. 18A), (ii) the temperature dependence of the microscopic rate constants was described with thermodynamic parameters (Eq. 27, §2.4.3.6) for each of the solvent conditions, (iii) a photocycle model (Eq. 61, §2.7.4) which contains the SAS including the GSB and a fraction cycling parameter. The SAS are assumed to be temperature independent [106], (iv) an anisotropy model (Eq. 67 and Eq. 68), and (v) spectral assumptions (§2.7.2) on the SAS, i.e. that the M intermediates did not absorb above 540 nm, that the N and O intermediates did not contribute to the difference absorption below 460 nm, and that the L and N intermediates did not absorb above 680 nm. Relative to the analysis in [21], parts (iii) and (v) were the major improvements. This model consistently described the five-way data. The estimated SAS are shown in Fig. 18B. From the thermodynamic parameters the free energy changes during the photocycle can be calculated, and ΔG relative to the first intermediate is depicted in Fig. 18C. Thus this target analysis is instrumental in monitoring the energetics of the BR photocycle. Although this photocycle has been studied for thirty years, there is still some controversy [22,23,24]. In an alternative approach, all possible kinetic schemes using a full K-matrix (Eq. 18) were tested [20], successively eliminating superfluous microscopic rate constants. The method was tested on
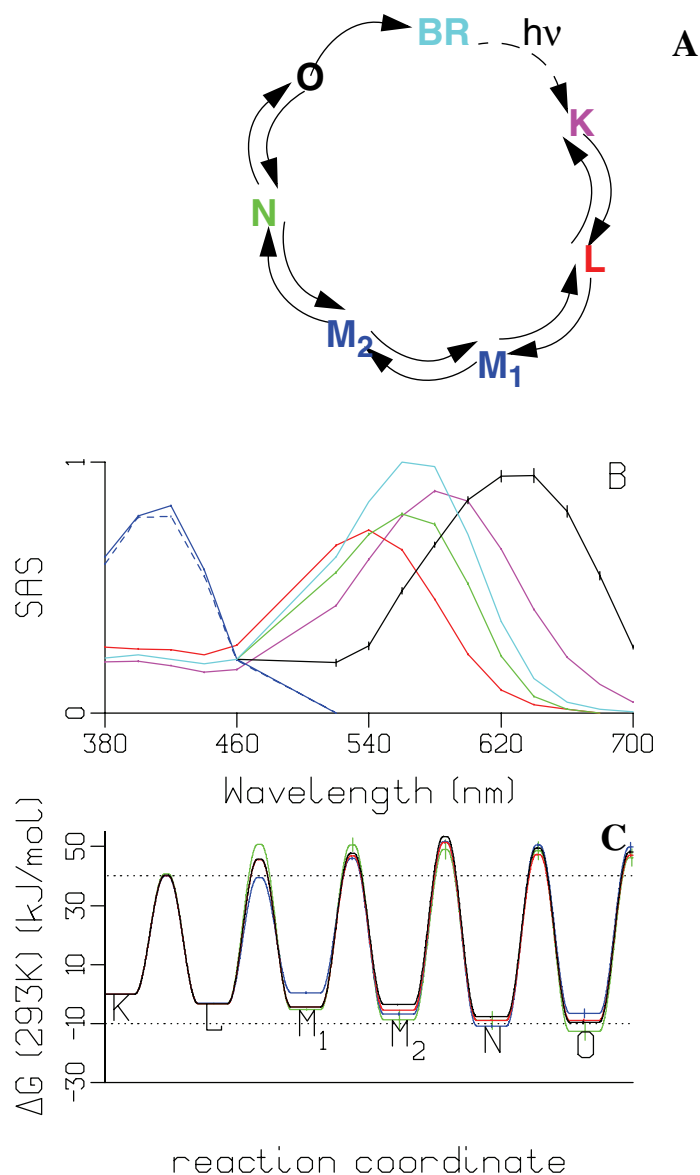
Fig.18. Target analysis of bacteriorhodopsin photocycle, using a reversible compartmental scheme (A) with six intermediates $K \rightleftharpoons L \rightleftharpoons M1 \rightleftharpoons M2 \rightleftharpoons N \rightleftharpoons O \rightarrow BR$.

(B) Estimated SAS, note the congestion of K,L,N and BR. Key: magenta K, red L, blue M, green N, black O and cyan BR. (C) Free energy changes relative to K (at 293K) during the photocycle. Key: black pH 5, red pH 7, blue pH 9, green pD 7. Vertical bars indicate plus or minus standard error.

concentration profiles estimated from resonance Raman data, and arrived at a similar kinetic scheme [90], except for a few small differences. An alternative description with two parallel, irreversible photocycles is proposed in [22,23]. This exemplifies that the explanation of multiexponential decay of components by equilibria or by heterogeneity (§2.4.3.3) is a

recurring theme in target analysis.

2.7.6 Multi-pulse excitation models

Multi-pulse spectroscopy can provide extra information with which complicated compartmental schemes can be unravelled [7]. When a second pulse interacts with an excited state three processes can occur: (i) when the state is emissive at the wavelength of the second exciting pulse, stimulated emission occurs, and the excited state returns to the ground state or to a ground state intermediate (GSI), which subsequently relaxes to the ground state. In this case the second pulse is termed a dump pulse. (ii) When the state shows excited state absorption at the wavelength of the second exciting pulse a higher excited state is created, which subsequently relaxes, possibly through the same or very similar intermediate states. (iii) When an excited state is already present, and a new excited state is created, singlet-singlet annihilation can occur. To isolate the effect of the second exciting pulse from a possible ground state excitation, the effect of the second pulse alone is subtracted from the measurement with both exciting pulses. Formally this can be described as follows:

$$\Delta A_{PP}(t, \lambda) \ = \ \left(\sum_{l=1}^{n_{\text{comp}}} c_{l,PP}(t) \varepsilon_l(\lambda) \right) \oplus i_{PP}(t) \tag{69}$$

where the subscript PP indicates the ordinary pump-probe difference absorption experiment. Now the second pulse (which we call here the dump pulse) has an additional effect, which can be approximated by:

$$\Delta A_{PDP}(t, \lambda) \ = \ \left(\sum_{l=1}^{n_{\text{comp}}} c_{l,PDP}(t) \varepsilon_l(\lambda) \right) \oplus i_{PDP}(t - \Delta\mu) \tag{70}$$

with $i_{PDP}(t - \Delta\mu)$ being the instrument response function of the dump pulse administered after an interval $\Delta\mu$ with respect to the first, pump pulse. The interaction of the dump pulse with an Excited State Intermediate (ESI) results in the disappearance of part of the population of the ESI and possible appearance of GSI or another, higher ESI. Thus the compartmental scheme of the $c_{l,PDP}(t)$ is in general extended with these new states. In the ideal case the dump pulse results in the enhanced population of a GSI, and introduces no new states. When this GSI decays faster than it is formed by normal decay of the ESI, the dump pulse uncovers the GSI. This is the case with the Green Fluorescent Protein [107,108]. After excitation of

GFP a proton is transferred resulting in a long lived (ns) excited state, which is responsible for the green fluorescence. A second pulse resonant with this emission dumps this excited state resulting in a GSI. This GSI relaxes in two steps of about 3 and 440 ps, whereby the proton is transferred back to the ground state. The pump-dump-probe technique in combination with the target analysis reveals the dynamics and the SADS of the states involved in this ground state proton transfer.

3 Parameter estimation

3.1 Simple linear regression

To better understand the variable projection algorithm which will be explained below, it is helpful to first consider a simpler problem of fitting an observed spectrum $\underline{\psi}(\lambda)$ as a linear combination of a priori known component spectra:

$$\underline{\psi}(\lambda) = \sum_{l=1}^{n_{comp}} c_l \varepsilon_l(\lambda) + \underline{v}(\lambda) \tag{71}$$

where c_l and $\varepsilon_l(\lambda)$ denote, respectively, the unknown concentration and the known spectrum of component l. $\underline{v}(\lambda)$ is additive noise, and the underscore is used to indicate that it is a stochastic quantity, the noise is a random variable. And because of the noise the observations are stochastic too. In vector and matrix form this equation reads:

$$\underline{\Psi}_{\lambda_j} = \sum_{l=1}^{n_{comp}} \varepsilon_{\lambda_j l} c_l + \underline{v}_{\lambda_j} \qquad \underline{\psi} = Ec + \underline{v} \tag{72}$$

$\varepsilon_{\lambda_j l}$ denotes the contribution of component l at wavelength λ_j, and is the jl-th element of the spectral matrix E, of dimensions $n_\lambda \times n_{comp}$. $\underline{\psi}$ and $\underline{v}$ are column vectors of length n_λ, representing the stochastic observations and additive noise. Note that when the spectra of the components are known this represents a linear regression. The vector c contains the n_{comp} concentrations of the components which are the unknown parameters to be estimated. Note that n_λ must be greater than n_{comp}; otherwise the problem is underdetermined, i.e. there are less observations than parameters. Formally, the least squares estimates are

$$\hat{\underline{c}} = (E^T E)^{-1} E^T \underline{\psi} \equiv E^\dagger \underline{\psi} \tag{73}$$

where $E^\dagger$ is called the Moore-Penrose generalized inverse of E, and the circumflex ^ denotes "estimator of". Note that this estimator is unbiased:

$$EXP[\hat{\underline{c}}] = EXP[(E^T E)^{-1} E^T \underline{\psi}] = (E^T E)^{-1} E^T EXP[\underline{\psi}] = (E^T E)^{-1} E^T Ec = c \qquad (74)$$

where $EXP[\ ]$ denotes expectation, and we have used Eq. 72 to find $EXP[\underline{\psi}] = Ec$.

It is usually assumed that the n_λ observations, which are stochastic because of the additive noise, are statistically independent and have a common variance σ^2. Because we assume additive normally distributed noise (§2.1.2) the least squares estimator is also the maximum likelihood estimator, which in the ideal case results in maximal parameter precision [25, 26]. The statistical properties of a vector of stochastic variables are expressed by the (auto)covariance matrix, which contains on its diagonal the variances of all elements of the vector, whereas an off-diagonal element is the covariance of two elements of the vector. The covariance of two random variables is the product of their standard deviations and their correlation coefficient:

$$\text{cov}(\underline{v}_i, \underline{v}_j) = \sigma_i \sigma_j r_{ij} \qquad (75)$$

which in this case is zero when i and j differ, because the observations are statistically independent, and equal to the variance σ^2 when i equals j.

The correlation coefficient r between random variables x and y is estimated with

$$\hat{\underline{r}}_{x,y} = \frac{\frac{1}{n}\sum_{i=1}^{n}(\underline{x}_i - \bar{\underline{x}})(\underline{y}_i - \bar{\underline{y}})}{\underline{s}_x \underline{s}_y} \qquad (76)$$

where the mean and the variance are estimated as usual: $\bar{\underline{x}} = \frac{1}{n}\sum_{i=1}^{n}\underline{x}_i$ and $\underline{s}_x^2 = \frac{1}{n}\sum_{i=1}^{n}(\underline{x}_i - \bar{\underline{x}})^2$. Illustrations of correlations are shown in Fig.19. The plot at the right

Fig.19. Examples of correlations between random variables *x* and *y*.

hand side of Fig. 19 shows a positive correlation.

Thus the (auto)covariance matrix of the vector $\underline{v}$ is given by

$$\text{cov}(\underline{v}) = \sigma^2 I_{n_\lambda} \tag{77}$$

where I_{n_λ} is the unity matrix of dimension n_λ. In that case the variance of the estimated parameters $\hat{\underline{c}}$ can be calculated from Eq. 73, because the estimator $\hat{\underline{c}}$ is a linear combination of the observations $\underline{\psi}$. Using the relation

$$\text{cov}(A\underline{y} + b) = A\text{cov}(\underline{y})A^T \tag{78}$$

we find by combining Eq. 73, Eq. 77 and Eq. 78:

$$\text{cov}(\hat{\underline{c}}) = E^\dagger \text{cov}(\underline{\psi})E^{\dagger T} = E^\dagger \sigma^2 I_{n_\lambda} E^{\dagger T} = \sigma^2 E^\dagger E^{\dagger T} = \sigma^2 (E^T E)^{-1} \tag{79}$$

Now what have we gained so far? We started with n_λ observations, and provided the component spectra are sufficiently different we now describe the data with n_{comp} parameters $\hat{\underline{c}}$ with hopefully a small variance. To quantify the difference between two spectra i and j one can calculate the inner product of the two columns of the E matrix, and normalize it by dividing by their lengths, this will result in a (sort of) correlation coefficient r_{ij}. When $|r_{ij}|$ is close to 1 the spectra are considered to be very similar, and are termed collinear. Collinearity results in problems with the inversion of the $E^T E$ matrix (needed in Eq. 79) and introduces large variances of the accompanying parameters c_i and c_j.

The differences between the observations and the fitted values are called the residuals which are given by:

$$\underline{z} = \underline{\psi} - \hat{\underline{\psi}} = \underline{\psi} - EE^\dagger \underline{\psi} = (I - EE^\dagger)\underline{\psi} \tag{80}$$

Now the matrices $EE^\dagger$ and $(I - EE^\dagger)$ are projection matrices. $EE^\dagger$ projects on the space spanned by columns of the E matrix, whereas $(I - EE^\dagger)$ projects on the space orthogonal to it. The geometry of linear least squares is illustrated in Fig. 20 . The sample space is an n-dimensional space which contains the observation vector y and the column vectors of X. An example with $n = 3$ and $q = 2$ is shown, where X is an $n \times q$ matrix and the model $y = X\theta$ has q unknown parameters. This model has the same form as $\psi = Ec$ from Eq. 72. The tilted plane shown in Fig. 20 represents the space spanned by columns of the X (or E) matrix (depicted by vectors P_1 and P_2). The observation or measurement vector y (or ψ) is projected onto this plane, resulting in the projection $\hat{y}$. The plane is also called *expectation plane*, because with different θ each point on the plane can be reached. The difference of the two vectors, $\underline{z} = \underline{y} - \hat{\underline{y}}$, is the residual vector which is orthogonal to the plane. It is obvious
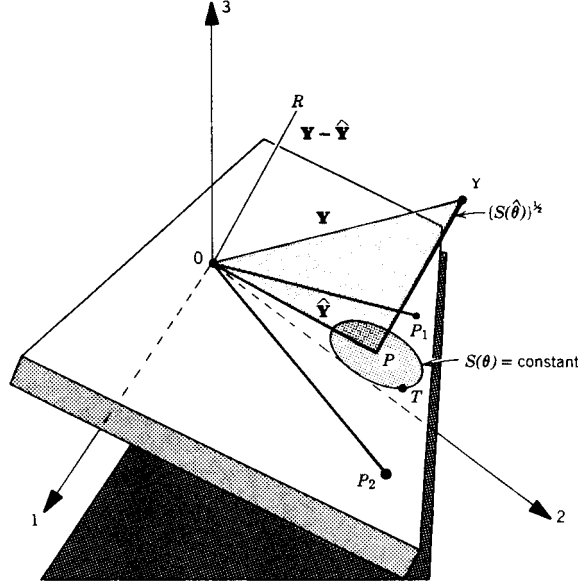
Fig.20. Example of geometry of linear least squares in sample space.

that this orthogonal projection results in the shortest possible residual vector, and thus the θ

that corresponds to the point *P* is the least squares estimate.

3.1.1 Least squares estimation using the QR decomposition

Now to actually calculate the least squares estimates a matrix decomposition of the design

matrix (*X* or *E)* is helpful. The advantage of using the QR decomposition (or alternatively, the

SVD) is numerical stability. In this section we use a matrix *X* of dimensions $n \times q$, and the

model used is $y = X\theta$ as illustrated in Fig. 20. Let the QR decomposition of the $n \times q$ matrix

X be

$$X = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} \tag{81}$$

R is a $q \times q$ upper triangular matrix. Q_1 and Q_2 are, respectively, $n \times q$ and $n \times (n-q)$

matrices such that $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ is an $n \times n$ orthogonal matrix, i.e. both columns and rows

are orthogonal:

$$Q^T Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^T \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} = \begin{bmatrix} Q_1^T Q_1 & Q_1^T Q_2 \\ Q_2^T Q_1 & Q_2^T Q_2 \end{bmatrix} = \begin{bmatrix} I_q & 0 \\ 0 & I_{n-q} \end{bmatrix} = I_n \tag{82}$$

and of length 1. From Eq. 81 we have

$$X = Q_1 R \tag{83}$$

which is termed the skinny QR decomposition. The QR decomposition is illustrated in Fig. 21, where a sample space (with 3 samples) and a tilted *expectation plane* is transformed into
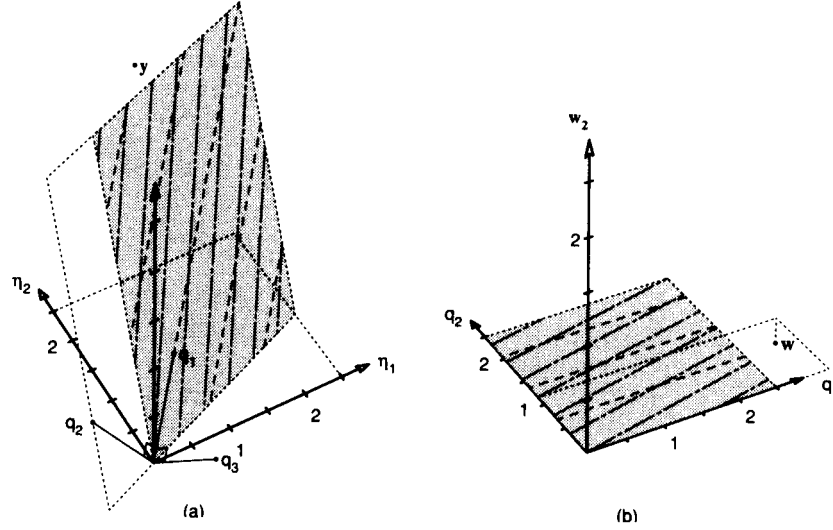

Fig.21. Expectation surface for 3 data points η_i and 2 parameters. Part (a) shows the expectation plane (shaded) in the response space together with the response vector y. Also shown are the orthogonal unit vectors q_1 and q_2 (in the plane) and q_3 (orthogonal to the plane). Part (b) shows the same in the rotated coordinates given by Q, $w = Q^T \underline{y}$.

Q-space. Note that now the residuals are along the $w_2 = Q_2^T \underline{y}$ axis. Using $Q_1^T X = R$ we have $Q_1^T \underline{y} = R\hat{\theta}$, a triangular system which can easily be solved for $\hat{\theta}$. The Moore-Penrose generalized inverse is given by

$$X^\dagger = R^{-1} Q_1^T \tag{84}$$

Thus $X^\dagger X = I_q$ and $XX^\dagger = Q_1 Q_1^T$ is the orthogonal projection operator which projects the observation vector $\underline{y}$ onto the column space of Q_1 (and thus of X), cf. the tilted plane shown in Fig. 20. Likewise, the residual vector is orthogonal to this plane and is given by

$$\underline{y} - \hat{\underline{y}} = (I - XX^\dagger)\underline{y} = Q_2 Q_2^T \underline{y} \tag{85}$$

The sum of squares of errors SSE is given by the square of the length of this residual vector:

$$SSE = (\underline{y} - \hat{\underline{y}})^T (\underline{y} - \hat{\underline{y}}) = (Q_2 Q_2^T \underline{y})^T (Q_2 Q_2^T \underline{y}) = \underline{y}^T Q_2 Q_2^T Q_2 Q_2^T \underline{y} = \underline{y}^T Q_2 Q_2^T \underline{y} \tag{86}$$

where we have used Eq. 82. Theoretically, the expectation of the SSE equals $(n-q)\sigma^2$, because the dimension of the Q_2 space is $n-q$. This is because the residual vector has no

component in Q_1 space. We can use this to estimate an unknown σ^2:

$$\hat{\underline{\sigma}}^2 = \frac{SSE}{n-q} = \frac{SSE}{\text{df}} \tag{87}$$

In words: the estimator of the variance is equal to the residual sum of squares divided by the number of degrees of freedom (df) for this sum, $n - q$. The estimate of σ^2 can be used in Eq. 79 to estimate the covariance matrix of the estimated parameters $\hat{\underline{c}}$. The parameter uncertainty can also be visualized. In parameter space confidence intervals and regions can be drawn, based on the increase in SSE when a parameter is changed from the estimate. These are shown as horizontal and vertical lines in Fig. 22a, whereas the data which were fitted with a straight
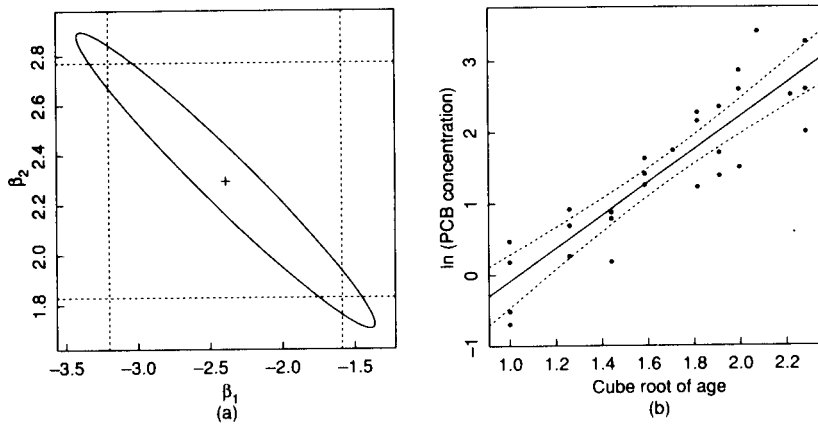

Fig.22. Inference regions. In (a) the least squares estimates (+), the 95% confidence region (solid line) and the 95% confidence intervals (dotted lines) are depicted. The data points and fitted response are shown in (b) as a solid line, the dotted lines indicate the 95% inference band. β_2 is the slope of the line, whereas β_1 is the intercept.

line $y = \beta_1 + \beta_2 x$ are shown in Fig. 22b. Also shown is the 95% confidence region, which is an ellipse. The negative slope of the long axis of the ellipse indicates a strong anticorrelation of the parameters β_1 and β_2, an increase in β_1 can be compensated by a decrease of β_2. In the transformed sample space $Q^T y$ a confidence interval can be calculated as those values of β for which the test $\beta = \beta^0$ is not rejected, which is depicted in Fig. 23. This is based upon the ratio of tangential and orthogonal components of the residual vector. Finally, Fig. 24 visualizes the relation between a disk on the expectation plane in sample space, for all points on this disk the SSE is decreased by no more than a certain amount, and a confidence region in
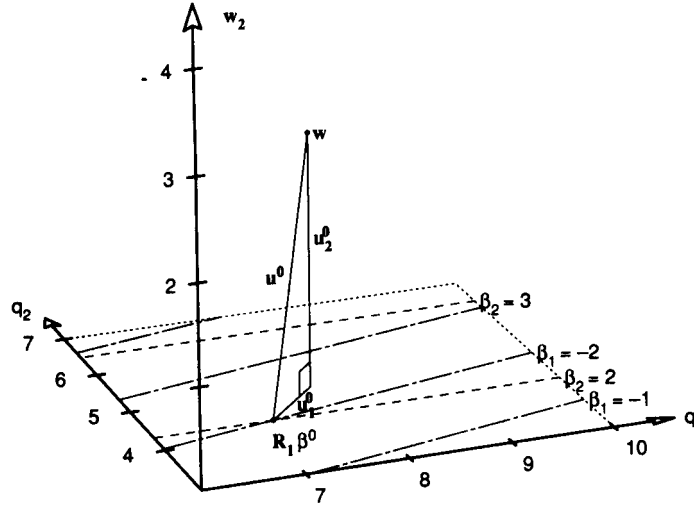
Fig.23. Geometric interpretation of the test of $H_0: \beta = \beta^0$. A portion of the expectation plane is shown, projected into the three dimensional space given by q_1, q_2 and w_2. For the test point β^0 the residual vector u^0 is decomposed into a tangential component u_1^0 and an orthogonal component u_2^0.
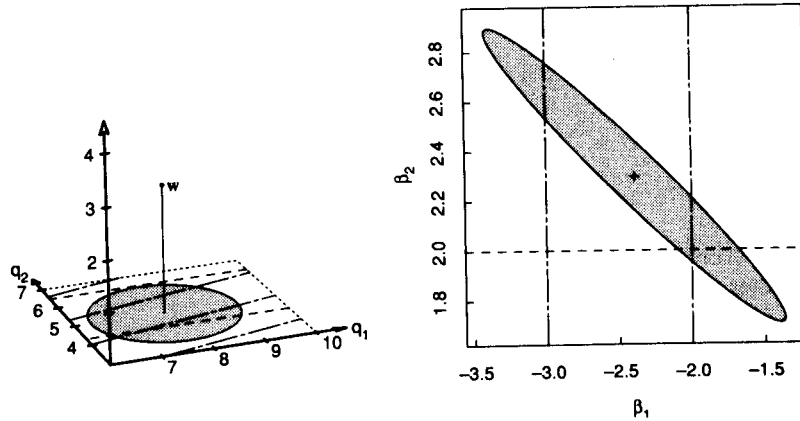
Fig.24. The 95% confidence region is a disk in (q_1, q_2, w_2) sample space (left) and an ellipsoid in parameter space (right) (β are the linear parameters).

parameter space, which is an ellipse.

For further explanation the reader is referred to Chapter 1 of Bates and Watts [26].

3.2 Multiple linear regression

Let us now generalize the model from Eq. 71 and Eq. 72 to multiple observed spectra, say n_{obspec}. Disregarding the noise, the model for the matrix of observations then reads

$$\Psi^T \;=\; EC^T \tag{88}$$

which is the transposed version of Eq. 23. Note that now all columns of the $n_{\text{comp}} \times n_{\text{obspec}}$ matrix C^T are unknown parameters, which can be estimated by the generalization of Eq. 73:

$$\hat{C}^T \;=\; E^{\dagger}\Psi^T \tag{89}$$

Note that the c-parameters are independent for each observed spectrum. There is no global analysis here, and consequently there can be no gain in precision of the c-parameters. However, if we would introduce a model for the C matrix, we would arrive at a spectrotemporal model (§2.7.1) and could estimate a smaller number of e.g. kinetic parameters more precisely.

3.3 Nonlinear regression model

As an intermediate step towards the variable projection algorithm, we consider the model where instead of the spectra, the concentrations are known a priori. Disregarding the noise, the model for the matrix of observations then reads

$$\Psi \;=\; CE^T \tag{90}$$

which is the noise-free version of Eq. 23. Note that now all columns of the matrix E^T are unknown parameters, which by analogy with Eq. 89 are estimated by:

$$\hat{E}^T \;=\; C^{\dagger}\Psi \tag{91}$$

where $C^{\dagger}$ is the Moore-Penrose generalized inverse of C. Thus if we would have a model for the C matrix, we could estimate the unknown spectral parameters by means of linear regression. The model for Ψ now reads (see Eq. 25)

$$\Psi \;=\; C(\theta)E^T \tag{92}$$

E^T are now called conditionally linear parameters [26], because they can be solved for conditionally on the unknown parameters θ, using Eq. 91. Since the model for the C matrix depends on θ in a non-linear way, there is no closed form solution like with linear regression, and iterative numerical methods have to be used. There are several methods. Methods based on evolutionary optimization techniques, like genetic algorithms [109], are unattractive, because the total number of unknown parameters is generally huge, which precludes interactive data analysis. Most applicable to the problem domain are non-linear regression

methods, which come in different flavours [110]. These methods distinguish between the conditionally linear parameters and the intrinsically nonlinear parameters θ. The alternating least squares method involves alternating between multiple linear regression (Eq. 91) and optimizing θ assuming the previously estimated $\hat{E}^T$. The variable projection method combines these two steps by implicitly solving Eq. 91, and directly minimizing the SSE. This is detailed in the next section. The number of unknown parameters that need to be estimated can easily amount to several thousand, e.g. estimation of many spectra at hundreds of wavelengths with the help of a kinetic model (Eq. 38). Therefore it is crucial to reduce the dimensionality of the parameter search space by implicitly solving for the conditionally linear parameters. This is done by the Variable Projection algorithm developed in the 1970s [95, 111, 112], which later became widely used in global analysis [20, 113, 114]. This algorithm exploits the bilinear structure of the model function.

Crucial for precise parameter estimation is careful weighting of the observations [36]. As discussed in §2.1.2, in some cases the weights are known a priori, in other cases they must be estimated as well, a procedure which is called *iteratively reweighted least squares* [35].

3.4 Variable Projection

Now everything is ready for introducing the main algorithm. Recall that, after formulation of a model for the observations (Eq. 38, Eq. 39, Eq. 53, Eq. 56, Eq. 57, Eq. 61) the unknown parameters have to be estimated from the data. It is important to recognize that, as discussed in §3.3, these equations represent a separable nonlinear model, also called a partially linear model [25]. This means that, conditionally on the intrinsically nonlinear parameters θ, the conditionally linear parameters [26] (E in Eq. 38, C in Eq. 39, A in Eq. 54, a in Eq. 55, $\tilde{E}$ in Eq. 56, $\tilde{C}$ in Eq. 57, $\tilde{E}$ and fc in Eq. 61) can be solved for (see e.g. §3.3) using the variable projection algorithm. First we present the equations for the kinetic model (Eq. 38). Conditionally on the kinetic parameters θ, Eq. 38 represents a linear model. Therefore the spectral parameters are estimated by (Eq. 91):

$$\hat{E}^T(\theta) \; = \; C^{\dagger}(\theta)\Psi \; = \; R^{-1}(\theta)Q_1^T(\theta)\Psi \tag{93}$$

where $C^{\dagger}(\theta)$ is the Moore-Penrose generalized inverse of $C(\theta)$, and where we have used the QR decomposition of $C(\theta)$ (Eq. 83) resulting in $C^{\dagger}(\theta)$ (Eq. 84). Using Eq. 92 and Eq. 93 we find for the residual matrix Z:

$$Z(\theta) = \Psi - C(\theta)\hat{E}^T(\theta) = (I - Q_1(\theta)Q_1^T(\theta))\Psi = Q_2(\theta)Q_2^T(\theta)\Psi \qquad (94)$$

where $Q_2(\theta)Q_2^T(\theta)$ is an orthogonal projection matrix, cf Eq. 85. Referring back to Fig. 20 we can now interpret the name variable projection: the columns of the $C(\theta)$ span the expectation plane (the tilted plane in Fig. 20) and the data are projected onto this plane. Optimizing the unknown parameters θ means that this plane becomes variable, optimizing the columns of $C(\theta)$ in order to span the best possible plane. Now this can be generalized to arrive at a Partitioned Variable Projection algorithm, which is applicable for all models that have been discussed, including models for multiple data sets. The only exception is the BR photocycle models of Eq. 61 or Eq. 67, where all observations must be dealt with simultaneously, because all wavelengths are linked through the fc parameter, therefore no partitioning is possible, and the optimization is most compute intensive.

Let the data be decomposed into P parts:

$$\psi_p = X_p(\theta)\beta_p \qquad p = 1, \ldots, P \qquad (95)$$

where each part ψ_p contains n_p observations, e.g. each column of Ψ in Eq. 92 which represents a trace, or each column of Ψ^T in Eq. 92 which represents a spectrum, or all observations at a particular wavelength linked in more complicated spectrotemporal models, e.g. Eq. 62. Eq. 95 further contains b_p conditionally linear parameters β_p, and an $n_p \times b_p$ model matrix $X_p(\theta) = Q_1 R$. For each part, the β_p are estimated as:

$$\hat{\beta}_p = X_p^{\dagger}(\theta)\psi_p = R^{-1}Q_1^T\psi_p \qquad (96)$$

and the residuals are:

$$\psi_p - X_p(\theta)\hat{\beta}_p = (I - Q_1 Q_1^T)\psi_p = Q_2 Q_2^T\psi_p \qquad (97)$$

The sum of squares to be minimized as a function of the unknown parameters θ is given by, cf Eq. 86

$$SSE(\theta) = \sum_{p=1}^{P} \psi_p Q_2 Q_2^T \psi_p \qquad (98)$$

Thus elimination of the $\sum_{p=1}^{P} b_p$ conditionally linear parameters results in the $\sum_{p=1}^{P} n_p - b_p$ residuals $Q_2^T(\theta)\psi_p$ which leaves as degrees of freedom: df $= (\sum_{p=1}^{P} n_p - b_p) - n_{par}$, where n_{par} is the number of intrinsically nonlinear parameters θ. Numerical minimization of $SSE(\theta)$ in Eq. 98 requires the derivative $\frac{\partial}{\partial\theta}Q_2^T(\theta)\psi_p$. Again

there are different methods [110]. Most straightforward, but not so precise, is a finite difference derivative. more precise is an analytical derivative, which can also be approximated. Following Kaufman [112] and Golub and Leveque [113] we recall from Eq. 82 that $Q_2^T(\theta)X_p(\theta) = 0$ and thus

$$0 = \frac{\partial}{\partial\theta}(Q_2^T(\theta)X_p(\theta)) = \left(\frac{\partial}{\partial\theta}Q_2^T(\theta)\right)X_p(\theta) + Q_2^T(\theta)\left(\frac{\partial}{\partial\theta}X_p(\theta)\right) \tag{99}$$

This leads to the approximation $\frac{\partial}{\partial\theta}Q_2^T(\theta) \approx -Q_2^T(\theta)\left(\frac{\partial}{\partial\theta}X_p(\theta)\right)X_p^\dagger(\theta)$ which when multiplied with ψ_p and using Eq. 96 gives the desired (approximate) derivative of the residuals:

$$\frac{\partial}{\partial\theta}Q_2^T(\theta)\psi_p \approx -Q_2^T(\theta)\left(\frac{\partial}{\partial\theta}X_p(\theta)\right)\hat{\beta}_p \tag{Eq. 100}$$

According to [110] the Kaufman approximation is very good, but this still has to be investigated for the Partitioned Variable Projection algorithm. Note that the model connects the parts ψ_p via the unknown intrinsically nonlinear parameters θ. The Partitioned Variable Projection algorithm is an example of a divide-and-conquer strategy. Suppose there are $P = 1000$ parts, each with $n_p = 1000$ observations, and that each part has $b_p = 10$ conditionally linear parameters. Then, without partitioning the model matrix would have dimensions 1000000×10000 (which will rarely fit in computer memory). With partitioning, however, the dimensions of each of the 1000 X_p model matrices are 1000×10 (which will easily fit in computer memory). The computational savings of this partitioning are at least a factor of 1000. Only in this way a million observations can be simultaneously analysed. Standard nonlinear regression algorithms like Gauss-Newton or Levenbergh-Marquardt [25, 26] can be used to minimize $SSE(\theta)$.

3.5 Convergence and summary statistics

After convergence of the nonlinear least squares fit routine a further check of the neighbouring parameter space can be done using new starting values for the unknown parameters. This is elaborately discussed in [17]. (describe ...) Having estimated the parameters careful checking of the residuals is of paramount importance. With matrices of data, and thus also of residuals, the SVD is an ideal tool to check for structured residuals [102]. When the residuals are satisfactory, the parameter precision needs to be investigated. Linear approximation standard

errors can be calculated from the Jacobian of the model function. The Jacobian is the derivative of the model function with respect to all parameters. It is the analogon of the design matrix X (or E in Eq. 70) in linear regression. With the Variable Projection algorithm the situation is somewhat more complicated, and the covariance matrix of the conditionally linear parameters has to be calculated separately. The usual linear approximation summary statistics are based upon a Taylor expansion of the criterion function $SSE(\theta)$ around the maximum likelihood estimates $\hat\theta$ and $\hat\beta_p$. First the Jacobian has to be calculated:

$$J = \begin{bmatrix} \dfrac{\partial}{\partial\theta} \\ \dfrac{\partial}{\partial\hat\beta_P} \end{bmatrix} X_p(\theta)\beta_p = \begin{bmatrix} \left(\dfrac{\partial}{\partial\theta}X_p(\theta)\right)\beta_p \\ X_p \end{bmatrix} = \begin{bmatrix} \partial_\theta X_p \beta_p \\ X_p \end{bmatrix} \tag{101}$$

omitting the θ dependence for clarity. Analogous to Eq. 79 the full linear approximation covariance matrix is given by:

$$\mathrm{cov}\left(\begin{bmatrix} \theta \\ \hat\beta_P \end{bmatrix}\right) = \hat\sigma^2 (J^T J)^{-1} \tag{102}$$

with $\hat\sigma^2$ the variance estimate $\hat\sigma^2 = SSE(\hat\theta)/\mathrm{df}$ (cf Eq. 87). We now need to invert the block matrix $J^T J$:

$$J^T J = \begin{bmatrix} (\partial_\theta X_p \beta_p)^T \partial_\theta X_p \beta_p & (\partial_\theta X_p \beta_p)^T X_p \\ X_p^T \partial_\theta X_p \beta_p & X_p^T X_p \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{103}$$

Then we can use a matrix inversion theorem (e.g. Eq.2 in appendix A3 of [25]) to find the lower right hand block of

$$(J^T J)^{-1} = \begin{bmatrix} C_{11}^{-1} & -C_{11}^{-1}C_{12} \\ -C_{21}C_{11}^{-1} & A_{22}^{-1} + C_{21}C_{11}^{-1}C_{12} \end{bmatrix} \tag{104}$$

where $\sigma^2 C_{11}^{-1} = \mathrm{cov}(\hat\theta)$ which is returned by the nonlinear regression routine. Furthermore, using $A_{22}^{-1} = X_p^\dagger X_p^{\dagger^T}$ we find

$$C_{21} = A_{22}^{-1} A_{21} = X_p^\dagger X_p^{\dagger^T} X_p^T \partial_\theta X_p \hat\beta_p = X_p^\dagger \partial_\theta X_p \hat\beta_p \equiv G$$
$$C_{12} = A_{12}A_{22}^{-1} = G^T \tag{105}$$

where matrix G consists of, for each parameter θ_i, a column $X_p^\dagger \dfrac{\partial}{\partial\theta_i}X_p\hat\beta_p$.

Applying these equations we find for the kinetic model (Eq. 38) that the approximate covariance matrix of ε_λ (which is the λ-th row of $\hat{E}$) is given by:

$$\text{cov}(\hat{\varepsilon}_\lambda) \;=\; \hat{\sigma}^2 (C^\dagger C^{\dagger T}) + G\,\text{cov}(\hat{\theta})\,G^T \tag{106}$$

with $n_p \times n_{\text{par}}$ matrix G consisting of, for each parameter θ_i, a column $C^\dagger \dfrac{\partial C}{\partial \theta_i}\hat{\varepsilon}_\lambda$.
In a simulation study [98] (see §3.5.1) it was shown that the linear approximation standard error is adequate for the intrinsically nonlinear parameters (in this case the kinetic parameters θ), thus the model is functionally linear [115]. However, the linear approximation standard error is less adequate for the conditionally linear parameters (in this case the spectral parameters ε_λ) [116]. In that case likelihood-based confidence regions can be calculated [26, 39, 116].

3.5.1 Simulation study of parameter precision

By means of simulation parameter estimation in global analysis of time-resolved spectra has been studied [98,116]. Here we present the results for simulated data with two components. The concentrations of the components are described by exponential decays $\exp(-kt)$ with rate parameter k, whereas the spectral shapes are described by a Gaussian in the energy domain (Eq. 30), with parameters $\bar{\nu}_{\text{max}}$, $\Delta\bar{\nu}$ for, respectively, location and Full Width at Half Maximum (FWHM). Thus the simulated data are a function of eight parameters: for each component four parameters: k, $\bar{\nu}_{\text{max}}$, $\Delta\bar{\nu}$ and amplitude a (f_{max} in Eq. 30). The time-resolved spectrum was simulated at $m = 51$ time points equidistant in the interval 0-2 ns and $n = 51$ wavelengths equidistant in the interval 350-550 nm. The overlap between the spectra and concentration profiles of the components could be large (EL,CL) or small (ES,CS). The nonlinear parameters of the combinations used in the simulations, whose values are inspired by experimental data [34, 69], are summarized in Table 1. Normally distributed noise was

Table 1. Parameters k (in 10^9s^{-1}), $\bar{\nu}_{\text{max}}$, $\Delta\bar{\nu}$ (in 10^3 cm^{-1}) and a of components.

Overlap	CSmall (CS)	CLarge (CL)
ESmall (ES)	0.25, 22, 9, 1	0.5, 22, 9, 1
	1.0, 18, 8, 2	0.6, 18, 8, 2
ELarge (EL)	0.25, 19, 9, 1	0.5, 19, 9, 1
	1.0, 18, 8, 2	0.6, 18, 8, 2

added to the simulated data. With the low noise level the standard deviation σ of the noise

was equal to 6×10^{-3} of the maximum of the data (CS,ES and CS, EL cases, small kinetic

overlap) or 3×10^{-3} (CL,ES (see Fig. 25) and CL,EL cases, large kinetic overlap). The high

noise level was ten times higher.

In order to compare the parameter precision with different types of model an ensemble of

datasets was simulated with certain parameters and noise level. The deviation

$\mathrm{dev}(\theta) = \hat{\theta} - \theta$, the difference between the estimated and true value of a parameter, the

linear approximation standard error from Eq. 102, and the ratio of these two, which is the

studentized parameter [26,115], were calculated. From this ensemble of realizations the rms

value was calculated and a smoothed probability density was estimated using the Splus

function *ksmooth* [117]. Kinetic, spectral as well as spectrotemporal models were used to

describe a system consisting of an inherent mixture of components whose concentrations

change with time. A typical example of a global analysis with the help of a kinetic model of a
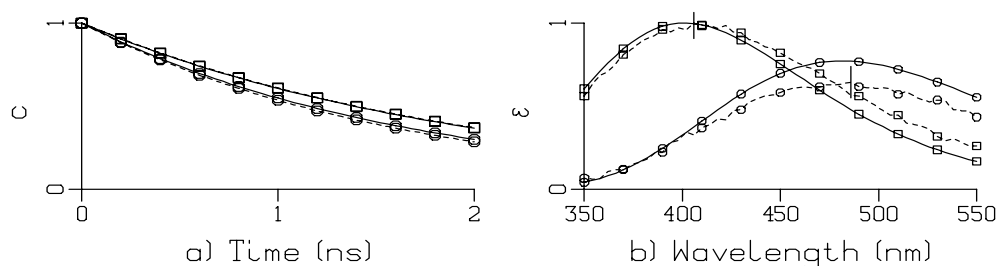

Fig.25. Global analysis with the help of a kinetic model of CL, ES combination data set

with low noise. (a) concentration profiles. Squares and circles indicate first and second

component, respectively. Solid and dashed lines indicate true and fitted, respectively. (b)

spectra, vertical bars indicate plus or minus standard error.

CL,ES combination data set with low noise is shown in Fig. 25. Note that the fitted curves

(dashed lines in Fig. 25a,b) are close to the simulated curves (solid lines). Fig. 26 shows the

distributions for k_1 estimated from an ensemble of realizations of this kinetic model fit. The

distribution (Fig. 26a) of deviation shows no signs of bias (peaking at zero deviation) but it

appears a bit skewed. The studentized parameter distribution (Fig. 26c) deviates somewhat

from a t_{df} distribution which would apply when the model is linear. The improvement gained

with a spectrotemporal model is clearly visible in Fig. 27 (note the differences in scale). Note

that here the studentized parameter is more closely distributed as t_{df} (Fig. 27c) which indicates that the spectrotemporal model is functionally linear [115]. Thus overlap combinations were studied at two noise levels. The averaged results for the two rate constants are collated in Table 2. With CS,EL overlap the spectrotemporal model (column ST) provides only a small improvement over the kinetic model (column K) . However, with small spectral overlap (CL,ES and CS,ES) the spectrotemporal model is clearly superior (compare the K and ST columns of Table 2), the precision is increased about tenfold (CL,ES) and twofold (CS,ES). Surprisingly, with CL,EL overlap (and low noise) the spectrotemporal model increased the precision about fivefold. With high noise the kinetic model could no longer resolve two components, which is indicated by the rms deviations being larger than the parameter values (Table 1). With the spectrotemporal model the precision was close to the CS,EL overlap, where the noise was two times larger, indicating a trade off between (temporal) overlap and noise level. With the spectral parameters the results (not shown) are analogous. With CL,ES overlap the spectrotemporal model provides only a small improvement over the spectral model. However, with small temporal overlap (CS,EL and CS,ES) the spectrotemporal model is clearly superior, providing an increase in the parameter precision by a factor of 1.5 to 10. Summarizing, with a single component the parameter precision does not differ between the different model types. With two or more components the precision decreases with a kinetic or spectral model because of overlap between concentration profiles or spectra of the components. However, with a spectrotemporal model the decrease is much less, and equals zero with zero spectral overlap.

Fig.26. Distributions estimated from kinetic fit of CL, ES combination with low noise. (a) deviation of estimated rate constant k_1, (b) standard error, (c) studentized parameter (solid) and t_{df} distribution (dotted).
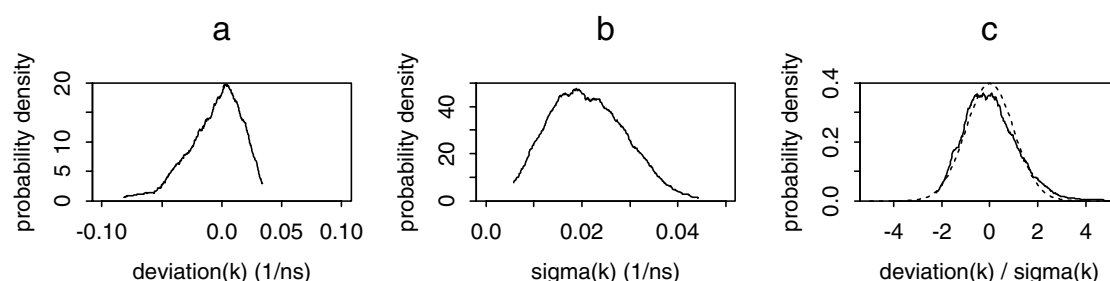
Fig.27. Distributions estimated from spectrotemporal fit of CL, ES combination with low noise. (a) deviation of estimated rate constant k_1, (b) standard error, (c) studentized parameter (solid) and t_{df} distribution (dotted).
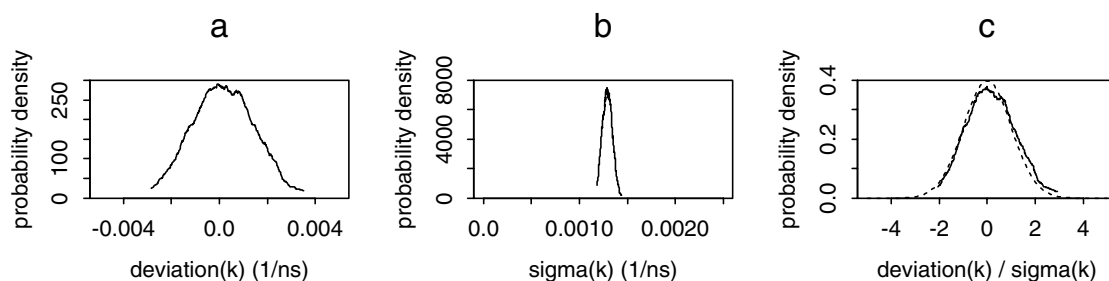

Table 2. Rms deviation of rate constants (in $10^6 s^{-1}$) with kinetic (K) or spectrotemporal (ST) model as a function of temporal and spectral overlap.

		CS, ES		CL, ES		CS, EL		CL, EL	
	component	K	ST	K	ST	K	ST	K	ST
low noise	1	5	2	24	1	14	13	78	9
	2	11	7	26	3	19	18	77	18
high noise	1	43	24	240	20	150	130	2200	130
	2	110	73	340	170	290	200	1800	230

3.6 Incorporating multiple experiments

A commonly occurring situation is when multiple experiments are done which are believed to be described by a single (most often kinetic) model, however, each experiment may possess some experiment-specific parameters. Examples are (i) measurement of time-gated spectra in two or more different wavelength ranges, or under different polarization angles; (ii) trace measurements in different time ranges; (iii) measurements with varying instrument responses. In most of these cases global analysis is still applicable, but extra nuisance parameters may have to be introduced. E.g. because of laser power fluctuations scaling parameters may be necessary to simultaneously analyse data from different experiments (case (ii)) [102]. This number of scaling parameters can become very large. The generalization of Eq. 61 to a simultaneous target analysis of multiple experiments is straightforward. For each extra experiment an overall scaling parameter is needed, to account for variations in the product of sample OD and intensity of the actinic flash. With the five way BR photocycle this amounts to 235 scaling parameters, which is much more than the number of kinetic and spectral

parameters [24].

3.7 Software

The modelling of time-resolved spectra is an iterative process, which benefits from a dedicated Problem Solving Environment [13,14,17,61] that incorporates a wide variety of models. The data analysis environment described in [13,14] concentrates on fluorescence data. The program described in [17] applies a simulated annealing strategy for estimation of the globally optimal model parameters. The implementation and application of the variable projection algorithm to global analysis was described in [113].

4 Case study: ultrafast dynamics of PYP

The purpose of this section is to apply the above presented methodology in a typical case study with real three-way data sets. A chromophore-protein complex, the photoactive yellow protein (PYP) [5, 6, 101,102,118], was studied by time-resolved polarized difference absorption spectroscopy.　The PYP chromophore (p-coumaric acid, pCA) is covalently bound to a 14 kDa protein, and then absorbs maximally at 446 nm (black spectrum in Fig. 34A). Upon excitation PYP traverses a photocycle that eventually leads to a signaling state, most probably the pB state (also called I_2) mentioned in §2.7.4. The dynamics of the PYP photocycle extend from femtoseconds to seconds. In this experiment PYP was excited at 400 nm (blue pump spectrum in Fig. 34A). The early photophysics and photochemistry of this model system are discussed in detail elsewhere [8]. Here we present the modelling of three types of ultrafast data: we start with the magic angle (MA) data, then we add anisotropy data, and finally we test the kinetic scheme with multipulse data.

4.1 Residual analysis of MA data

The MA data (typical traces are shown in Fig. 28) were globally analysed using a parallel kinetic scheme (§2.4.1, Eq. 4) with five lifetimes plus a coherent artefact. The uncorrected data (black) exhibit common noise, which is most clearly visible around 6, 100 and 400 ps. This is likely caused by baseline fluctuations in the data. To investigate this further, the matrix of residuals was subjected to Singular Value Decomposition (Eq. 37, §3). Fig. 29C shows that the first singular value is significantly larger than the remainder. The first right singular vector (w res, Fig. 29B) is almost flat in wavelength, whereas the first left singular vector (u res, Fig.
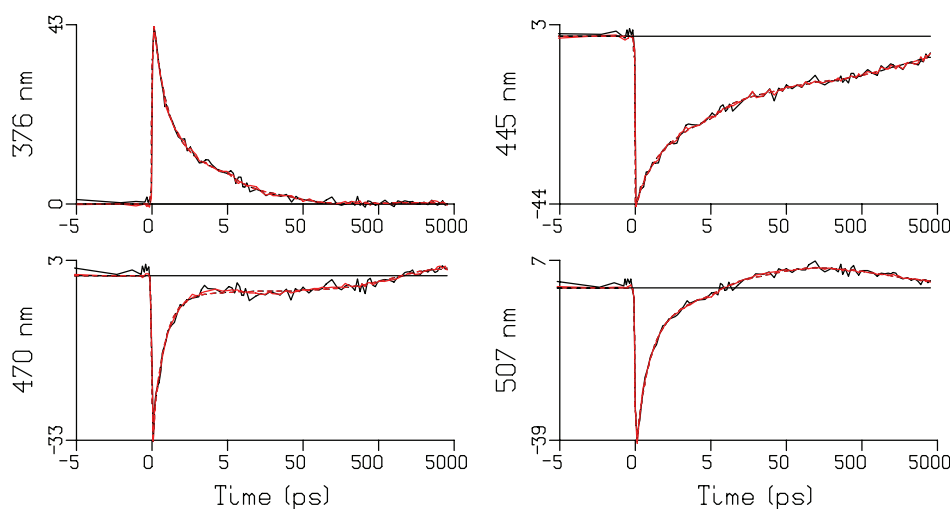
Fig.28. Selected difference absorption traces (in mOD) of PYP after 400 nm excitation, measured at magic angle. Probe wavelength indicated along the ordinate. Uncorrected (black) and after correction for baseline fluctuations (red). Dashed lines indicate fit. Note that the time axis is linear from -5 to +5 ps relative to the maximum of the IRF, and logarithmic thereafter.

29A) displays no clear structure in time. These baseline fluctuations can be estimated from the residuals, and subtracted from the data, thus refining the analysis. The refined data are depicted in red in Fig. 28. This procedure results in a decrease of the root mean square error (rmse) from 0.85 to 0.42 mOD. Note that most of the fluctuations have been corrected for, as evidenced from the difference between the black and red solid lines. The results of SVD of the residual matrix from the refined global analysis have been depicted in Fig. 29D-F. Note that there appears to be no significantly larger first singular value in Fig. 29F. Judging from this residual analysis, and from the estimated parameters (vide infra) this global analysis is considered satisfactory.

4.2 Global analysis of MA data: DADS and EADS

The results from the global analysis of the MA data are shown in Fig. 30, using respectively a parallel (left) or a sequential (right) kinetic scheme. The concentrations of the components are depicted in Fig. 30A,B. When using a parallel scheme the estimated Decay Associated Difference Spectra (DADS) are shown in Fig. 30C. Alternatively a sequential scheme with increasing lifetimes (§2.4.3.2) can be used resulting in Evolution Associated Difference Spectra (EADS), Fig. 30D. Note that both schemes result in exactly the same residuals and
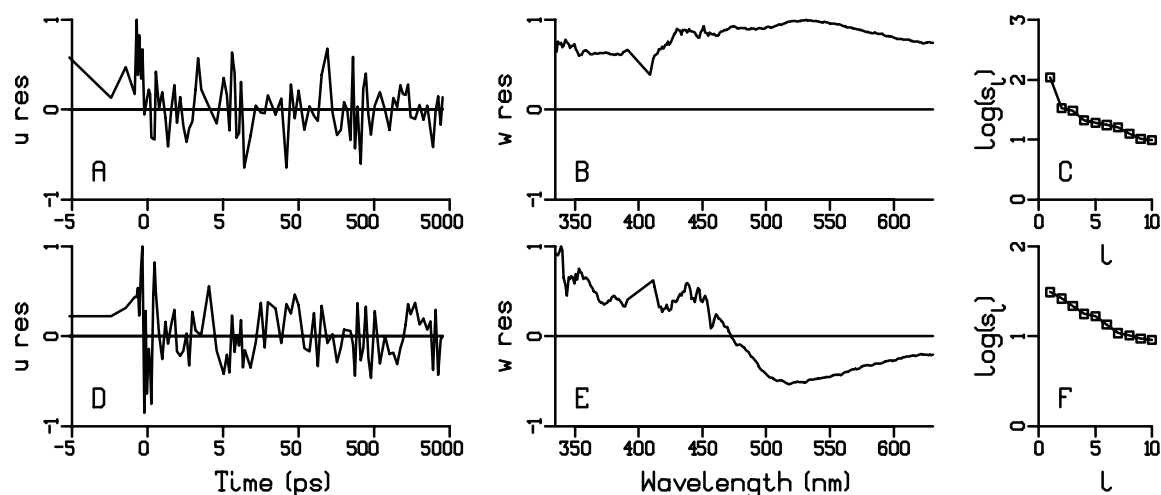
Fig.29. Results from Singular Value Decomposition of the residual matrix. Top panels: uncorrected data. Bottom panels: residuals from data after correction for baseline fluctuations. (A,D): first left singular vector, showing dominant temporal structure. (B,E): first right singular vector, showing dominant spectral structure. (C,F): Singular values on logarithmic scale.

quality of fit of the data (§2.6.2). The DADS are interpreted as loss or gain of absorption with a certain lifetime, whereas the EADS represent the spectral evolution, e.g. the third EADS rises with the second lifetime and decays with the third lifetime. With simple systems the interpretation of the DADS or EADS can be straightforward. E.g. when the sequential scheme with increasing lifetimes represents the correct physicochemical picture, the EADS correspond to true Species Associated Difference Spectra (SADS) characterizing the intermediate states. In that case the DADS represent decay and rise of these states. With the complicated system here at hand the interpretation becomes an iterative process, and in fact the target analysis described below allows for a better interpretation of the DADS or EADS. At this point a number of comments are in order. The first DADS (black) which decays in 0.8 ps shows a major loss of Excited State Absorption (ESA) from 340-420 nm and of Stimulated Emission (SE) from 460-630 nm. Less clear is a loss of Ground State Bleach (GSB, 420-460 nm). From the shape of the first DADS it cannot be decided which states rise with 0.8 ps. The second DADS (red) decays in 3.3 ps. Again ESA loss is evident from 340-410 nm, as well as loss of SE (480-600 nm) and of GSB (410-460 nm). Note the large differences in shape between the first and second DADS which indicate the presence of other states. The third DADS (blue) decays in 34 ps. Its shape is similar to the second DADS (compare the normalized DADS in Fig. 30E). The fourth DADS (green, lifetime ~1 ns) shows loss of GSB
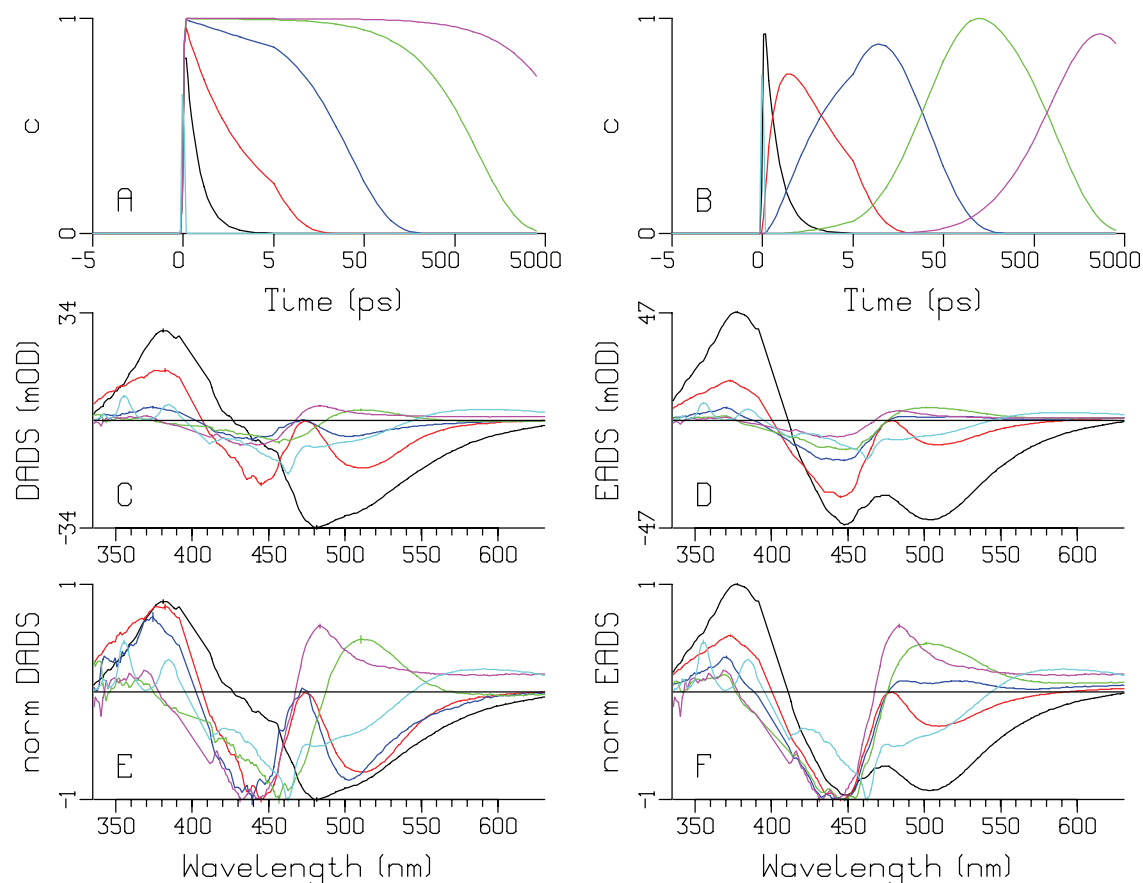
Fig.30. Global analysis of PYP magic angle data using a parallel (left) or sequential kinetic scheme (right). (A,B) show $c(t)$ of the five components. Key: lifetime 0.8 ps (black), 3.3 ps (red), 34 ps (blue), 1 ns (green), long lived (magenta). Coherent artefact with IRF time profile is in cyan.

(C,D) depict estimated Decay Associated Difference Spectra (DADS) and Evolution Associated Difference Spectra (EADS). (E,F) show normalized spectra. Vertical bars indicate approximate standard errors.

as well as loss of product state absorption (490-570 nm) which will later be ascribed to intermediate I_0. The fifth and final DADS (magenta) is long lived and represents the intermediate I_1. The coherent artefact is described by the cyan spectrum which is associated to the IRF. Depicted here is its contribution at the IRF maximum. This spectrum steals some amplitude from fast processes, e.g. early SE (470-540 nm) and solvated electron production (above 550 nm, explained below). Unless noted otherwise, the relative errors in the estimated lifetimes are about 10%. The estimated errors in the spectral parameters (see Eq. 106) are also small, as evidenced by the error bars (vertical lines at extrema).

The first EADS (black in Fig. 30D) equals the sum of all DADS. Apart from the coherent artefact contribution it represents the difference spectrum at time zero. It is characterized by a large GSB (420-460 nm). Note that the subsequent EADS show a gradual decrease of this GSB. The final EADS (magenta) equals the final DADS. If we assume that the evolution from the fourth to the fifth EADS corresponds to the transition from photocycle intermediates I_0 to I_1, then the fourth and fifth EADS represent the SADS of I_0 and I_1, respectively. The apparent loss of GSB in this transition could be due to non-unity quantum yield of this transition. Alternatively, the absorption of I_1 largely overlaps with the GSB, thus resulting in a smaller SADS. Both alternatives will be tested below. The second and third EADS are most difficult to interpret because at least three states contribute: excited state intermediates, I_0, and also a ground state intermediate (vide infra).

4.3 Target analysis of anisotropy data: SAS

To disentangle these complicated dynamics more information is needed. Measurements at parallel and perpendicular polarization angles add information on anisotropy differences of the intermediates relative to the anisotropy of the GSB. Representative traces and spectra are shown in Fig. 31 and Fig. 32, respectively. Disregarding the dispersion, the 0.5, 300 and 2500 ps spectra resemble the first, fourth and fifth EADS (Fig. 30F), respectively. The 4 ps spectrum resembles a mixture of the second and third EADS. For illustrative purposes only, the raw anisotropies $r(t, \lambda)$ (calculated according to Eq. 63) are depicted in the insets of Fig. 31 and Fig. 32. Note that the $r(t, \lambda)$ are noisy, as explained in §2.7.5.1, and not suitable for further analysis.

A necessary prerequisite for the simultaneous analysis of the anisotropy data is a target model (§2.7.5.2). The target model consists of four parts: a compartmental scheme (§2.4.3) for the MA concentrations, a photocycle model (§2.7.4) which contains the Species Associated Spectra (SAS) including the GSB (see Eq. 58), an anisotropy model (§2.7.5.2) and spectral assumptions (§2.7.2) on the SAS. The compartmental scheme (Fig. 33A) consists of three Excited State Intermediates ESI1, 2 and 3 each decaying monoexponentially (symbolized by the triple arrows), thus describing the three time scales of SE and ESA decay. This heterogeneity is not unusual for a biological system (§2.2.1). An alternative scheme with three interconverting states is discussed in [8]. An ESI can either decay into the photocycle intermediate I_0 or into the unrelaxed Ground State Intermediate (GSI). Subsequently, GSI relaxes to the ground state. As described above, the photocycle intermediate I_0 relaxes to the
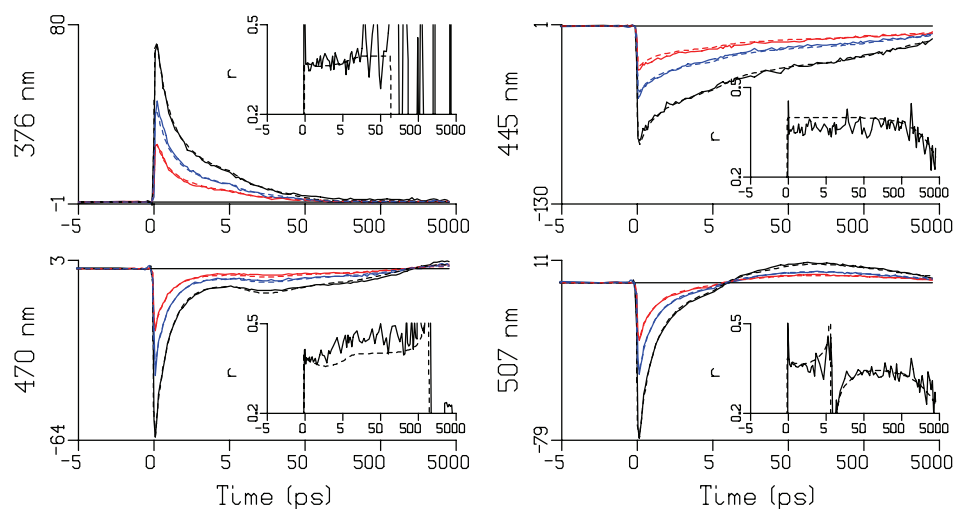
Fig.31. Selected difference absorption traces (in mOD) of PYP after 400 nm excitation, measured at magic angle (blue), parallel (black), and perpendicular angle (red). Probe wavelength indicated along the ordinate. Insets depict raw anisotropy. Dashed lines indicate fit. Note that the time axis is linear from -5 to +5 ps relative to the maximum of the IRF, and logarithmic thereafter.
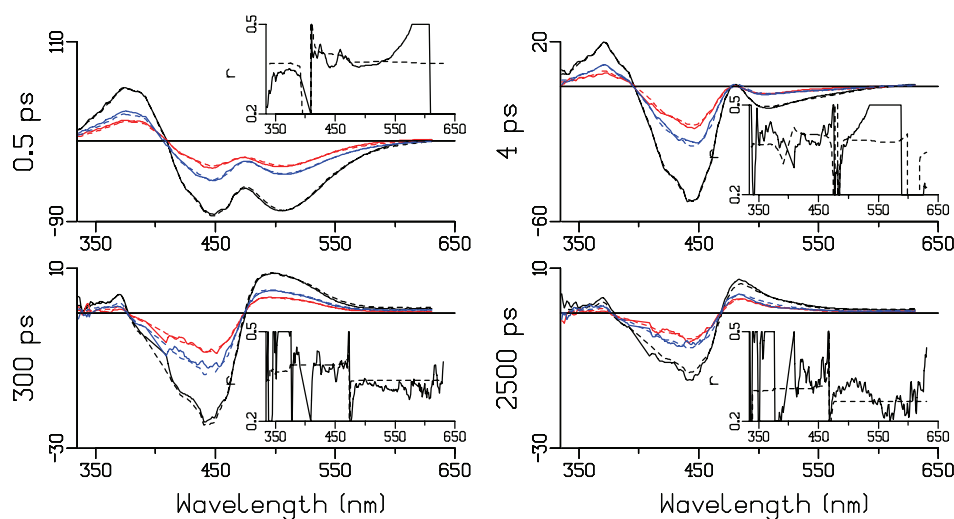

Fig.32. Selected difference absorption spectra (in mOD) of PYP after 400 nm excitation, measured at magic angle (blue), parallel (black), and perpendicular angle (red). Probe time indicated along the ordinate. The two early spectra are affected by dispersion. Insets depict raw anisotropy. Dashed lines indicate fit.
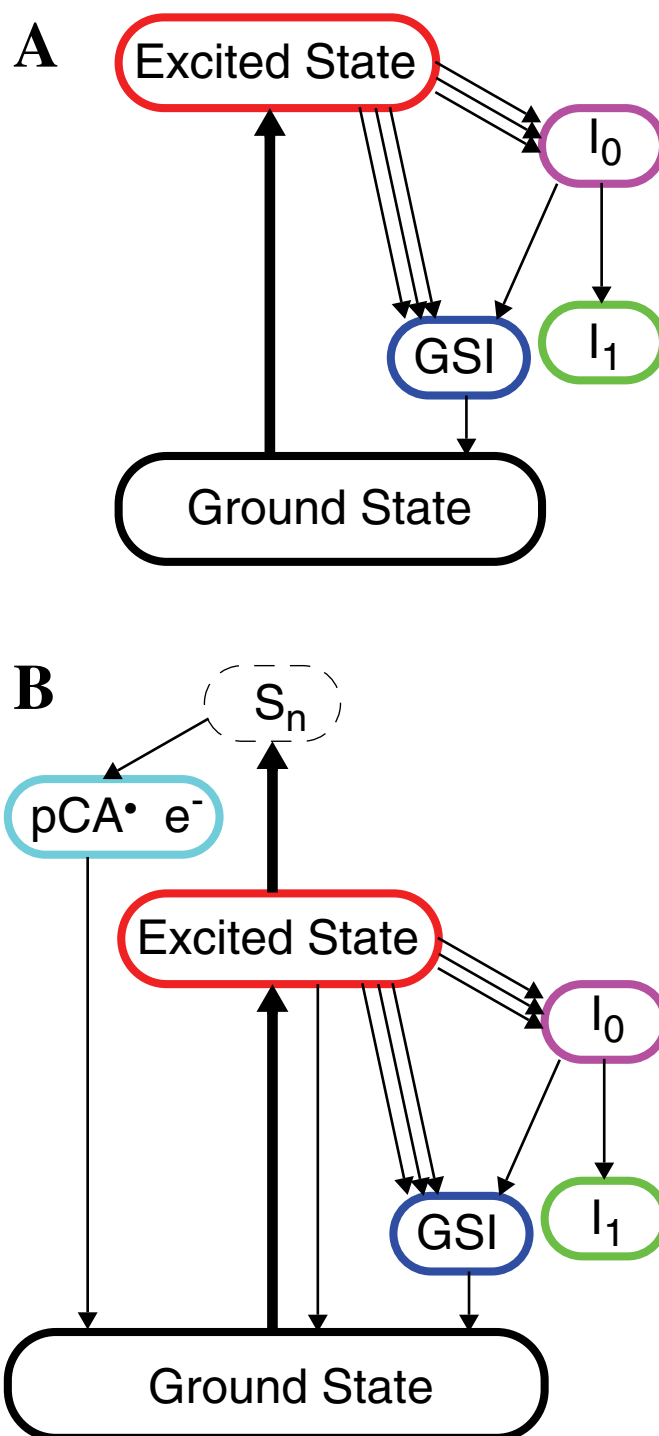
Fig.33. Compartmental schemes for the target analysis of polarized transient difference absorption of PYP after 400 nm excitation. Vertical upward arrows indicate excitation. Excited state intermediates ESI1, 2 and 3 decay into I_0 or GSI. Scheme B differs from A by an additional decay from ESI1 to the ground state, and by a parallel photo ionization. long lived I_1. Of paramount importance for the fit of these anisotropy data with coexisting

intermediates are the spectral assumptions which allow to describe parts of the data with a subset of the intermediates. The spectral assumptions used are: (i) GSB zero above 475 nm, (ii) SAS of ESI1, 2 and 3 identical, (iii) SAS of GSI nonzero from 427-500 nm, (iv) SAS of I_0 zero below 475 nm, (v) SAS of I_1 zero below 440 nm. Although the anisotropy differences are small, they do allow resolution of the GSB spectrum (to which an anisotropy of 0.4 is assigned), and of the SAS of the intermediates (which possess somewhat lower anisotropies). The estimated SAS are shown in Fig. 34B. Instead of the five lifetimes from the global

Fig.34. (A) Normalized steady-state absorption (black) and emission (red) spectra of PYP, spectra of the pump pulse (blue) and of the dump pulse (green, applied in Fig. 35). (B,C,D) Species Associated Spectra (SAS) estimated from the PYP anisotropy data using the compartmental schemes of Fig. 33. Panel B, corresponding to Fig. 33A, differs from panel C (which corresponds to Fig. 33B) by the absence of a decay from ESI1 to the ground state. Panel D differs from panel C in the SAS of I_1, it corresponds to Fig. 33B without the decay from I_0 to GSI. Key: GSB (black), ESI1,2,3 (red), GSI (blue), I_0 (magenta), I_1 (green), pCA radical (cyan). Coherent artefact has been omitted for clarity. Vertical bars indicate approximate standard errors.
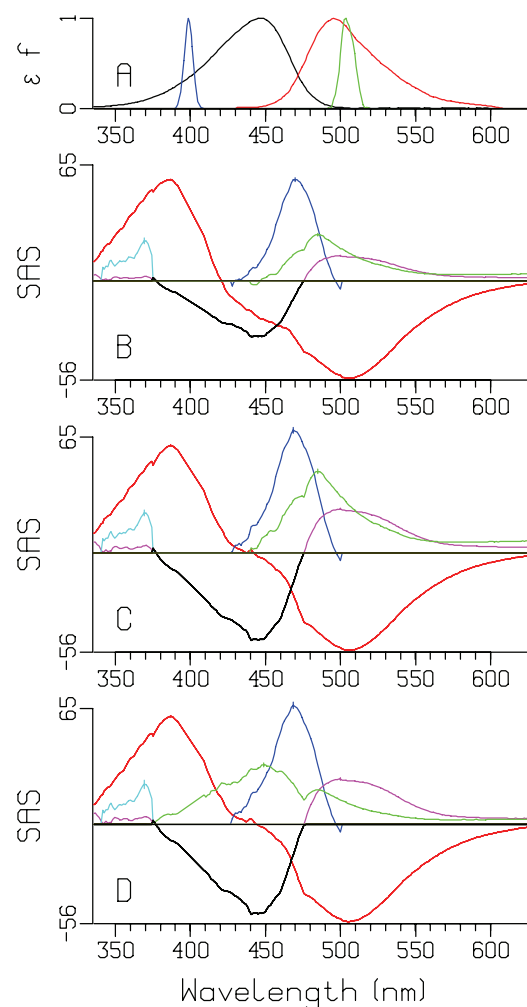
analysis now seven lifetimes are estimated from the target analysis. Next to the long lived GSB an extra lifetime of 2.3 ps is estimated for the GSI. The SAS are all smooth, except for that of the ESI (red) which still appears to contain some bleach contribution from 420-460 nm. To improve the shape of the SAS of the ESI an additional decay path from the ESI1 state directly to the ground state was added. Since this rate cannot be estimated from the fit, it was adjusted iteratively in order to produce a satisfactory shape. Actually a contribution of this

path of 50% was used. Note that in Fig. 34C, the negative part of the ESI SAS (red) resembles the mirror image of the ground state SAS (black), analogous to the steady state emission and absorption spectra in Fig. 34A. This shows that a rate parameter that does not influence the quality of fit of the data can be determined indirectly from the resulting SAS. In retrospect, we can now interpret the shape difference of the first and second DADS between 460 and 490 nm (Fig. 30E) and ascribe it to a rise of the GSI which accompanies the ESI1 decay. Furthermore, a small long lived product state absorption is present below 375 nm. This absorption was not influenced by the dump experiments (described in §4.4) and could therefore be attributed to resonantly enhanced photo ionization of the chromophore (p-coumaric acid, pCA) resulting in a radical and a solvated electron. It is possible to isolate this contribution from the GSB using spectral constraints. When the GSB is assumed to be zero below 375 nm, the long lived signal below 375 nm can be attributed to the radical, resulting in the cyan SAS. The absorption of the solvated electron cannot be resolved, it is clearly visible in the long lived DADS (magenta) above 570 nm in Fig. 30E. The anisotropies of ESI2, I_0 and I_1 were estimated to be, respectively, 0.33, 0.35 and 0.35, significantly lower than the 0.4 anisotropy of the GSB. The other anisotropies could not be precisely estimated.

As discussed in §4.2, an alternative interpretation of the loss of bleach in the transition from I_0 to I_1, is that the absorption of I_1 largely overlaps with the GSB. Thus the spectral assumption (v) SAS of I_1 zero below 440 nm, can be replaced by a different assumption, namely that the quantum yield of the transition from I_0 to I_1 is unity, instead of 41% as estimated using the spectral assumption (v). This results in the I_1 SAS depicted in Fig. 34D. Although this modified kinetic scheme cannot be excluded, the spectral shape of the I_1 SAS is considered unlikely. This question can be resolved by a precise measurement of the wavelength-dependence of the anisotropy at a few ns (before it disappears by tumbling, Eq. 68, and 445 nm inset in Fig. 31), provided there is no photoionization present.

4.4 Target analysis of multipulse data: SADS

The compartmental scheme of Fig. 33B was put to the test with a multipulse experiment (§2.7.6). Administering a second laser pulse (green in Fig. 34A) resonant with the SE transfers population from excited to ground state intermediates. Thus different concentrations of intermediates are created, providing new clues to resolve their properties. From the representative traces in Fig. 35 it becomes evident that the dump pulse depletes the ESI population resulting in decreased SE (530 nm), and decreased ESA (378 nm), concomitant
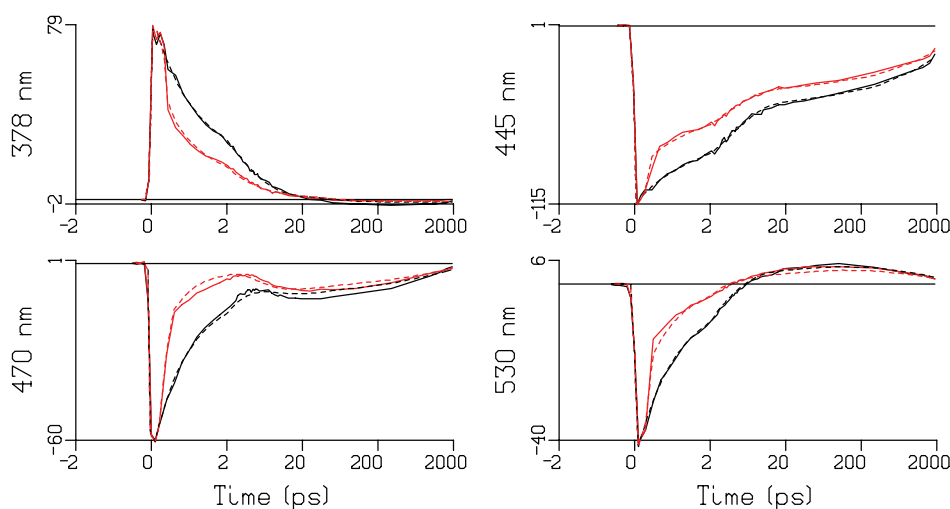
Fig.35. Selected difference absorption traces (in mOD) of PYP after 400 nm
excitation, (black) and after a second dump pulse (red) centered around 505 nm
(green in Fig. 34A) and administered ~400 fs later. Probe wavelength indicated
along the ordinate. Dashed lines indicate fit. Note that the time axis is linear from -2
to +2 ps relative to the maximum of the IRF, and logarithmic thereafter.

with transient increased absorption at 470 nm. The last is attributable to the GSI. Also note the
relatively small amount of GSB recovery (445 nm) upon application of the dump pulse, which
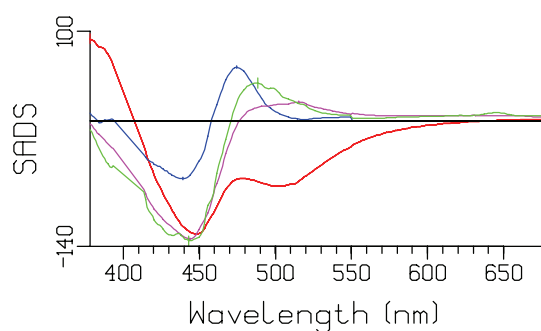also points to the presence of a GSI. Overall, the estimated SADS in Fig. 36 are consistent

Fig.36. Species Associated Difference Spectra (SADS) estimated from the PYP
data with and without dump pulse (cf. Fig. 35) using the compartmental scheme
of Fig. 33B. Key: SADS of ESI1,2,3 (red), GSI (blue), I_0 (magenta), I_1 (green).
Coherent artefact has been omitted for clarity. Vertical bars indicate approximate
standard errors.

with the SAS of Fig. 34C. Unfortunately, the contribution of the radical and the solvated electron could not be resolved with these data, which were measured above 378 nm. The absorption by the solvated electron is clearly visible in the SADS of I_0 and I_1 above 570 nm. Both the anisotropy and the multipulse experiments are satisfactorily described by the kinetic scheme of Fig. 33B in combination with the anisotropy and IRF parameters, the SAS of Fig. 34C and the SADS of Fig. 36.

4.5 Hierarchies and the role of the PSE in this case study

When the elements of the target analysis are combined with the flow chart of Fig. 2 we get Fig. 37. The target analysis was based on a hierarchy of models, which is summarized in

Table 3 Hierarchical modelling of time-resolved polarized difference absorption

level of modelling	parametric description of	
linking of experiments	relative scaling, linkage schemes	§2.7.5.2
contribution of component l $c_l(t)\varepsilon_l(\lambda)$	spectrum of component l, $\varepsilon_l(\lambda)$	§2.7.2
convolution $a_l(t)c_l^\delta(t) \otimes i(t)$	Instrument Response Function $i(t)$	§2.4.2
depolarization	anisotropy $a_l(t)$	§2.7.5.2
MA concentration $c_l^\delta(t)$ with δ-input	compartmental scheme with microscopic rates	§2.4.3

Table 3 and Fig. 38. Some numbers are collated in Table 4. The total parameter vector Θ can

Superposition model:
for each state

Instrument
Response
Function

concentration

depolarization
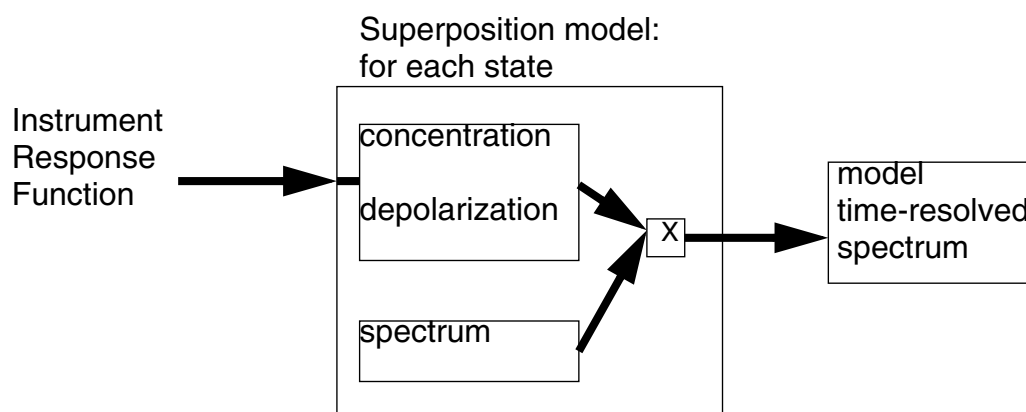
X

model
time-resolved
spectrum

spectrum

Fig.38. Flow chart of hierarchical modelling of time-resolved spectrum.

be split into $\Theta = (\Theta_K, \Theta_L, \Theta_I, \ldots)$ where each subvector $\Theta_.$ expresses a different aspect of the model. Spectral assumptions, that $\varepsilon_l(\lambda) = 0$ for certain components l in certain λ-
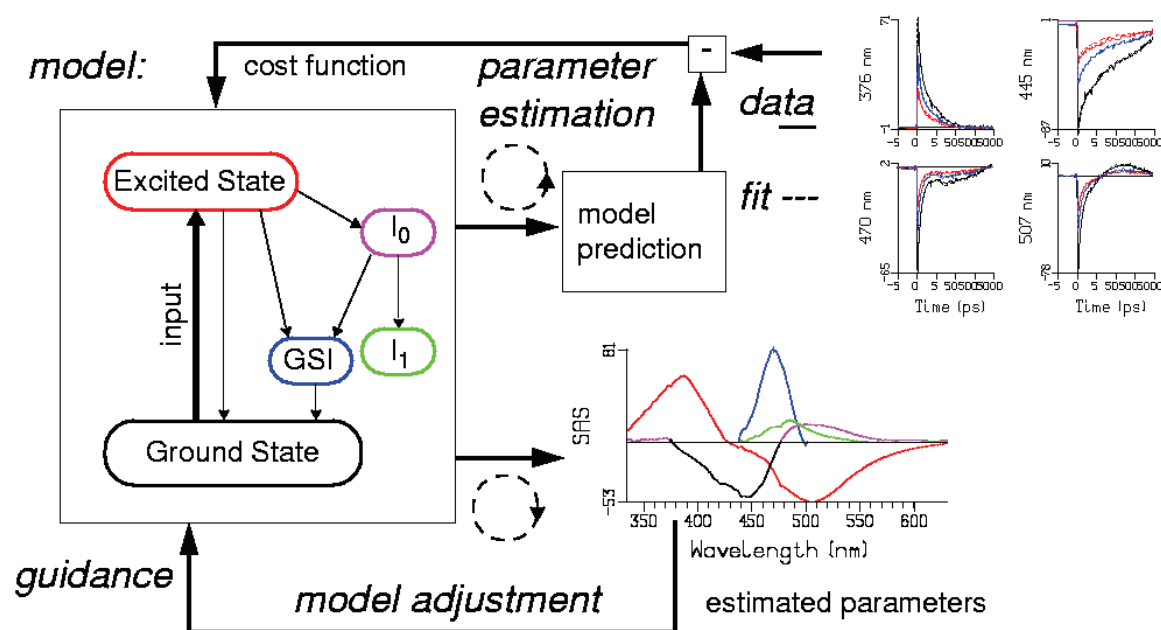
Fig.37. Possible graphical user interface of prototype PSE showing the key concepts and some pictures from the PYP case study with three-way data. Data (solid lines) at four representative wavelengths (indicated by the vertical label) are depicted in the upper right hand panel. Note that the time axis is linear from -5 to +5 picoseconds (ps) relative to the instrument response function (IRF) maximum, and logarithmic thereafter. Dashed lines represent the fit, calculated from the model prediction after traversing the parameter estimation loop. The compartmental scheme (left hand panel) contains five different states. The model scheme was adjusted in the model adjustment loop in order to arrive at a satisfactory description, in particular realistic shapes of the estimated spectral parameters, the SAS (bottom right hand panel). The colors of the SAS correspond to those of the boxes in the compartmental scheme. The units of the SAS and of the data are milliOD. The vertical lines indicate the estimated error. Further explanation in text.

ranges (§2.7.2 and §2.7.5.2), or that some spectra are equal up to a scaling parameter Θ_ε, are essential to estimate overlapping spectra and branching ratios. Adjusting and testing such assumptions (Fig. 34) necessitates a flexible PSE user interface.

5 Conclusion

State-of-the-art experiments and sophisticated analysis methodology mutually benefit. Global and target analysis have become indispensable tools in the investigation of complex systems

Table 4. Numbers and parameters with PYP polarized difference absorption data

	range	#	#parameters	
wavelength λ	334 – 631 nm	240		
time t	-1ps – 5 ns	100		
polarization	Magic Angle, //, $\perp$	3		
data points		72000		
compartments		7		
spectra		7+1+1		
microscopic rate constants		18	18-6=12	Θ_K
instrument response location, width, dispersion			4	Θ_I
anisotropy			8-1=7	Θ_ϕ
scaling parameter for linking experiments			2-1=1	Θ_L
spectral parameters (VARPRO)	$9n_\lambda$ – assumptions		984	$\epsilon_l(\lambda)$
spectral assumptions		28	1-1=0	Θ_ϵ
Thus 24 truly nonlinear parameters and 984 conditionally linear parameters must be estimated.				

with time-resolved spectroscopy. In particular multipulse experiments enable testing of complicated kinetic schemes. Spectrotemporal modelling offers a solution to the inverse problem, and allows precise estimation of the kinetic and spectral parameters that describe the complex system dynamics.

6 References

[1] A.R. Holzwarth, Time-resolved fluorescence spectroscopy. *Methods Enzymol.* 246 (1995) 334-362.
[2] H. van Amerongen, R. van Grondelle. Transient absorption spectroscopy in study of processes and dynamics in biology. *Methods Enzymol.* 246 (1995) 201-226.
[3] R.H. Lozier, R.A. Bogomolni, W. Stoeckenius, Bacteriorhodopsin: a light-driven proton pump in *Halobacterium halobium. Biophys. J. 15 (*1975*)* 955-962.
[4] T.G. Ebrey, Light energy transduction in bacteriorhodopsin. In: M. B. Jackson (Ed.) *Thermodynamics of Membrane Receptors and Channels*; CRC: Boca Raton, FL, 1993; Chapter 10.

[5]   T.E. Meyer, E. Yakali, M.A. Cusanovich, G. Tollin, Properties of a water-soluble, yellow protein isolated from a halophilic phototrophic bacterium that has photochemical activity analogous to sensory rhodopsin. *Biochemistry 26* (1987) 418-423.

[6]   K.J. Hellingwerf, J. Hendriks, T. Gensch, Photoactive Yellow Protein, a new type of photoreceptor protein: Will this "yellow lab" bring us where we want to go? *J. Phys. Chem. A* 107 (2003) 1082-1094

[7]   D.S. Larsen, M. Vengris, I.H.M. van Stokkum, M.A. van der Horst, F.L. de Weerd, K.J. Hellingwerf, R. van Grondelle, Photoisomerization and photoionization of the photoactive yellow protein chromophore in solution, *Biophys. J. 86* (2004) 2538-2550.

[8]   D.S. Larsen, I.H.M. van Stokkum, M. Vengris, M.A. van der Horst, F.L. de Weerd, K.J. Hellingwerf, R. van Grondelle, Incoherent manipulation of the photoactive yellow protein photocycle with dispersed pump-dump-probe spectroscopy, *Biophys. J. 87* (2004) 1858-1872.

[9]   Houstis EN, Rice JR, Gallopoulos E, Brambley R (eds.). *Enabling Technologies for Computational Science: Frameworks, Middleware and Environments*. Kluwer Academic Publishers, 2000.

[10]   R. Bro, J.J. Workman, P.R. Mobley, B.R. Kowalski, Review of chemometrics applied to spectroscopy: 1985-95. 3. Multi-way analysis. *Appl. Spectrosc. Rev. 32* (1997) 237-261.

[11]   S. Bijlsma, A.K. Smilde, Estimating reaction rate constants from a two-step reaction: a comparison between two-way and three-way methods. *J. Chemom. 14* (2000) 541-560.

[12]   J.M. Beechem, M. Ameloot, L. Brand, Global and target analysis of complex decay phenomena. *Anal. Instrum. 14* (1985) 379-402.

[13]   J.M. Beechem, A second generation global analysis program for the recovery of complex inhomogeneous fluorescence decay kinetics. *Chem. Phys. Lipids 50* (1989) 237-251.

[14]   J.M. Beechem, Global analysis program of biochemical and biophysical data. *Methods Enzymol. 210* (1992) 37-54.

[15]   M. Ameloot, N. Boens, R. Andriesen, V. van den Bergh, F.C. de Schryver, Compartmental analysis of fluorescence decay surfaces of excited state processes. *Methods Enzymol. 210* (1992) 37-54.

[16]   A.R. Holzwarth, Data analysis of time-resolved measurements, in: J. Amesz, A.J. Hoff (Eds.), Biophysical Techniques in Photosynthesis, Kluwer, Dordrecht, the Netherlands, 1996, pp. 75-92.

[17]   A.K. Dioumaev, Evaluation of intrinsic chemical kinetics and transient product spectra from time-resolved spectroscopic data. *Biophys. Chem. 67* (1997) 1-25.

[18]   R. Bonneau, J. Wirz, A.D. Zuberbühler, Methods for the analysis of transient absorbance data. *Pure & Appl. Chem.* 69 (1997) 979-992.

[19]   J.F. Nagle, L.A. Parodi, R.H. Lozier, Procedure for testing kinetic models of the photocycle of bacteriorhodopsin. *Biophys. J. 38* (1982) 161-174.

[20]   J.F. Nagle, Solving complex photocycle kinetics. Theory and direct method. *Biophys. J. 59* (1991) 476-487.

[21]   R.H. Lozier, A. Xie, J. Hofrichter, G.M. Clore, Reversible steps in the bacteriorhodopsin photocycle. *Proc. Nat. Acad. Sci. USA* 89 (1992) 3610-3614.

[22]   R.W. Hendler, R.I. Shrager, S. Bose, Theory and procedures for finding a correct kinetic model for the bacteriorhodopsin photocycle. *J. Phys. Chem. B. 105 (*2001) 3319-3328.

[23]   R.I. Shrager, R.W. Hendler, Critical evaluation of kinetic models for bacteriorhodopsin photocycles. *J. Phys. Chem. B. 107 (*2003) 1708-1713.

[24]   I.H.M. van Stokkum, R.H. Lozier, Target analysis of the bacteriorhodopsin photocycle

using a spectrotemporal model. *J. Phys. Chem. B* 106 *(*2002) 3477-3485.

[25] G.A.F. Seber, C.J. Wild, *Nonlinear regression*. Wiley, New York, 1989.

[26] D.M. Bates, D.G. Watts, *Nonlinear regression and its applications*. Wiley, New York, 1988.

[27] I.H.M. van Stokkum, D. S. Larsen, R. van Grondelle, Global and target analysis of time resolved spectra, *Biochim. Biophys. Acta* 1657 *(*2004) 82-104, and 1658, 262 (Erratum)

[28] B. Gobets, *The life and times of Photosystem I. Excitation energy transfer and trapping unravelled.* Vrije Universiteit Amsterdam, 15 October 2002, PhD Thesis.

[29] T. Ito, M. Hiramatsu, M. Hosoda, Y. Tsuchiya, Picosecond time-resolved absorption spectrometer using a streak camera. *Rev. Sci. Instrum. 62* (1991) 1415-1419.

[30] M. van den Zegel, N. Boens, D. Daems, F.C. de Schryver, Possibilities and limitations of the time-correlated single photon counting technique: a comparative study of correction methods for the wavelength dependence of the instrument response function. *Chem. Phys. 101* (1986) 311-335.

[31] N. Boens, N. Tamai, I., Yamazaki, T. Yamazaki, Picosecond single photon timing measurements with a proximity type microchannel plate photomultiplier and global analysis with reference convolution. *Photochem Photobiol. 52* (1990) 911-917.

[32] M. Zuker, A.G. Szabo, L. Bramall, D.T. Krajcarski, B. Selinger, Delta function convolution method (DFCM) for fluorescence decay experiments. *Rev. Sci. Instrum. 56* (1985) 14-22.

[33] S.A. Kovalenko, A.L. Dobryakov, J. Ruthmann, N.P. Ernsting, Femtosecond spectroscopy of condensed phases with chirped supercontinuum probing. *Phys. Rev .A 59* (1999) 2369-2384.

[34] I.H.M. van Stokkum, A.M. Brouwer, H.J. van Ramesdonk, T. Scherer, Multiresponse parameter estimation and compartmental analysis of time resolved fluorescence spectra. Application to conformational dynamics of charge-separated species in solution. *Proc. Kon. Ned. Akad. v. Wetensch*. 96 (1993) 43-68.

[35] S. Weisberg, Applied linear regression. 2nd. ed. Wiley, New York, 1985.

[36] K.-H. Müller, Th. Plesser, Variance reduction by simultaneous multi-exponential analysis of data sets from different experiments. *Eur. Biophys. J. 19* (1991) 231-240.

[37] K.J. Laidler, *Chemical kinetics*, 3rd ed. Harper & Row: New York, 1987.

[38] G.R. Fleming, T. Joo, M. Cho, Femtosecond Chemical Dynamics in Condensed Phases. *Adv. Chem. Phys. 101* (1997) 141-183.

[39] T.A. Roelofs, C.H. Lee, A.R. Holzwarth, Global target analysis of picosecond chlorophyll fluorescence kinetics from pea chloroplasts. A new approach to the characterization of the primary processes in photosystem II α- and β-units. *Biophys. J. 61* (1992) 1147-1163.

[40] S.W. Provencher, An eigenfunction expansion method for the analysis of exponential decay curves. *J. Chem. Phys. 64*, (1976) 2772-2777.

[41] S.W. Provencher, A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comp. Phys. Comm.* 27 (1982) 213-227.

[42] J.C. Brochon, Maximum-entropy method of data-analysis in time-resolved spectroscopy. *Methods Enzymol.* 240 (1994) 262-311.

[43] L.M.P. Beekman, I.H.M. van Stokkum, R. Monshouwer, A.J. Rijnders, P. McGlynn, R.W. Visschers, M.R. Jones, R. van Grondelle, Primary electron transfer in membrane-bound reaction centers with mutations at the M210 position. *J. Phys. Chem.100* (1996) 7256-7268.

[44] C.R. Cantor, P.R. Schimmel. *Biophysical Chemistry. Part II: Techniques for the study of biological structure and function.* Freeman, New York, 1980.

[45] J.C. Hamilton, P.J. Gemperline, Mixture analysis using factor analysis. II: Self modeling curve resolution. *J. Chemometrics 4* (1990) 1-13.

[46] W. Windig, Self-modeling mixture analysis of spectral data with continuous concentration profiles. *Chemometrics Intell. Lab. Syst. 16* (1992) 1-16.

[47] L. Zimányi, A. Kulcsár, J.K. Lanyi, D.F. Sears, J. Saltiel, Singular value decomposition with self-modeling applied to determine   bacteriorhodopsin intermediate spectra: Analysis of   simulated data. *Proc. Natl. Acad. Sci. USA.* 96 (1999) 4408-4413.

[48] L. Zimányi, A. Kulcsár, J.K. Lanyi, D.F. Sears, J. Saltiel, Intermediate spectra and photocycle kinetics of the Asp96 →Asn mutant   bacteriorhodopsin determined by singular value decomposition with self-modeling. *Proc. Natl. Acad. Sci. USA. 96* (1999) 4414-4419.

[49] A. Kulcsár, J. Saltiel, L. Zimányi, Dissecting the photocycle of the bacteriorhodopsin E204Q mutant from kinetic multichannel difference spectra. Extension of the method of singular value decomposition with self-modeling to five components. *J. Am. Chem. Soc. 123 (*2001*) 3332-3340.

[50] S. Solar, W. Solar, N. Getoff, A pulse radiolysis-computer simulation method for resolving of complex kinetics and spectra. *Radiat. Phys. Chem. 21* (1983) 129-138.

[51] K.J. Visscher, M.L. Hom, H. Loman, H.J.W. Spoelder, J.B. Verberne, Spectral and kinetic properties of intermediates induced by reaction of hydrated electrons with adenine, adenosine, adenylic acid and polyadenylic acid; a multi component analysis. *Radiat. Phys. Chem. 32* (1988) 465-473.

[52] F.J. Knorr, J.M. Harris, Resolution of multicomponent fluorescence spectra by an emission wavelength-decay time data matrix. *Anal. Chem. 53* (1981) 272-276.

[53] J.M. Beechem, M. Ameloot, L. Brand, Global analysis of fluorescence decay surfaces: Excited-state reactions. *Chem. Phys. Lett. 120* (1985) 466-472.

[54] J.R. Knutson, J.M. Beechem, L. Brand, Simultaneous analysis of multiple fluorescence decay curves: a global approach. *Chem. Phys. Lett. 102* (1983*) 501-507.

[55] J.E. Löfroth, Time-resolved emission spectra, decay-associated spectra, and species-associated spectra. *J. Phys. Chem. 90* (1986) 1160-1168.

[56] V.N. Petushkov, I.H.M. van Stokkum, B. Gobets, F. van Mourik, J. Lee, R. van Grondelle, A.J.W.G. Visser, Ultrafast fluorescence relaxation spectroscopy of 6,7-dimethyl-(8-ribityl)-lumazine and riboflavin, free and bound to antenna proteins from bioluminescent bacteria. *J. Phys. Chem. B 107* (2003) 10934-10939.

[57] A. Fellinger. Deconvolution of overlapping skewed peaks. *Anal. Chem. 66* (1994) 3066-3072.

[58] S.J. Kok, E. M. Kristenson, C. Gooijer, N. H. Velthorst, U. A. Th. Brinkman, Wavelength-resolved laser-induced fluorescence detection in capillary electrophoresis: naphthalenesulphonates in river water, *J. Chromatogr. A 771* (1997) 331-341.

[59] M. Bouchy, J.Y. Jezequel, J.C. Andre, J. Bordet, Problems arising from the quantization of the data in the mathematical treatment of single photon counting experiments. In: Deconvolution and reconvolution of analytical signals. Application to fluorescence spectroscopy, E.N.S.I.C.-I.N.P.L., Nancy, France, 1982, pp. 393-409.

[60] P.J.M. van Kan, I.H.M. van Stokkum, S. Styring, Time-resolved EPR spectroscopy of the triplet state of the primary electron donor P680 in PS II. In: P. Mathis (ed.), Photosynthesis: from Light to Biosphere, Vol. I, 543-546. Dordrecht, Kluwer.

[61] I.H.M. van Stokkum, H.E. Bal, A Problem Solving Environment for interactive modelling of multiway data, *Concurrency Computat: Pract. Exper.*, (in press)

[62] Van Stokkum, IHM. Parameter Precision in Global Analysis of Time Resolved Spectra. IMTC Proceedings 1996, Vol. I, IEEE, Piscataway NJ, p.168-173.

[63] D.H. Anderson, *Compartmental modeling and tracer kinetics*. Springer, Berlin, 1983.

[64] K. Godfrey, *Compartmental models and their application*. Academic Press, London, 1983.

[65] A. Arcioni C. Zannoni, Intensity deconvolution in fluorescence depolarization studies of liquids, liquid crystals and membranes, *Chemical Physics 88* (1984) 113-128.

[66] I.H.M. van Stokkum, L.M.P. Beekman, M.R. Jones, M.E. van Brederode, R. van Grondelle, Primary electron transfer kinetics in membrane-bound *Rhodobacter sphaeroides* reaction centers: A global and target analysis. *Biochemistry 36* (1997) 11360-11368.

[67] S. Vajda, H. Rabitz, Identifiability and distinguishability of first-order reaction systems. *J. Phys. Chem. 92* (1988) 701-707.

[68] J. Eisenfeld, C.C. Ford, A systems-theory approach to the analysis of multiexponential fluorescence decay. *Biophys. J. 26* (1979) 73-84.

[69] I.H.M. van Stokkum, T. Scherer, A.M. Brouwer, J.W. Verhoeven, Conformational dynamics of flexibly and semirigidly bridged electron donor-acceptor systems as revealed by spectrotemporal parameterization of fluorescence. *J. Phys. Chem. 98 (*1994) 852-866.

[70] X.Y. Lauteslager, I.H.M. van Stokkum, H.J. van Ramesdonk, A.M. Brouwer, J.W. Verhoeven, Conformational Dynamics of Semiflexibly Bridged Donor-Acceptor Systems Studied with a Streak Camera and Spectrotemporal Parametrization of Fluorescence. *J. Phys. Chem. 103* (1999) 653-659.

[71] M.G. Müller, K. Griebenow, A.R. Holzwarth, Primary processes in isolated bacterial reaction centers from Rhodobacter sphaeroides studied by picosecond fluorescence kinetics. *Chem. Phys. Lett. 199* (1992) 465-469.

[72] M.G. Müller, J. Niklas J, W. Lubitz, A.R. Holzwarth, Ultrafast transient absorption studies on Photosystem I reaction centers from Chlamydomonas reinhardtii. 1. A new interpretation of the energy trapping and early electron transfer steps in Photosystem I *Biophys. J.* 85 (2003) 3899-3922.

[73] R. Croce, M.G. Müller, R. Bassi, A.R. Holzwarth, Carotenoid-to-chlorophyll energy transfer in recombinant major light-harvesting complex (LHCII) of higher plants. I. Femtosecond transient absorption measurements. *Biophys. J.* 80 (2001) 901-915.

[74] I. Chizhov, D.S. Chernavskii, M. Engelhard, K.-H. Mueller, B.V. Zubov, B. Hess, Spectrally silent transitions in the bacteriorhodopsin photocycle. *Biophys. J. 71* (1996) 2329-2345.

[75] M.E. van Brederode, W.D. Hoff, I.H.M. van Stokkum, M.L. Groot, K.J. Hellingwerf, Protein folding thermodynamics applied to the photocycle of photoactive yellow protein. *Biophys. J. 71* (1996) 365-380.

[76] I.H.M. van Stokkum, H. Linsdell, J.M. Hadden, P.I. Haris, D. Chapman, M. Bloemendal, Temperature-induced changes in protein structures studied by Fourier transform infrared spectroscopy and global analysis. *Biochemistry 34* (1995) 10508-10518.

[77] R. Kubo, A stochastic theory of lineshape. *Adv. Chem. Phys. 15* (1969) 101-127.

[78] R. A. Marcus, Relation between charge transfer absorption and fluorescence spectra and the inverted region. *J. Phys. Chem. 93* (1989) 3078-3086

[79] J.R. Lakowicz, *Principles of fluorescence spectroscopy*. Plenum Press, New York, 1983, Chapter 2.

[80] R.D.B. Fraser, E. Suzuki, Resolution of overlapping bands: functions for simulating band shapes. *Anal. Chem.* 41 (1969) 37-39.

[81] J.M. Sevilla, M. Dominguez, F. Garcia-Blanco, M. Blazquez, Resolution of absorption spectra. *Computers Chem.* 13 (1989) 197-200.

[82] T. Pullerits, R. Monshouwer, F. van Mourik, R. van Grondelle, Temperature dependence of electron-vibronic spectra of photosynthetic systems. Computer simulations and comparison with experiment. *Chem. Phys. 194* (1995) 395-407.

[83] G.H. Golub, C.F. van Loan, Matrix computations. 2nd ed., Johns Hopkins University Press, Baltimore, 1990, Chapter 8.

[84] R.I. Shrager, R.W. Hendler, Titration of individual components in a mixture with resolution of difference spectra, pKs, and redox transitions, *Anal. Chem.* 54 (1982) 1147-1152.

[85] R.I. Shrager, Chemical transitions measured by spectra and resolved using singular value decomposition, *Chemometrics Intell. Lab. Syst. 1* (1986) 59-70.

[86] E.R. Henry, J. Hofrichter, Singular Value Decomposition: Application to Analysis of Experimental Data. *Methods Enzymol. 210* (1992) 129-192.

[87] R.W. Hendler, R.I. Shrager, Deconvolutions based on singular-value decomposition and the pseudoinverse - a guide for beginners. *J. Biochem. Biophys. Meth. 28* (1994) 1-33.

[88] E.R. Henry, The use of matrix methods in the modeling of spectroscopic data sets. *Biophys. J.* 72 (1997) 652-673.

[89] H. Satzger, W. Zinth, Visualization of transient absorption dynamics - towards a qualitative view of complex reaction kinetics, *Chem. Phys. 295* (2003) 287-295.

[90] J.F. Nagle, Photocycle kinetics: analysis of Raman data from bacteriorhodopsin. *Photochem. Photobiol. 54* (1991) 897-903.

[91] I. Szundi, J.W. Lewis, D.S. Kliger, Deriving reaction mechanisms from kinetic spectroscopy. Application to late rhodopsin intermediates. *Biophys. J.* 73 (1997) 688-702.

[92] D.S. Burdick, An introduction to tensor-products with applications to multiway data-analysis, *Chemometrics Intell. Lab. Syst. 28* (1995) 229-237.

[93] Y. Imamoto, M. Kataoka, F. Tokunaga, T. Asahi, H. Masuhara, Primary photoreaction of photoactive yellow protein studied by subpicosecond-nanosecond spectroscopy, *Biochemistry 40* (2001) 6047-6052.

[94] M. Schmidt, S. Rajagopal, Z. Ren, K. Moffat, Application of singular value decomposition to the analysis of time-resolved macromolecular X-ray data, *Biophys. J. 84* (2003) 2112-2129.

[95] G. Golub, V. Pereyra, Separable nonlinear least squares: the variable projection method and its applications, *Inverse Probl. 19* (2003) R1-R26.

[96] M.L. Horng, J.A. Gardecki, A. Papazyan, M. Maroncelli, Subpicosecond measurements of polar solvation dynamics: coumarin 153 revisited. *J. Phys. Chem. 99* (1995) 17311-17337.

[97] V.H. Vilchiz, J.A. Kloepfer, A.C. Germaine, V.A. Lenchenkov, S.E. Bradforth, Map for the relaxation dynamics of hot photoelectrons injected into liquid water via anion threshold Photodetachment and above threshold solvent ionization. *J. Phys. Chem. A 105* (2001) 1711 - 1723.

[98] I.H.M. van Stokkum, Parameter precision in global analysis of time resolved spectra.

IEEE Trans. Instr. Meas. 46 (1997) 764-768.

[99] A. Pandit, R.W. Visschers, I.H.M. van Stokkum, R. Kraayenhof, R. van Grondelle, Oligomerization of light-harvesting I antenna peptides of *Rhodospirillum rubrum. Biochemistry 40* (2001) 12913-12924.

[100] O.F.A. Larsen, I.H.M. van Stokkum, M.L. Groot, J.T.M. Kennis, R. van Grondelle, H. van Amerongen, Different electronic states in 2-aminopurine revealed by ultrafast transient absorption and target analysis. *Chem. Phys. Lett.* 371 (2003) 157-163.

[101] J. Hendriks, I.H.M. van Stokkum, K.J. Hellingwerf, Deuterium isotope effects in the photocycle transitions of the photoactive yellow protein. *Biophys. J.* 84 (2003) 1180-1191.

[102] W.D. Hoff, I.H.M. van Stokkum, H.J. van Ramesdonk, M.E. van Brederode, A.M. Brouwer, J.C. Fitch, T.E. Meyer, R. van Grondelle, K.J. Hellingwerf, Measurement and global analysis of the absorbance changes in the photocycle of the photoactive yellow protein from *Ectothiorhodospira halophila. Biophys. J. 67* (1994) 1691-1705.

[103] K. Ekvall, P. van der Meulen, C. Dhollande, L.-E. Berg, S. Pommeret, R. Naskrecki, J.-C. Mialocq, Cross phase modulation artifact in liquid phase transient absorption spectroscopy, *J. Appl. Phys. 87* (2000) 2340-2352.

[104] B. Borucki, H. Otto, M.P. Heyn, Reorientation of the retinylidene chromophore in the K, L, and M intermediates of bacteriorhodopsin from time-resolved linear dichroism: Resolving kinetically and spectrally overlapping intermediates of chromoproteins. *J. Phys. Chem. B 103* (1999) 6371-6383.

[105] M.P. Heyn, B. Borucki, H. Otto, Chromophore reorientation during the photocycle of bacteriorhodopsin: experimental methods and functional significance. *Biochim. Biophys. Acta 1460* (2000) 60-74.

[106] J.F. Nagle, L. Zimányi, J.K. Lanyi, Testing BR Photocycle kinetics. *Biophys. J. 68* (1995*)* 1490-1499.

[107] M. Chattoraj, B.A. King, G.U. Bublitz, S.G. Boxer, Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer, *Proc. Nat. Acad. Sci. USA* 93 (1996) 8362-8367.

[108] J.T.M. Kennis, D.S. Larsen, I.H.M. van Stokkum, M. Vengris, J.J. van Thor, R. van Grondelle, Uncovering the hidden ground state of green fluorescent protein, *Proc. Natl. Acad. Sci. USA, 52*, 17988-17993.

[109] W. Wohlleben, T. Buckup, J.L. Herek, R.J. Cogdell, M. Motzkus, Multichannel carotenoid deactivation in photosynthetic light harvesting as identified by an evolutionary target analysis. *Biophys. J. 85* (2003) 442-450.

[110] D.M. Bates, M.J. Lindstrom, Nonlinear least squares with conditionally linear parameters. In: Proc. of the Statistical Computing Section, American Statistical Association, New York, (1986), pp. 152-157.

[111] G.H. Golub, V. Pereyra, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM J. Numer. Anal. 10* (1973) 413-432.

[112] L. Kaufman, A variable projection method for solving separable nonlinear least squares problems. *BIT 15 (*1975*)* 49-57.

[113] G.H. Golub, R.J. LeVeque, Extensions and uses of the variable projection algorithm for solving nonlinear least squares problems. *Proc. of the 1979 Army Numerical Analysis and Comp. Conf.*, ARO Report 79-3 (1979) 1-12.

[114] S. Solar, W. Solar, N. Getoff, A semi-linear optimization model for resolving fast processes. *J. Chem. Soc., Faraday Trans. 2, 79* (1983) 123-135.

[115] I. Brooks, D.G. Watts, K.K. Soneson, P. Hensley, Determining confidence intervals for parameters derived from analysis of equilibrium analytical ultracentrifugation data, *Meth. Enzymol.* 240 (1994) 459-478.

[116] R. Pribić, I.H.M. van Stokkum, A. van den Bos, Inference regions of nonlinear parameters in global analysis, in: *Proceedings SYSID '94, 10th IFAC Symposium on System Identification*, M. Blanke, T. Söderström (Eds.) Copenhagen: Danish Automation Society, 1994, pp. 2.423-2.428.

[117] *Splus Reference Manual,* Statistical Sciences Inc., Seattle, 1991.

[118] L. Ujj, S. Devanathan, T.E. Meyer, M.A. Cusanovich, G. Tollin, G.H. Atkinson, New photocycle intermediates in the photoactive yellow protein from *Ectothiorhodospira halophila*: picosecond transient absorption spectroscopy. *Biophys. J. 75* (1998) 406-412.