Data analysis in transient electronic spectroscopy – an experimentalist's view

Beckwith, Joseph; Rumble, Christopher Allen; Vauthey, Eric

How to cite

BECKWITH, Joseph, RUMBLE, Christopher Allen, VAUTHEY, Eric. Data analysis in transient electronic spectroscopy – an experimentalist's view. In: International Reviews in Physical Chemistry, 2020, vol. 39, n° 2, p. 135–216. doi: 10.1080/0144235X.2020.1757942

This publication URL: https://archive-ouverte.unige.ch/unige:137443
Publication DOI: 10.1080/0144235X.2020.1757942

International Reviews in Physical Chemistry
Vol. 00, No. 00, Month–Month 20xx, 1–90

Review

Data Analysis in Transient Electronic Spectroscopy - An Experimentalist's View

Joseph S. Beckwith,[†][*] Christopher A. Rumble and Eric Vauthey[*]

Department of Physical Chemistry, University of Geneva, 30 Quai Ernest-Ansermet, CH-1211 Geneva, Switzerland; [†]Present Address: Department of Chemistry, Frick Laboratory, Princeton University, Princeton, New Jersey 08544, United States

(*v4.0 released August 2011*)

Time-resolved electronic spectroscopy has grown into a technique that provides hundreds to thousands of electronic spectra with femtosecond time resolution. This enables complex questions to be interrogated, with an obvious cost that the data are more detailed and thus require accurate modelling to be properly reproduced. Data analysis of these data comes in a variety of forms, starting with a variety of assumptions about how the data may be decomposed. Here, four different types of analysis commonly used are discussed: band-shape analysis, global kinetic analysis, lifetime distribution models, and soft-modelling. This review provides a 'user's guide' to these various methods of data analysis, and attempts to elucidate their successes, domains in which they may be useful, and potential pitfalls in their usage.

Keywords:
Band-Shape Analysis
Lifetime Density Analysis
Maximum Entropy
Global Analysis
Global Lifetime Analysis
Global Target Analysis
Soft-Modelling
SVD
MCR-ALS
Transient Electronic Spectroscopy
Transient Absorption
Transient Fluorescence
Data Analysis

Contents

PAGE

Introduction

Ever since the inception in 1950 of flash photolysis,[1, 2] time-resolved spectroscopy in the electronic domain (hereafter referred to as transient spectroscopy or a

---

[*] Email: joseph.beckwith@princeton.edu
[*] Email: eric.vauthey@unige.ch

ISSN: 0144235X print/ISSN 1366591X online

http://www.tandfonline.com

transient experiment) has been widely recognised as a significant tool for the unravelling of chemical dynamics and kinetics.[3, 4] The initial implementation of flash photolysis had a time resolution of 1 millisecond at best and very few spectra could be collected for each reaction, meaning that the interpretation of the results had to be relatively simple. Now, the terrain is considerably different - pulsed lasers and precise, automateable delay stages mean that time-resolved techniques that record hundreds to thousands of electronic spectra with femtosecond time resolution are now widespread.[5, 6] This allows one to answer more detailed questions, but this comes with a cost: the data that these instruments provide has also become more detailed and thus requires accurate modelling to be properly reproduced. This review discusses this issue - data analysis in transient electronic spectroscopy.

This review is organised in the following fashion: in Section 1 we will describe the format of electronic transient spectroscopy data, as well as any spectral artefacts that can impede analysis of the underlying kinetics. Furthermore, we will describe how one can remove and/or model these artefacts.

In Section **??**, we will introduce and explain 'band-shape analysis' methods - here, the shape of the electronic spectrum is mathematically described and/or empirically decomposed, and the temporal behaviour is inferred afterwards from the spectral decomposition.

In Section 3, we will describe here-called 'global kinetic' models, as they combine the assumptions of classical kinetics with the simultaneous analysis of all time traces measured in the experiment.

In Section 4, we will discuss 'lifetime distribution' models, including the maximum entropy model. These models all assume that the underlying lifetimes in the experiment may be described by a smooth distribution and all attempt (in various manners) to recover said distribution.

Finally, in Section 5 we will discuss 'soft-modelling' methods - these include multivariate curve resolution, singular value decomposition, and related methods. These methods do not insist upon a formal mathematical description of the time traces nor the band shapes, and can in some sense be described as dimensionality reduction methods (though, formally, one could think of all the methods described in this review as such).

This review does not discuss an in a sense 'alternative' method, i.e. the use of (a variety of) simulation techniques (possibly parametrised using experimental data, though not necessarily) to calculate observables and compare them to transient spectroscopy data. This does not in any sense denigrate this approach - it is most valuable, and has provided a variety of interesting insights in electron transfer,[7–10] vibrational relaxation,[11, 12] donor-acceptor dyads [13] and donor-acceptor complexes'[14] structural dynamics. It is however different in conceptual approach to the data analysis techniques reviewed here.

1. The Experiments and their Data

We will concern ourselves in this review primary with two experimental methods: broadband electronic transient absorption (TA) and broadband fluorescence upconversion spectroscopy (FLUPS). These spectroscopies deal primarily with electronic absorption and emission covering a range of ∼350-800 nm, with TA experiments sometimes extending into the near-IR. Familiarity with such experiments has been assumed for the purposes of this review, and we only provide here a cursory overview of the experimental principles. For more detailed descriptions, the reader is referred to the papers of Lang,[15] Megerle *et al.*[16], Zhang *et al.*[17] and Gerecke *et al.*[6]

Both TA and FLUPS experiments begin with impinging a UV-vis laser pulse upon a sample in order to create a small excited state population at a well-defined point in time. This pulse is termed the ‘pump’ and the time at which it arrives is taken to be $t = 0$ (‘time-zero’). In a TA experiment, a white-light ‘probe’ pulse covering the UV-vis region is passed through the sample at some later time in order to investigate differences in the sample’s absorption spectrum due to the effect of the pump pulse. The time between the pump and probe pulses is then varied and a ΔA spectrum recorded at each time step. For FLUPS, the emission of the pumped sample is focused onto a non-linear crystal and crossed with a delayed ‘gate’ pulse in order to upconvert a small temporal slice of the fluorescence through sum frequency generation. This upconverted fluorescence is then separated from the rest of the emission and its spectrum imaged and recorded. The delay between the pump and gate is then scanned and an emission spectrum recorded at each position. The data resulting from both TA and FLUPS experiments are typically stored as an $m \times n$ matrix, $\mathbf{D}$, with m representing the rows that make up the wavelength/wavenumber axis and n the columns that constitute the time axis.

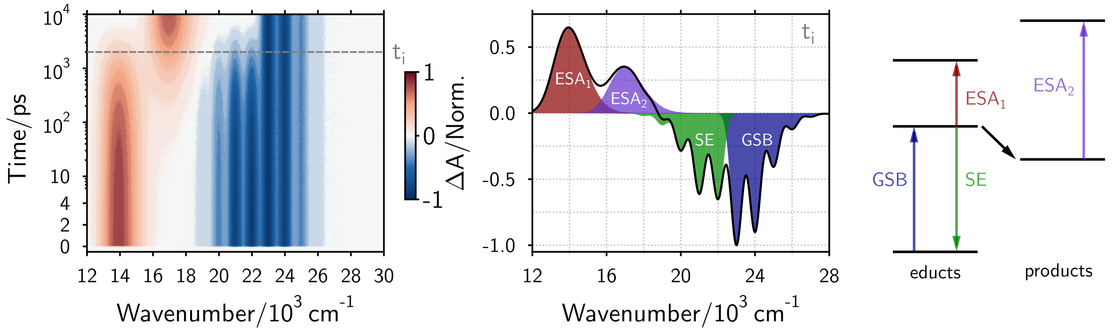


Figure 1. Simulated transient electronic absorption spectra and a schematic energy level diagram of the transitions involved. The line marked ‘t_i’ on the left panel corresponds to the time from which the spectrum in the middle panel was extracted. These data have been simulated according to the procedure described in Appendix A. The bands in the middle panel and transitions outlined on the right panel are the ground state bleach (GSB), stimulated emission (SE), and two excited state absorptions (ESA_1 and ESA_2) as described in the text.

Because TA experiments are difference measurements, the signal may have both positive and negative contributions. Positive contributions arise from new absorptions induced by the pump. These may arise from excited state absorption, ESA, of the initially prepared state or photoproducts from state changes or reactions, or from e.g. a non-equilibrated ground-state population. Negative contributions arise from the missing ground state population, giving rise to a ground-state bleach (GSB), as well as from stimulated emission (SE) due to interactions between the probe pulse and excited state population. These contributions are illustrated schematically in Figure 1. For FLUPS experiments, the signal is the gated emission intensity at each time step and is always positive. This signal arises from emission of the excited population prepared by the pump in addition to any emissive states populated thereafter. The stimulated emission measured by TA and fluorescence collected by FLUPS originate from the same photophysical process and should, in principle, exhibit the same dynamics (though the fluorescence dynamics may be difficult to extract from the TA due to the presence of overlapping bands).

The remainder of this section will discuss six important considerations when processing TA or FLUPS data: background contributions and how they may be removed; the instrument response function and the limitations it puts on the experimental time resolution; various artefacts arising from the interaction of the laser pulses with the solvent, solute, sample cell, and themselves; calibration of pixel vs. wavelength; chirp correction and identification of time-zero; and, finally, spectral response corrections.

1.1. Background Subtraction

Background contributions in transient spectroscopy are defined here as any contributions to the data which are not time-dependent, i.e. contributions which have no time-dependence upon pump and probe but which nonetheless manifest in the data. The most common background contributions in transient experiments are scattered light from the pump and spontaneous emission (which, being emission that is not 'gated' by the white-light pulse, arrives at the detector as a background contribution). In order to remove these, an average of multiple spectra from negative time delays (i.e. well before the temporal overlap of pump and probe) is subtracted from the overall data matrix, under the assumption that they do not change throughout the experiment. For the scattered pump this does not always hold due to fluctuations of the laser intensity during the experiment. Thus, the pump region is often simply removed from the experimental data matrix (or not measured at all, as is the case in the vast majority of FLUPS experiments). An example of the subtraction procedure is shown in Figure 2.

1.2. The Instrument Response Function

As in all time-resolved spectroscopies, the ultimate time-resolution of a TA or FLUPS experiment is mostly limited by temporal widths of the pump and probe/gate pulses. One way of characterizing this resolution is to measure the response generated by the instrument to an infinitely fast stimulus, typically accomplished using measurements of a scattering process or the optical Kerr effect (OKE).[6, 16, 18, 19] The resulting measurement is termed the 'instrument response function' (IRF). In a real measurement, the signal detected by the instrument will always have this response convolved

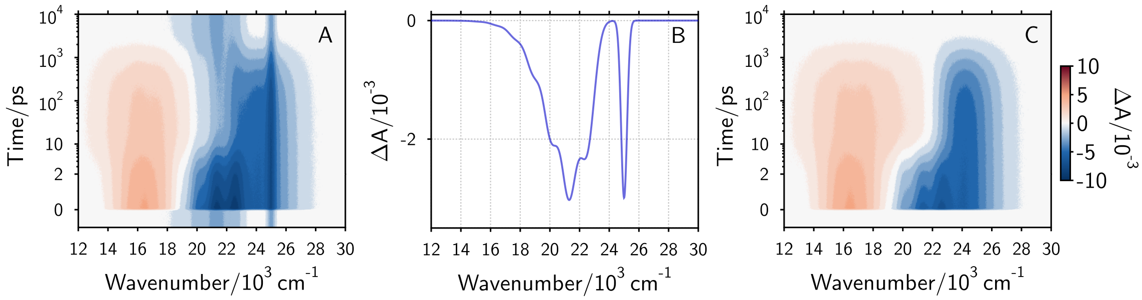


Figure 2. **A**. TA signal with pump scatter and spontaneous background emission present. **B**. An average of signals measured at negative time delays. **C**. TA signal with the spectrum plotted in **B** subtracted from every time step - only time-dependent components remain. For FLUPS data the principle is in essence the same, though the spectral region of the pump is not often measured, meaning only background emission is subtracted.

with that of the system being investigated:

$$S(t;\tilde{\nu}) = \mathrm{IRF}(t) \otimes I(t) \tag{1}$$

$$S(t;\tilde{\nu}) = \int_{-\infty}^{+\infty} \mathrm{IRF}(t-\tau;\tilde{\nu}) I(t;\tilde{\nu}) d\tau, \tag{2}$$

where $S(t;\tilde{\nu})$ is the signal recorded by the instrument at time t and wavenumber $\tilde{\nu}$, $\mathrm{IRF}(t;\tilde{\nu})$ is the instrument response function, and $I(t;\tilde{\nu})$ is the signal arising directly from the sample. In our case, the signals are either ΔA for TA or the fluorescence intensity for FLUPS.

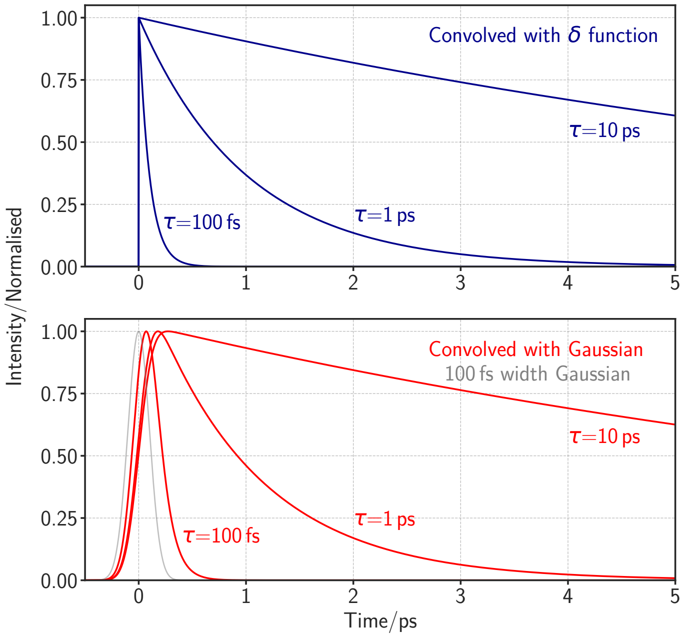


Figure 3. A comparison of decays convolved with a δ function (black, upper panel) and convolved with a 100 fs width Gaussian IRF (decays in red, Gaussian IRF in grey; lower panel). As may be seen, around the region where the IRF is prominent the decay is considerably broadened with respect to the decay convolved with the δ function. This thus becomes most relevant where the lifetime is comparable to the IRF width.

As shown in Figure 3, the IRF serves to temporally broaden the observed decay during the time when the IRF is prominent. The width of the IRF, and the corresponding achievable time-resolution, depends strongly on factors such as sample thickness, the walls of the cuvette, and group velocity dispersion from other transmissive optical elements. These sources of broadening can be compensated for in the experiment with the use of prism compressors or chirped mirrors to recompress pulses before they reach the sample and maximize time-resolution. Due to the dependence of the solvent refractive index on wavenumber, the width of the IRF also changes with the probe wavenumber [5, 20]. This can effect a 500 fs difference in IRF width across the visible spectrum as demonstrated in Figure 4. Therefore, one must keep this in mind when simultaneously fitting data collected at multiple wavenumbers. No matter the IRF, broadening can mask and distort data collected during the timescale of the IRF, and one must find a way of dealing with this in order to extract reliable dynamics from an experiment.

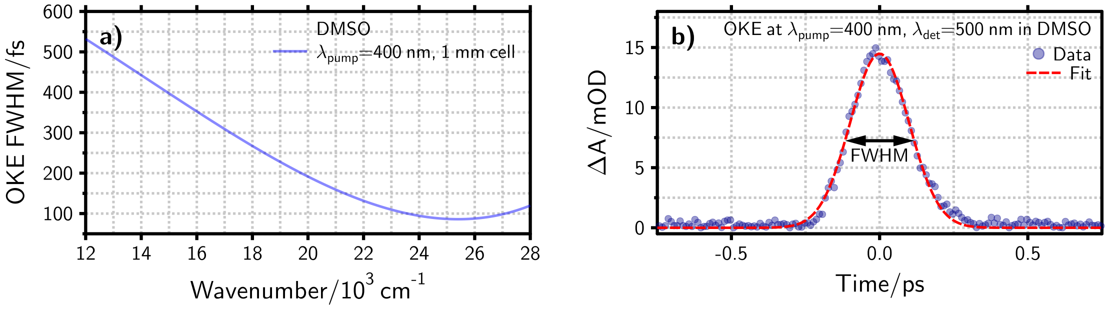


Figure 4. **a)** A 4th-order polynomial fitted to experimental FHWM data points derived from an OKE measurement of DMSO, excited at 400 nm, done by the authors (i.e. a measurement of the instrument IRF). **b)** Example OKE trace (blue dots) and a Gaussian fit (red, from which the FWHM was extracted). Note in **a)** that the FWHM changes by almost 500 fs across the electronic spectrum.

1.2.1. Deconvolution

For processes that are significantly slower than the IRF, one can simply ignore the short-time portion of the experiment and deal only with data at later times which are unperturbed by the IRF. In a typical ultrafast experiment this amounts to truncating all data before 0.3–0.5 ps. On the other hand, processes such as solvation dynamics, vibrational relaxation, and ultrafast reactions can approach the time scale of the IRF and be impossible to treat without removing the IRF broadening. One must then turn to deconvolution procedures which are intended to remove the effect of the IRF on the data and extract the 'intrinsic' system response from the experiment.

IRF deconvolution is typically accomplished by convolute-and-compare fitting algorithms which construct $S(t;\tilde{\nu})$ according to Eq. 2 in one of two ways. In the first method, a measured IRF is convoluted numerically with a parameterized model of $I(t;\tilde{\nu})$. Least squares algorithms are then used to optimize the parameters of $I(t;\tilde{\nu})$ versus the experimental data in order to extract the ideal system response.[19] While effective, this method can encounter problems due to slight differences between the systems used to measure the IRF and those of the sample.[21]. A more robust method of deconvolution is to build $S(t;\tilde{\nu})$ by assuming an analytical form for the IRF and convolving that with the $I(t;\tilde{\nu})$ model. Such a method allows one to optimize the IRF and $I(t;\tilde{\nu})$ parameters simultaneously and typically results in higher quality fits. This can be accomplished with numerical convolution as before, but analytical expressions exist for the convolution of some common IRF and $I(t;\tilde{\nu})$ models. Such is the case for the commonly encountered sum-of-Gaussians IRF and multi-exponential $I(t;\tilde{\nu})$, where the analytical expression for this convolution is known and used to fit experimental data directly.[22] It can be shown (see e.g. Section 5.1.7 of [23]) that careful IRF deconvolution can allow the determination of time constants shorter than the IRF, but only under the conditions high point-density and signal-to-noise. As demonstrated in the following simulations, time-constants extracted in this way will come with corresponding higher uncertainties.

In order to demonstrate the limits of deconvolution, we present the results of fitting large simulated data sets with varying IRF widths, signal-to-noise ratios (SNRs), and underlying dynamics. In the simulations that follow, we will assume our signal originates from multi-exponential dynamics convoluted with a Gaussian IRF. Such a signal can be calculated according to

$$I(t) = \sum_{1}^{N} \frac{A_N}{2} \exp\{-\tau_N^{-1}(t-t_0)\} \exp\{0.5 \cdot \tau_N^{-2} \cdot \sigma^2\} \left[1 + \mathrm{erf}\left(\frac{t - t_0 - \tau_N^{-1} \cdot \sigma^2}{\sqrt{2}\sigma}\right)\right], \quad (3)$$

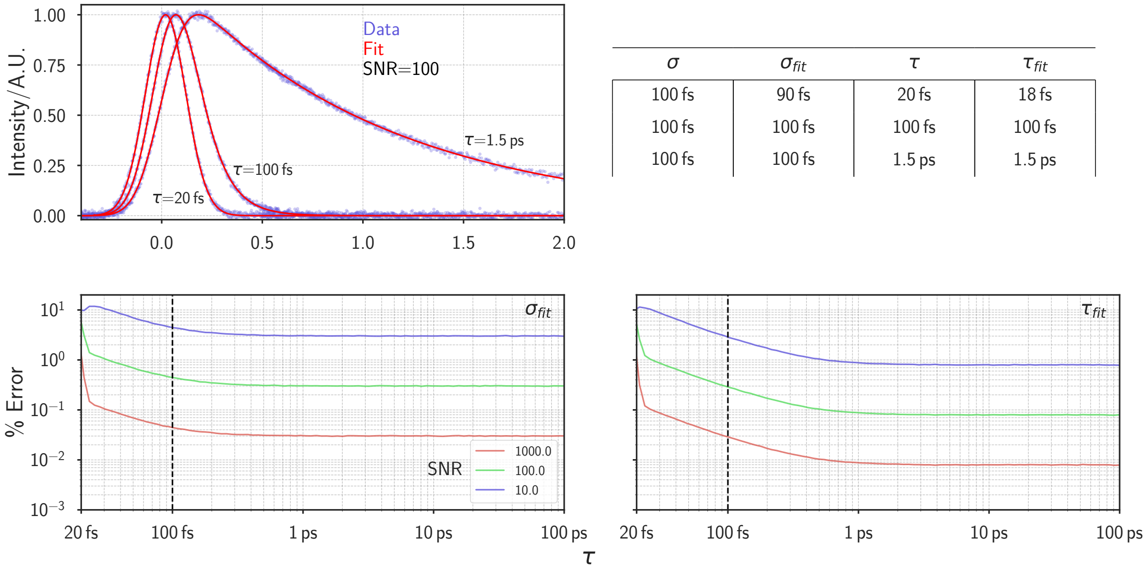
where N is the number of exponential components, A_N and τ_N are the amplitude and lifetime of component N, t_0 is the time-zero offset assumed here to be 0, σ is the width of the Gaussian IRF, and 'erf' is the error function. Gaussian noise is then added to each point to simulate real experimental conditions.

The simulated noisy data are then fit using Eq. 3 to attempt to extract the parameters used to generate the simulations. The uncertainty in the fitted parameters will be quantified using the standard error:

$$\theta_n(err) = \frac{\theta_n - G_n}{G_n} \times 100, \quad (4)$$

where θ_n is the nth recovered parameter from the fit and G_n is the corresponding parameter used to simulate the data. Further details of the simulations are provided in Appendix A.

The first set of simulations we will consider deals with the effect of the IRF and SNR on the determination of fast dynamics. Here we have simulated monoexponential signals according to Eq. 3 with a fixed IRF width (σ) of 100 fs, exponential time constants (τ) ranging from 20 fs to 100 ps, and Gaussian noise with SNRs of 10, 100, and 1000. Each $[\sigma, \tau]$ pair is simulated 10,000 times then fit with Eq. 3. The standard errors of σ_{fit} and τ_{fit} are then calculated and averaged. The results of these simulations are presented in Figure 5. We see in the bottom panel of Figure 5 that the relative errors in σ_{fit} and τ_{fit} track strongly with SNR, but are $< 3\%$ for all simulated



σ	σ_{fit}	τ	τ_{fit}
100 fs	90 fs	20 fs	18 fs
100 fs	100 fs	100 fs	100 fs
100 fs	100 fs	1.5 ps	1.5 ps

Figure 5. Upper panels: Example fits (red) to simulated data (blue). Longest trace shows an accurate fit due to a large τ relative to the σ, the middle trace an accurate fit where τ=σ and on the shortest trace a less accurate fit where the $\tau << \sigma$. Lower panels: Percentage error on recovered parameters from fitting the equation 3, where N=1, to data simulated using the same equation. Data points are averaged accuracy of 10,000 fits. σ = 100 fs. Further details on all simulations may be found in appendix A.

time constants longer than 0.5 ps. Therefore, one can be quite confident when extracting time constants a factor of 5 or more greater than the IRF width even under poor SNR conditions. These errors begin to grow as τ approaches the IRF width and reach up to ∼10% for SNR = 10 when τ is a factor of 5 smaller than the IRF. These errors are still reasonably good for such short time constants and demonstrate the power of good deconvolution for extracting fast dynamics from experimental data.

Although deconvolution can produce statistically good results even under poor experimental conditions, IRF broadening is not the only source of uncertainty in TA and FLUPS experiments. Day-to-day variation in instrument performance and sample preparation can have profound impacts on the reproducibility of the measured dynamics, and can greatly outweigh the uncertainty from just the data processing method. It is therefore imperative that, when possible, replicate measurements be conducted in order to properly characterize uncertainties in the reported dynamics. Additionally, deconvolution requires making an assumption about the dynamics within the IRF. These are, often, assumed exponential. However, this needs not always be the case - early solvation dynamics are non-exponential.[21]

1.3. Additional Ultrashort Pulse Induced Artefacts

Sub-picosecond pulses can induce additional spectral artefacts which can contaminate the dynamics under investigation. For TA data, the artefacts liable to be present are *cross-phase modulation*, *Raman scattering*, *two-photon absorption of the solvent*, the *coherent artefact* and the *perturbed polarisation decay*. For FLUPS data, only Raman scattering of the solvent is observed. Cross-phase modulation

(XPM) is a signal that that originates from the time-dependent modulation of the refractive index due to the intense pump pulse.[24] When the pump and probe pulses overlap in time and space, the modulation of the refractive index is experienced by the spectrally broad probe and thus the probe's spectral distribution is modified, giving rise to a signal near time-zero. Raman artefacts (RAs) arise when the pump wavelength is close to the probe wavelengths. RAs arise from low-frequency vibrations excited by an ultrashort pump pulse, and the amplitude of the change can be considerable.[5, 24] Two-photon absorption (2PA) arises from the high incident peak intensity of short laser pulses, which may enable the simultaneous absorption of two photons. This can lead to a distortion in the signal around time-zero in a transient experiment, though if the excitation wavelength is above 350 nm, the only solvents that have been observed to show 2PA are benzene and toluene.[24, 25] The coherent artefact is an effect that arises when the pump and probe pulses overlap in both frequency and time, and may be thought of as originating from a third-order non-linear polarization produced by the pump and probe beams. This polarization then radiates into the direction of the probe pulse, resulting in an increase of the probe intensity, the coherent artefact.[26] As this arises when the pump and probe are at the same frequency, it is often obscured by scattering from the pump. The perturbed polarisation decay occurs when the probe pulse interacts with the system before the pump pulse (i.e. before time zero). It may be interpreted as the pump pulse perturbing the polarisation created by the probe pulse, leading to a variation of the light transmitted just before time zero, which in the spectral domain translates to an oscillation with a frequency proportional to the time delay between pump and probe pulses.[27, 28] This effect is only significant when the dephasing time is longer than the duration of the pulses involved.[29] Both the coherent artefact and the perturbed polarisation are intrinsic to any pump-probe measurement, and do not arise from the solvent. In principle this means that they cannot be simply subtracted, as XPM an RA signals may be. They are both, in principle, able to be simulated using time-dependent perturbation theory, and are often more relevant in degenerate pump-probe experiments (i.e. where the pump and probe frequencies are identical). An example of XPM and RA signals in TA data is shown in Figure 6.

These signals originate from the temporal overlap of the pump and probe pulses, therefore they disappear outside of the cross-correlation region. For a rigorous treatment of the origin of the coherent artefact, XPM and RAs the reader is referred to the derivations by Kovalenko *et al.*[5] and Ekvall *et al.*[20] There are multiple methods by which these signals may be eliminated, and three most common approaches will be reviewed here. It should be noted that, as mentioned in Section 1.2.1, if extremely fast dynamics are not under investigation it is sufficient to only consider data outside of the cross-correlation region and avoid these artefacts completely.

1.3.1. Subtraction

The simplest method entails subtraction of the neat solvent response measured under the same experimental conditions. This involves a relatively simple, but non-trivial,[30, 31] measurement and is easy to implement in the data post-processing

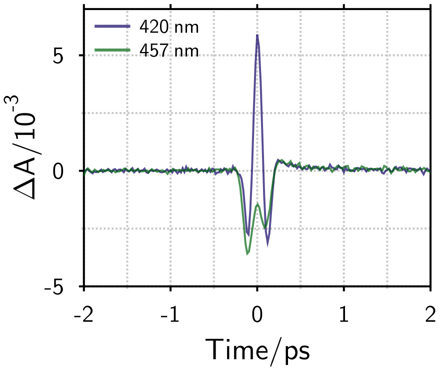


Figure 6. Cross-phase modulation (blue) and combination of stimulated Raman scattering and cross-phase modulation (green) measured in dichloromethane upon 400 nm excitation. Signal at 457 nm assigned as being partially due to Raman scattering due to its similarity to simulated RA signals (see e.g. Figure 10 of reference [5]) and its occurrence at the anticipated Raman shift for dichloromethane excited at 400 nm.

procedure. Often in this method involves subtracting the measured signal scaled by a factor typically determined by simple visual inspection. One constructs the corrected signal $\mathbf{D_{corr}}$ by

$$\mathbf{D_{corr}} = \mathbf{D_{init}} - \overrightarrow{\eta} \cdot \mathbf{S} \tag{5}$$

where $\mathbf{S}$ is the measurement of the neat solvent, $\mathbf{D_{init}}$ is the initial data matrix and $\mathbf{D_{corr}}$ is the corrected data matrix and $\overrightarrow{\eta}$ is a wavenumber dependent scaling factor, dependent upon the sample absorption.

1.3.2. Modelling of the Solvent Signal

A second, considerably more involved method for subtracting the artefacts involves modeling the signals explicitly [5, 20] from a measurement of the solvent alone, followed by subtraction. This method is more established in transient absorption, where more artefacts are present. To model the 2PA, one can use a simple Gaussian or skewed-Gaussian function.[24, 25] To model the XPM, the simplest method is to fit it with the the frequency-dependent cross-correlation function [5] $F_{cc}(\omega_2, t)$ (where ω_2 is probe frequency and t is time delay) between the pump and supercontinuum probe, as well as its 1st and 2nd derivatives. For the RAs, two equations (one for the low-frequency impulsive-stimulated Raman and one for the high-frequency Raman excitations) are possibly needed for signal modelling.

It must be noted that in the ‘best’ case, i.e. only XPM contributions at the particular wavenumber under consideration, the model is a fitting function with 5 adjustable parameters, thus iterating over all the wavenumbers measured and subtracting all of these modelled signals is a potentially quite laborious process. Where Raman contributions become relevant, the amount of parameters involved in the model becomes almost absurd. It is for this reason that it is not recommended to explicitly model the artefact signals for their subtraction.

1.3.3. Asymmetric Least-Squares

A third method of artefact removal is to use penalised least-squares. This method is based on the 1922 mathematical work of Whittaker [32] and was implemented by Eliers.[33] This was then slightly modified to use an asymmetric least-squares method and applied to ultrafast TA data by Devos *et al.*[30] The goal of this method is to 'smooth out' the sharp signals that arise from XPM and other contributions, leaving only the 'smooth' signals originate from the chemical dynamics.

Devos *et al.* note that, in the implementation for TA artefact removal, cross-validation procedures (automated procedures to select optimal parameters, discussed in Section 4.4.5) cannot be used for the optimisation of the smoothing procedure,[30] meaning that one has to use a grid-search method on 'similar' synthetic data to optimise the root mean square error. This means that *a priori* one should be able to separate artefact and kinetics, which is the goal of the artefact removal - a somewhat circular piece of reasoning. Additionally, the authors admit that 'mild deformation' of their simulated data are obtained. Finally, one loses statistical information in the smoothing as one generates new data rather than in e.g. the subtraction/simulation methods, where one simply subtracts a signal. Thus for these reasons, this method is not preferred to the subtraction method.

1.4. Wavelength Calibration

For TA, the wavelength calibration (i.e. the conversion of pixel in the experiment to wavelength) may be performed by e.g. comparing the spectrum measured by inserting a Holmium oxide glass filter in the probe pulse path to a reference absorption spectrum of the same filter. The pixel-to-wavelength calibration is then performed by finding the best overlap of the measured sample with the reference using the equation

$$\lambda = (p \times scale) + constant \tag{6}$$

where p corresponds to the pixel in question, and *scale* and *constant* are the fitting parameters to optimise the overlap. It must be noted that the above equation assumes a grating spectrograph - for a prism spectrograph, the equation is different (taking into account the focal length of the optical element that couples into the spectrograph) and a full discussion of the pixel to wavelength correction procedure for this type of spectrograph is found in Section 3.1 of [16]. This correction procedure should be re-done after each realignment of the white-light generation optics, thus in general before each set of experiments.

For FLUPS, the situation (naturally) differs, as one may not take an absorption spectrum of a known standard. The simplest method is to take a light source with known spectral lines (e.g. a mercury lamp, or to scan the excitation source of a calibrated steady-state fluorometer) and couple this into the detection system. One may then record this data and calibrate the known spectral locations of the lines vs. pixels, with either equation 27 in reference [34] in the case of a prism spectrograph or equation 6 in the case of a grating spectrograph. This procedure should only need to be repeated if the spectrograph itself is realigned, thus a pixel-to-wavelength calibration for the FLUPS need not be done as frequently as in the TA.

1.5. Time-Zero Correction

In transient absorption, as well as FLUPS experiments, the time zero of the experiment will generally be dependent upon the probe wavelength, due to the manner of the generation of the white light pulse and/or the optics traversed before the pump and probe (or gate) pulses interact. A number of procedures exist to correct for this, the two simplest and most robust in terms of implementation (in our experience) are outlined here.

In transient absorption, in order to correct for the wavelength-dependent time-zero, the optical Kerr effect (OKE) of the pure solvent may be measured in identical conditions to the TA experiment itself. The OKE experiment utilises the ability of isotropic samples (liquids in this case) to become anisotropic under the application of an applied electric field (the Kerr effect) - in this case, the electric field associated with an intense light pulse. Thus, in essence, the polarization analyser before the probe detector is rotated such that the probe intensity is virtually extinguished. Then, when the intense pump pulse induces a birefringence in the solvent, the probe pulse is again visible due to the rotation of its polarisation upon propagation through this birefringent medium. OKE responses have two components: an instantaneous electronic response and a nuclear response with a finite time constant. The instantaneous response dominates for simple liquids,[35] and so the peak position in time of the OKE signal with wavelength may be used as a measure of the wavelength-dependent time zero and the time behaviour of the instantaneous response may be taken as a measure of the pump-probe cross-correlation.[36] The central position of the OKE signal may be easily determined at each wavelength (as the electronic part of the OKE had a well defined Gaussian shape in time and a constant sign over the region of interest) and these values may be then fitted with a polynomial of the form [18]

$$t_0(\lambda) = a + \frac{10^5 b}{\lambda^2} + \frac{10^6 c}{\lambda^4}. \tag{7}$$

This whole procedure allows one to precisely determine the wavelength-dependent time zero, $t_0(\lambda)$. The absorbance values at each wavelength are then reassigned to a modified time delay based on equation 7, via interpolation of the measured time traces. The algorithm that does so should be designed such that generation of a large number of data points on a ‘synthetic’ time grid that is far finer than that the time grid measured in experiment is avoided.

Alternately, some groups have discussed using the artefacts present during the cross-correlation period (discussed in more detail in the subsection preceding) to perform a similar procedure.[16] This is not preferred to the measurement of the OKE as to fit the coherent artefacts requires the use of a model requiring a very large number of parameters, as discussed above. By contrast, to fit the centre of the OKE is a simple Gaussian fit, which is extremely quick to calculate.

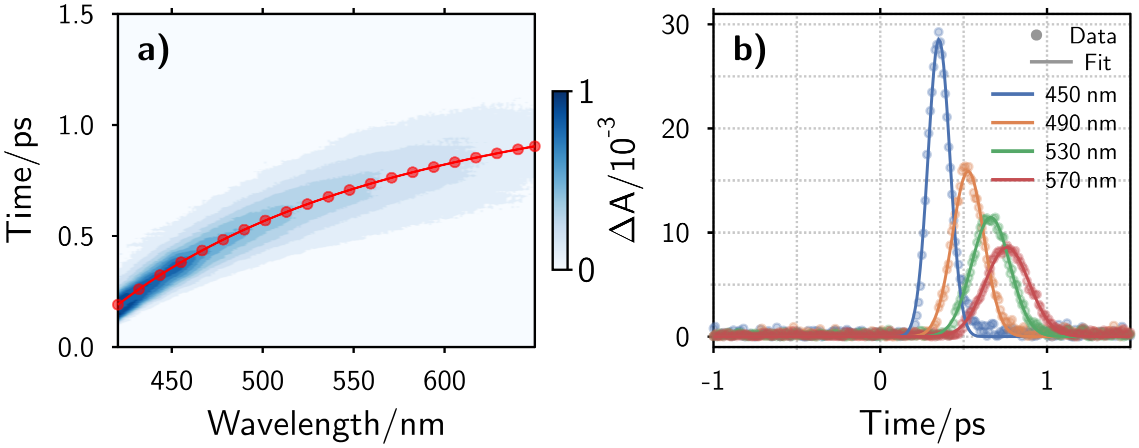


Figure 7. **a)** Electronic optical Kerr effect signal measured in DMSO upon 400 nm excitation. The solid line represents a fit of equation 7 to the central positions of the electronic OKE signal at each wavelength. Electronic OKE centre point density thinned for clarity. **b)** Example data (dots) and fit (lines) of the electronic OKE signal at several wavelengths, showing the wavelength-dependent time zero and FWHM of the electronic OKE signal.

For FLUPS experiments, the procedure for the correction of the time-zero may involve using a high-intensity pump pulse to generate a white-light continuum in the solvent used for experiments,[17] fitting these time profiles at each wavelength to a Gaussian then using equation 7 as above. This may be experimentally difficult however - Zhang *et al.* used 10 μJ to make white-light in methanol, which is a considerable amount of power (particularly when one considers the short pulse width of c. 40 fs used in these experiments). A more simple option experimentally (also discussed in reference [17]) is to measure the fluorescence of the laser dye BBOT, whose intensity is assumed to rise promptly upon excitation, in the solvent used in the experiments and fit the time profiles with a sum-of-exponentials (with the rise time reflecting the IRF) to accurately recover the time-zero as a function of wavelength, correcting again using equation 7.

1.6. Spectral Response Correction

If one wishes to extract meaningful properties (shapes, positions, populations) in a reliable fashion (i.e. in a way that facilitates inter-laboratory comparison) from transient spectra, it is essential that they accurately reflect the absorption or emission response of the chromophore in question - i.e. one needs to reach enough photometric precision to be able to compare transient spectroscopy data to steady-state data. Correction of the spectral response of the detection system may therefore be needed. For transient absorption spectroscopy, this (if chromatic aberration effects are excluded) is not in principle necessary due to the balanced detection scheme.[15]

For FLUPS however, the raw data must be corrected for the wavelength-dependent response characteristics of the instrument in order for the data to accurately report the correct emission spectrum of the sample as a function of time (assuming no effects such as reabsorption or the secondary inner-filter effect [37] are significant). In essence, the correction of the spectral response involves the measurement of an emissive 'standard' (or series of standards) with a known spectrum (or known spectra), and comparing the measured response with the correct emission spectrum. Various options for this exist - the use of a spectral irradiance standard, the use of a calibrated excitation source/calibrated reference detector, and the use of secondary emissive standards. The first option is, for most laboratories, prohibitively expensive [38] (in addition to aging effects on calibration lamps being a significant issue [39, 40]) and the second quite difficult to apply accurately.[38] For a fuller discussion of issues relating to these, as well as the construction of a correction function, references [38, 40] are both very thorough.

The use of secondary emission standards (i.e. a set of dyes with known spectral responses) is considered the easiest method to determine the wavelength-dependent spectral response function of one's instrument. Molecular dyes in liquid solutions are preferred, as these largely do not suffer from the problems of scattering that solid samples exhibit, nor the polarisation issues of phosphors (due to their rapid rotation in liquid solution). Multiple choices for such standards exist, the most 'common' of which we will discuss here. Gardecki and Maroncelli characterised a set of secondary emissive standards for steady-state fluorescence spectroscopy from an emission range of 300-800 nm and the spectral data may be found in reference [38]. These are all commercially available fluorophores and thus easily available, and were shown by Gardecki and Maroncelli to accurately recapitulate the National Bureau of Standards and Technology standard Quinine Sulfate. Whilst Resch-Genger *et al.* pointed out that some of these dyes exhibited photostability issues (in the case of two dyes) and minor anisotropy (in the case of one dye), these still represent a valuable set of emission standards due to their ease of availability and shown accuracy. The Federal Institute for Materials Research and Testing group in Berlin (the BAM) have also published and manufacture a set of secondary emission standards,[40, 41] which are less recommended as their chemical structures are unknown and the exact data of the spectral response unpublished, unlike the open-source nature of the procedure of Gardecki and Maroncelli. A set of secondary emissive standards specifically tested for use in FLUPS may be found in reference [17]. These, similar to the procedure of Gardecki and Maroncelli, are provided with functions that should recapitulate the standard shape of the spectrum needed for obtaining a correction function. As these dyes were specifically tested for the use in the FLUPS experiment, the dyes outlined in reference [17] are most recommended for this purpose (they are all also commonly available laser dyes).

For the FLUPS, the upconverted nature of the signal as well as the spectral response of the instrumentation must be taken into account when obtaining the fluorescence response. A thorough description may be found in Section III A of [17] and Section 6 of reference [34].

1.7. Conclusions

Thus we have a brief overview of the structure of transient spectroscopy data, and ways in which it may be processed before data analysis begins. We have been able to demonstrate the effects of the IRF upon the data, and that care must be taken when extracting parameters that have comparable time behaviour. We have also seen various methods of the removal of various ultrashort pulse induced artefacts in transient spectroscopy, and their relative advantages and disadvantages. Finally, correction for the probe-dependent time zero (essential in both TA and FLUPS experiments) and the spectral response (essential only in FLUPS experiments) was discussed.

2. Band-Shape Analysis

2.1. Introduction

We now discuss the methods of analysing these data and extracting the underlying dynamics. In order to do this, it is useful to first ask ourselves: 'what do absorption and emission spectra represent?' The observed absorption or emission intensity at a particular energy is related to the number of molecules undergoing transitions at that same energy. Therefore, the observed spectra are related to distributions of energies for populations of molecules in a particular state. By tracking changes in the intensity and shape of these distributions we may follow chemical and physical processes.

In a perfect world, one would be able to calculate all spectra and kinetics from first principles using quantum chemical methods. Methods for calculating steady-state and time-dependent electronic spectra do exist, but are technically demanding and prohibitively expensive for trajectories longer than a few tens of of picoseconds.[42, 43] Instead, experimentalists typically analyse spectra using empirical models from which physically meaningful parameters can be extracted. This has lead to the development of numerous forms of 'band-shape analysis' where particular functions or combinations thereof are used to simulate spectra. Fits are then performed independently at every time-point of the experiment and the time-dependent model parameters extracted for later analysis and interpretation.

Before covering a number of important band-shape functions and their uses, we must pause to consider two technical aspects of the fitting procedure. The first relates to the horizontal axis the spectra are presented in. Spectrometers based on grating spectrographs typically record spectra as intensity vs. wavelength due to the dispersion of a grating being linear with wavelength. As wavelengths are inversely proportional to energy, spectra presented in this representation appear compressed on the short-wavelength side and stretched on the long-wavelength side relative to their distribution in energy space. To correct for this, spectra can be presented as a function of wavenumber: $\tilde{\nu} = \lambda^{-1}$. For absorption spectra, which are ratios of intensities, this amounts to simply plotting the observed spectra vs. $\tilde{\nu}$ instead of λ. On the other hand, emission spectra are counting experiments and the non-linear energy bin size of wavelength spectra must be accounted for. This is accomplished by weighting the spectra according to

$$F(\tilde{\nu}) = F(\lambda)\lambda^2. \tag{8}$$

Although the spectra are now spaced linear in energy, they still cannot be used to directly calculate populations of states. The Einstein coefficients for absorption and emission are proportional to $\tilde{\nu}$ and $\tilde{\nu}^3$, respectively, and again distort the spectra with respect to the underlying population distribution.[44] By removing these dependencies we convert the spectra into the 'lineshape' or 'transition dipole moment'[45] representation:

$$\begin{aligned} a(\tilde{\nu}) &= A(\tilde{\nu})\tilde{\nu}^{-1} \\ f(\tilde{\nu}) &= F(\tilde{\nu})\tilde{\nu}^{-3}, \end{aligned} \tag{9}$$

where $a(\tilde{\nu})$ and $f(\tilde{\nu})$ are the absorption and emission lineshapes, respectively. Spectra in this representation are directly proportional to populations and can now be used for kinetic analyses. One drawback of fitting spectra in the lineshape representation is that the noise and uncertainty in the data are now weighted by $\tilde{\nu}^{-1}$ or $\tilde{\nu}^{-3}$ (or possibly $\tilde{\nu}^{-3}\lambda^2$). In order to fit lineshapes to spectra with proper weights without having to deal with error propagation, it is advisable to calculate the model fit in the lineshape representation and convert this fit to the wavelength representation for the calculation of residuals.[46]

2.2. Band-Shape Functions

A variety of functions for the fitting of band-shapes and approaches to their simulation exist in the physical and chemical literature, for numerous purposes. Here, we will describe popular and/or successful functions and approaches, though it is noted that an exhaustive description of all band-shape analysis approaches and functions is beyond the scope of this review.

2.2.1. Lorentzian

According to the phenomenological Bloch model,[47] an homogeneously broadened spectral line is Lorentzian and can be represented as

$$S(\tilde{\nu}) = \frac{I_0}{\pi\gamma}\left[\frac{\gamma^2}{(\tilde{\nu}-\mu)^2+\gamma^2}\right] \tag{10}$$

where the μ and γ refer to centre and width parameters and I_0 defines the band area.[48] Useful properties of the Lorentzian function are its analytical full width half maximum

$$S_l(\tilde{\nu})(FWHM) = 2\gamma \tag{11}$$

and analytical integral

$$\int_{-\infty}^{\infty} S_l(\tilde{\nu}) = I_0 \tag{12}$$

which enable easy quantification of the width and area of the band in question. It should be noted that this model is rarely useful in the simulation of electronic absorption and fluorescence spectra. However, the Lorentzian band shape is usually applicable to the infrared. This is due to the fact that the vibrational dephasing dynamics can often be well approximated by an exponential function.

2.2.2. Gaussian

Electronic dephasing dynamics in liquid are known to more complex than assumed by the Bloch model and cannot be accurately represented by a single exponential function. As a consequence, electronic band-shapes are generally not Lorentzian but can sometime be reproduced by a Gaussian function:

$$S_g(\tilde{\nu}) = \frac{I_0}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(\tilde{\nu}-\mu)^2}{2\sigma^2}\right\} \tag{13}$$

where the μ and σ refer to centre and width parameters and I_0 defines the band area.[48] This band shape does not often replicate true absorption and/or fluorescence spectra very well, though it has found use in the simulation of charge-transfer (CT) bands.[46, 49] Useful properties of the Gaussian function are its analytical full width half maximum

$$S_g(\tilde{\nu})(FWHM) = 2\sqrt{2\text{ln}2}\sigma \tag{14}$$

and analytical integral

$$\int_{-\infty}^{\infty} S_g(\tilde{\nu}) = I_0 \tag{15}$$

which enable easy quantification of the width and area of the band in question. It should be noted that this model will ill-fit spectra where some vibrational structure and/or asymmetry is observed.

2.2.3. Log-Normal

The log-normal function describes an asymmetric absorption or emission band, and is celebrated for its widespread applicability to the simulation of absorbtion and fluorescence spectra.[50] The widespread applicability arises from the fact that asymmetry is almost omnipresent in electronic spectra, due to the common scenario (discussed in a simplified model framework in reference [51]) of a differing curvature of the potential energy surface in the ground and excited electronic states. The log-normal function [48, 52] is defined as

$$S_{ln}(\tilde{\nu}) = I_0 \begin{cases} \exp[-ln(2)\{1+\alpha(\tilde{\nu})/\beta\}^2] \text{ if } \alpha(\tilde{\nu}) > -1 \\ 0 \text{ if } \alpha(\tilde{\nu}) \leq -1 \end{cases} \tag{16}$$

where

$$\alpha(\tilde{\nu}) = \frac{2\beta(\tilde{\nu} - \tilde{\nu}_0)}{\sigma} \tag{17}$$

with I_0 as the band intensity, β the asymmetry parameter, $\tilde{\nu}_0$ the peak position and σ the width parameter. Equally, a useful feature of the log-normal function is that of the analytical integral

$$\int_{-\infty}^{\infty} S_l(\tilde{\nu}) = I_0 \frac{\sigma}{2} \text{exp}\left[\frac{\beta^2}{4\text{ln}2}\right] \sqrt{\frac{\pi}{\text{ln}2}} \tag{18}$$

enabling easy quantification of band areas. The log-normal does not directly connect to a model of the physics of the chromophore and will not reproduce spectra with vibrational structure. However, it will perform considerably better than the Gaussian for asymmetric bands.

2.2.4. Displaced Harmonic Oscillator

The displaced harmonic oscillator (DHO) models for absorption,[53] and emission,[53, 54] are connected to a simplified model of the physics of absorption and emission, i.e. from a purely harmonic potential to another (displaced, thus the name). The DHO models, which may be derived from a classical treatment of the spectra,[53] are defined as

$$A_{abs}(\tilde{\nu}) = I\tilde{\nu}\sum_{m=0}^{\infty}\frac{S_{abs}^{m}e^{-S_{abs}}}{m!}\exp\left\{\frac{-(h\tilde{\nu}_{abs}^{0}+m\hbar\omega-h\tilde{\nu})^{2}}{2\sigma_{abs}^{2}}\right\} \tag{19}$$

and

$$F_{em}(\tilde{\nu}) = I\tilde{\nu}^{3}\sum_{m=0}^{\infty}\frac{S_{em}^{m}e^{-S_{em}}}{m!}\exp\left\{\frac{-(h\tilde{\nu}_{em}^{0}-m\hbar\omega-h\tilde{\nu})^{2}}{2\sigma_{em}^{2}}\right\}. \tag{20}$$

These represent a spectrum built on a progression of vibronic transitions with width $\Gamma = \sigma\sqrt{8\ln 2}$ built on a '0-0' frequency, $\tilde{\nu}^{0}$, and resulting from a single harmonic mode of frequency, ω, displaced by an amount $\Delta = (2S)^{1/2}$, with S representing the Huang-Rhys factor. Disadvantages of this model are that there are a large number of free parameters (though this may be remedied by e.g. measurement of steady-state absorption and emission spectra in non-polar solvents and fixing the vibrational frequency to these values)[55] and a lack of an analytical integral for the band area. Additionally, being a simplified physical model, the DHO model will not be able to fit spectra where significant anharmonicity plays a role in the spectra, or where the vibrational progression is more complex than a single 'effective' mode. Additionally, equation **??** only considers emission from the lowest vibrationally excited state. If emission from higher-lying vibrational states is to be considered, one must calculate the Franck-Condon factors for emission from an excited electronic vibrational state m to the n vibrational state of the electronic ground state and use

$$F_{m,n} = \exp(-S)m!n!\left[\sum_{r=0}^{min(m,n)}\frac{(-1)^{n-r}(\sqrt{S})^{m+n-2r}}{r!(m-r)!(n-r!)}\right]^{2} \tag{21}$$

and then use these to construct the fluorescence from all considered excited electronic vibrational states to all considered ground state vibrational states using

$$I_{em}(\tilde{\nu}) = I\tilde{\nu}^{3}\sum_{n}F_{m,n}\exp\left(-\frac{(\tilde{\nu}-\tilde{\nu}^{0}+(n-m)\tilde{\nu}_{vib})^{2}}{2\sigma^{2}}\right). \tag{22}$$

where, as above, the width, '0-0' frequency and single harmonic modes are defined by σ, $\tilde{\nu}^{0}$ and $\tilde{\nu}_{vib}$.

Recently, the formalism has been further extended by Fedunov *et al.* to explicitly include vibrational relaxation processes in the electronic excited state, as well as explicitly connecting the relaxation dynamics of the fluorescence spectrum to an arbitrary solvent relaxation function.[56] In addition, and significantly, they were also able to circumvent the long-standing problem [21] of the estimation of the fluorescence spectrum pre-solvation (discussed in Section **??**) by using information about the pump pulse used for excitation of the fluorophore to calculate the build-up dynamics of the excited-state wavepacket.[56]

2.2.5. Sulzer-Wieland Formalism

The Sulzer-Wieland formalism [57] is a formalism developed to describe the temperature dependence of the absorption spectrum of a diatomic molecule. It has also been used to compute the absorption of vibrationally 'hot' polyatomic molecules, e.g. after a rapid decay to a vibrationally excited electronic ground state after photoexcitation.[58] Within the formalism, the molar absorption coefficient of the molecule as a function of the temperature T may be computed as

$$\varepsilon(\tilde{\nu},T) = \varepsilon_0^m \frac{\tilde{\nu}}{\tilde{\nu}_{0,eff}} \left\{ \tanh\left(\frac{\Theta_{vib}}{2T}\right)\right\}^{1/2} \exp\left[-\tanh\left(\frac{\Theta_{vib}}{2T}\right)\cdot\left(\frac{\tilde{\nu}-\tilde{\nu}_{eff}}{\Delta\tilde{\nu}_0}\right)^2\right] \tag{23}$$

where

$$\Theta_{vib} = hc\tilde{\nu}_{vib}/k_B \tag{24}$$

is the characteristic temperature of the excited mode with vibrational frequency $\tilde{\nu}_{vib}$, h is Planck's constant, c is the speed of light and k_B is the Boltzmann constant. The peak molar absorption coefficient and the spectral width of the absorption band at a reference temperature (for example 0 K) are denoted by ε_0^m and $\Delta\tilde{\nu}_0$ respectively. The effective spectral position of the absorption band at temperature T is given by the spectral position at 0 K, $\tilde{\nu}_0$ and includes an instantaneous red shift that arises from the average vibrational energy in the excited mode i.e.

$$\tilde{\nu}_{eff} = \tilde{\nu}_0 - \tilde{\nu}_{vib}[\exp(\Theta_{vib}/T)-1]^{-1} \tag{25}$$

which introduces a temperature-dependent shift of the electronic resonance. Influence of vibrational excitation of a polyatomic molecule on the spectral shape of its electronic resonances is thus reduced to the coupling of a single 'effective' vibrational degree of freedom to the electronic transition dipole. The assumptions inherent in this model are as follows: the band is of a Gaussian shape, the upper state potential is assumed to be repulsive and is approximated by a straight line fit to the curved potential, and that the Franck-Condon integrals (and thus the oscillator strength) does not change with temperature. Despite the bold nature of these assumptions, the Sulzer-Wieland formalism has been successfully applied to the understanding of vibrational cooling of polyatomics. The Sulzer-Wieland formalism has been extended to polyatomics and systems with multiple effective relaxation modes, see Appendix A2 of the paper of Schalk *et al.*[58] for a more thorough discussion. For molecules with almost any structure of their absorption spectra however, it is often too crude to be useful.

2.2.6. Time-Zero Emission Estimates

The method of time-zero estimation of a fluorescence spectrum (i.e. pre-solvation) was introduced by Horng *et al.* [21] to estimate the position of the fluorescence spectrum prior to any solvent relaxation. This corresponds to the emission spectrum expected to be observed following vibrational relaxation, but before solvent relaxation. For some molecules (such as the Coumarin 153 the Maroncelli group used) this assumption may be assumed to be correct, but for other molecules this may not be accurate.[59] The calculation of the time-zero emission spectrum uses the absorption and emission spectrum of the probe in a reference solvent, and the absorption in the solvent under consideration. With the intrinsic line shape function of absorption denoted as $g(\tilde{\nu})$ and the distribution of spectral shifts δ caused by the solvent environment distribution denoted as $p(\delta)$, one expresses the absorption spectrum in a polar solvent (P) by

$$A_{\mathrm{P}}(\tilde{\nu}) \propto \tilde{\nu} \int g(\tilde{\nu} - \delta) p(\delta) \mathrm{d}\delta \tag{26}$$

and the time-zero emission spectrum is thus given as

$$F_{\mathrm{P}}(\tilde{\nu}, t = 0, \tilde{\nu}_{ex}) \propto \tilde{\nu}^3 \int g(\tilde{\nu}_{ex} - \delta) p(\delta) f(\tilde{\nu} - \delta) \mathrm{d}\delta. \tag{27}$$

The functions one thus needs to compute the time-zero emission spectrum are $g(\tilde{\nu})$, $f(\tilde{\nu})$ and $p(\delta)$. These are assumed to be related to the steady-state absorption and emission spectra in the non-polar reference solvent, using the relations

$$g(\tilde{\nu}) \propto \tilde{\nu}^{-1} A_{\mathrm{ref}}(\tilde{\nu}) \tag{28}$$

and

$$f(\tilde{\nu}) \propto \tilde{\nu}^{-3} F_{\mathrm{ref}}(\tilde{\nu}) \tag{29}$$

with the $p(\delta)$ function assumed to be Gaussian (equation **??**, with I_0 set to 1 and δ replacing the $\tilde{\nu}$). One may thus construct the time-zero emission estimate by iterative fitting of the polar absorption spectrum. This method was successfully used by Maroncelli when broadband fluorescence up-conversion techniques were not available, and is useful in any context where one has vibronic structure that will be lost as solvent relaxation occurs.[46, 49] Kumpulainen *et al.* recently showed it to be in good agreement with the time-zero emission spectra extracted using broadband fluorescence up-conversion techniques.[60] As noted above however, the recent extension of the DHO model by Fedunov *et al.*[56] was able to circumvent the need for this estimation.

2.2.7. Empirical Spectral Decomposition

If one does not have 'easily' computable band shapes for the measured spectra, then it is possible to decompose the spectra at time t using extracted band shapes from the separate experiments.[61] We take the example of a bimolecular photoinduced reaction:

$$\mathsf{M}^*(\mathsf{S}_1) + \mathsf{Q} \xrightarrow{\qquad} \mathsf{P}$$

where M represents the reactant that is photoexcited, Q the quencher and P the product. We may then write the observed TA spectrum as

$$\Delta A(c, \tilde{\nu}, t) = \Delta A_{\mathrm{M}}(c, \tilde{\nu}, t) + \Delta A_{\mathrm{P}}(c, \tilde{\nu}, t) \tag{30}$$

where Q is at concentration c, the spectrotemporal map $\Delta A(c, \tilde{\nu}, t)$ may be decomposed into a map of $\mathrm{M}^*(\mathrm{S}_1)$, $\Delta A_{\mathrm{M}}(c, \tilde{\nu}, t)$, and the products, $\Delta A_{\mathrm{P}}(c, \tilde{\nu}, t)$.

By measuring transient absorption maps of M*(S$_1$) alone, it is possible to access $\Delta A(c = 0, \tilde{\nu}, t)$ as well as the intrinsic excited state lifetime τ_{M} and recast the equation for the transient absorption map in the presence of the quencher as

$$\Delta A(c, \tilde{\nu}, t) = q_{\mathrm{M}}(c, t)\Delta A(0, \tilde{\nu}, t) + \Delta A_{\mathrm{P}}(c, \tilde{\nu}, t) \tag{31}$$

where $q_{\mathrm{M}}(c, t)$ is the survival probability of M*(S$_1$). To solve this problem, it may be assumed that the transition dipole is unaffected by any processes that give rise to spectral dynamics in ΔA_{M}. Thus, one may assume that the ratio of areas under different parts of the spectrum corresponding to the same given species are conserved. From this assumption, the dynamics of a photochemical reaction are obtained with the single assumption of conservation of transition dipole moment.[61] This, it should be noted, is not an assumption that may be reasonably made for most intramolecular processes due to many of these involving a spectral shift coupled to movement along a potential/free energy surface that will lead to a change in the transition dipole moment. This method (for which for full details see reference [61]) has key advantages, namely that there is no assumption of a specific functional form of the band shape, and that the time evolution of multiple different transients can be determined without the assumption of any model.

2.2.8. Numerical Band Integrals

Perhaps the simplest method of band-shape analysis is that of numerical band integrals. This is simply a method of taking the area of a spectral band numerically. It is also possible to access the band position, band width and band skewness using this procedure. To avoid effects of the frequency dependence, the data should be converted to the lineshape representation [45] before this analysis as detailed in Section **??**.

The band area, band position and band width may also be referred to as the 0th, 1st and 2nd moments of the specified range. The calculation of the 0th moment may be done by:

$$M_0(t, \Delta\tilde{\nu}) = \int_{\tilde{\nu}_1}^{\tilde{\nu}_2} D(t, \tilde{\nu})\ \mathrm{d}\tilde{\nu} \tag{32}$$

where $D(t, \tilde{\nu})$ refers to the time point of the data matrix in question. The calculation of the 1st moment may be done by:

$$M_1(t, \Delta\tilde{\nu}) = \frac{\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} D(t, \tilde{\nu}) \times \tilde{\nu}\ \mathrm{d}\tilde{\nu}}{M_0(t, \Delta\tilde{\nu})} \tag{33}$$

and the 2nd by:

$$M_2(t, \Delta\tilde{\nu}) = \sqrt{\frac{\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} D(t, \tilde{\nu}) \times (\tilde{\nu} - M_1(t, \Delta\tilde{\nu}))^2\ \mathrm{d}\tilde{\nu}}{M_0(t, \Delta\tilde{\nu})}}. \tag{34}$$

It should be noted that the higher order moments will have corresponding higher uncertainties and thus these values are correspondingly sensitive to data quality.

2.2.9. Absorptive Band-Shapes

A difficulty in the use of band-shape analysis in transient absorption experiments (as opposed to transient fluorescence experiments) is that one has multiple overlapping contributions of differing sign. The empirical spectral decomposition (Section **??**) method is able to deal with this, however does rely on experimental band shapes being available. For e.g. unimolecular systems, this may not be the case. A differing method to simplify transient absorption spectra is to add the steady-state spectra to the matrix in such a way that only positive contributions remain. This allows us to extract the excited state absorption lineshape without contamination of the negative ground state bleaches and stimulated emission. Thus explicitly we assume the GSB and SE are unchanging with time. The GSB we may describe using the steady state absorption spectrum, but the stimulated emission may not be described using the steady state emission spectrum, as this is spontaneous emission. To convert this into the stimulated emission spectrum, a correction factor,[62]

$$F(\lambda)_{stim} = F(\lambda)_{spont}\lambda^4, \tag{35}$$

must be applied. Here, $F(\lambda)_{stim}$ is the stimulated emission spectrum and $F(\lambda)_{spont}$ is the spontaneous emission spectrum. For systems with strong SE solvatochromism this approach will fail to remove the unrelaxed emission. In these cases, if one may use an alternate measurement (e.g. FLUPS) to fully describe this behaviour, it may then be possible to remove the unrelaxed emission. Additionally, this method may also have problems when relaxation back to the ground state potential energy surface is fast and thus 'hot ground state' features contaminate the GSB, as this may sufficiently distort the GSB area of the spectrum as to render the recovery of an entirely smooth filled-in spectrum difficult by only the addition of the steady-state spectrum.

Once negative features are removed, the data is more amenable to a band-integral analysis, or to other band-shape analysis. This has been used by Kovalenko *et al.*[63] and Wilcken *et al.*[64] for, in the former case, the analysis of vibrational cooling and, in the latter case, photoproduct quantum yield determination. In the case of Kovalenko *et al.* the removal of the GSB was accomplished visually, thus the procedure of using second derivatives (the procedure of Wilcken *et al.*) is preferred. This procedure involves taking the second derivative of a (smoothed) TA spectrum and comparing it to the second derivative of the corresponding steady-state spectrum of the same species. Where a 'mirror' image of the second derivative of the steady-state spectrum exists in the second derivative of the TA spectrum, this area provides a useful marker of how much of the steady-state spectrum one should add to remove it from the TA spectra. This mirror image peak should be removed i.e. the steady-state spectrum should be added until the second derivative is ideally flat (demonstrated in Figure **??**).

Thus, from this, one may add the steady-state spectra to one's TA data to remove these negative contributions. Wilcken *et al.* simply did this for a very small number

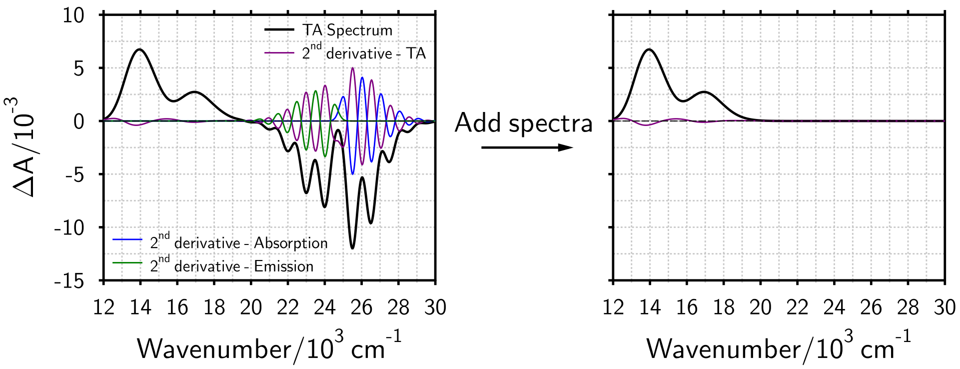


Figure 8. A demonstration of the addition method outlined by Wilcken *et al.*, as described above.

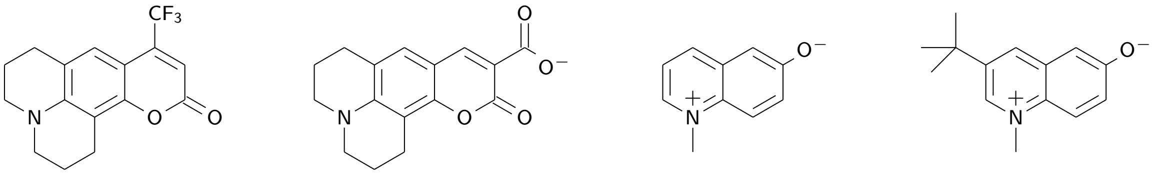
of spectra [64] for the measurement of quantum yields, and doing so for every time step in a TA experiment would likely be laborious. Thus, it is suggested that one may use Singular Value Decomposition (discussed in more detail in Section 5, and in reference [65]). One may add the steady-state spectra to the separated spectral components and then reconstruct the overall matrix using all of the SVD matrices with the spectral components edited to remove the steady-state components. In this way, the steady-state contributions are removed for the whole data matrix with the statistical information regarding the errors on one's measurement retained.

2.3. Published Implementations

Whilst many groups have used the methods of band-shape analysis, no implementations of a program (to our knowledge) exists in the ultrafast literature purely to fit band-shapes. The only program that does engage with fitting band-shapes is the program published by Grubb *et al.*, KOALA (Kinetics Observed After Light Absorption).[66] This program however only decomposes bands into sums of Gaussians or Lorentzians, considerably hampering its use in the electronic regime. In addition, Grubb *et al.* assert that shifting or narrowing of a spectrum is a result of a kinetic process, which is not necessarily accurate. The program also imposes a kinetic model whilst fitting the components of each spectra, which may be unhelpful for non-exponential processes.

2.4. Applications

A significant use of band-shape analysis has been in understanding solvation dynamics. In a landmark 1995 paper Horng *et al.* measured the time-resolved fluorescence of Coumarin 153 (**C153**, Figure **??**) in 24 solvents. This was achieved by using a single-wavelength fluorescence up-conversion instrument, taking time profiles at multiple different wavelengths for each of these solvents and reconstructing the time-resolved emission spectra.[21] These spectra were then fitted with log-normal functions and the band maximum and average frequencies taken from these fits were used as measures of solvent relaxation.



C153 **C343** **MQ** **3tBu-MQ**

Figure 9. Molecules used as fluorescent solvation probes.

These were then used to construct spectral response functions,

$$S_{\tilde{\nu}}(t) \equiv \frac{\tilde{\nu}(t) - \tilde{\nu}(\infty)}{\tilde{\nu}(0) - \tilde{\nu}(\infty)} \tag{36}$$

where $\tilde{\nu}(t)$ is the band maximum at time t, $\tilde{\nu}(\infty)$ is the band maximum after all solvent relaxation is complete, and $\tilde{\nu}(0)$ is the band maximum at $t = 0$ (accomplished in their case using the method outlined in Section **??**). A selection of these are shown in Figure **??**. These could then be analysed to give the characteristic solvation times in the 24 solvents measured. Using this method, they were able to demonstrate that a significant part of the solvent response is inertial, confirming predictions made by earlier classical molecular dynamics simulations. In addition, Horng *et al.* could compare their observed solvent response with predictions from models based on the dielectric response of the pure solvent and show that these provide semi-quantitative understanding of solvation dynamics, as well as showing cases of the failure of these theories - quadrupolar solvents such as dioxane and benzene are not well-described by the dielectric response.

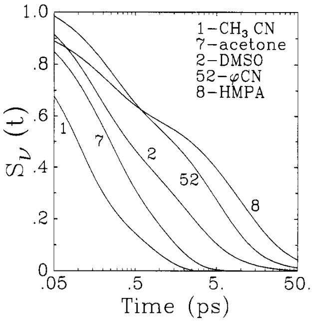


Figure 10. 5 spectral response functions constructed from fitting transient fluorescence data with log-normal functions, showing the variation of solvation dynamics with differing polar aprotic solvent. CH_3CN is acetonitrile, φCN is benzonitrile and HMPA is hexamethylphosphoramide. Reprinted with permission from reference [21]. Copyright 1995 American Chemical Society.

In the 1995 paper of Horng *et al.*, the effect of excess energy was perceived to be of minimal interference (determined by measuring the relaxation of **C153** in cyclohexane, where no solvent relaxation was anticipated to take place) and that vibrational relaxation should not affect the position of the fluorescence band and thus have no interference in the measure of solvent relaxation. This view was revisited with the considerably greater spectral resolution of the FLUPS experiment of Sajadi *et al.*, and it was shown that in fact if one excites a non-polar solute in a polar or non-polar solvent, the peak position may 'excessively' redshift (Figure **??**) if one excites with vibrational excess energy. This was also deduced using band-shape analysis, specifically from fitting a log-normal to the full spectrum at each time step and observing the time-dependence of the peak position and width.[59]

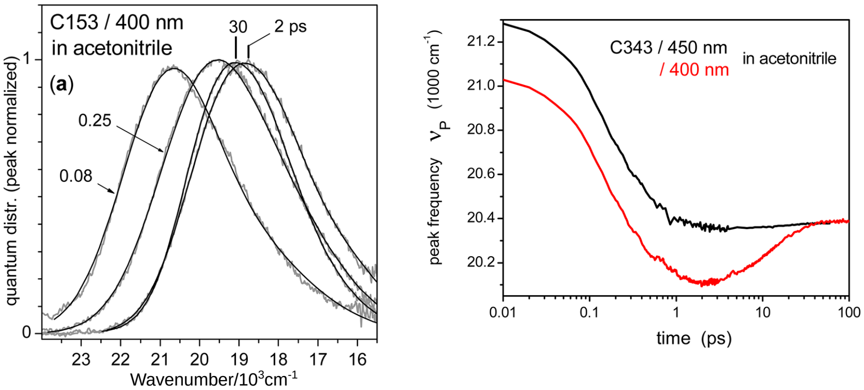


Figure 11. Example fits (left panel) of time-resolved fluorescence spectra of **C153** measured in acetonitrile. The band maximum (right panel, as deduced from a log-normal fit to the data) of **C343** (Figure **??**) if it is excited at 450 nm (black) or 400 nm (red). Note the magnitude of the shift maximum and minimum is the same, but the origins differ with vibrational excess energy (though the final position of the spectrum is the same). Adapted with permission from reference [59]. Copyright 2013 American Chemical Society.

The effect was interpreted in terms of the transient heating of the electronically excited chromophore, as well as the time-dependence of the cavity size of the molecular probe. As such, this improved spectral resolution coupled with the band-shape analysis techniques enabled Sajadi *et al.* to show that excitation at the red edge of the absorption spectrum is, in theory, a necessary condition before constructing the spectral response function and comparing this to theories of solvent response or simulations.[59]

Gerecke *et al.* also explored the behaviour of water molecules around nonpolar groups using N-methyl-6-oxyquinolinium betaine (**MQ**, Figure **??**) and a derivative where the H atom in position 3 is replaced with a tertiary butyl group (**3tBu-MQ**, Figure **??**).[67] This difference in the molecular structure between the two molecules lead to a minor difference in the solvation dynamics between the two (Figure **??**). This difference in solvation dynamics, by comparison with molecular dynamics simulations,[68] was ascribed to the coupling of solute vibrations to hydration water. This implied that a common view in the field of solvation dynamics - that the solvation dynamics measured should not be dependent on the molecular probe - is not necessarily correct. This elegant insight could only be made possible by the advent and usage of broadband fluorescence up-conversion, coupled with a band shape analysis that enabled the very small difference in shift times to be uncovered.[67]

Band-shape analysis has also recently been used in a study to detect excited

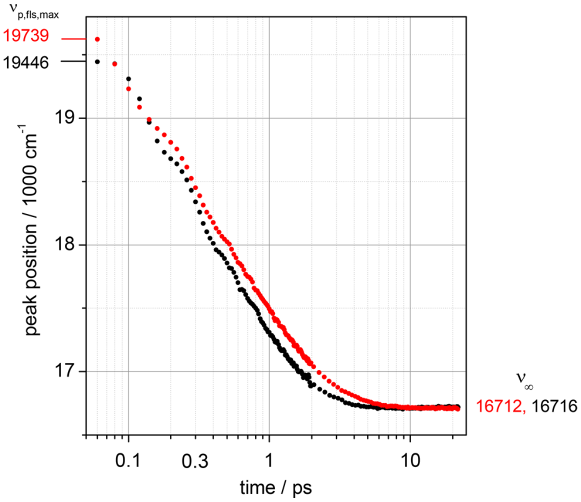


Figure 12. Fluorescence band maximum as a function of time for **MQ** (black) and **3tBu-MQ** (red) after excitation a 400 nm in water. Reprinted with permission from reference [67]. Copyright 2017 American Chemical Society.

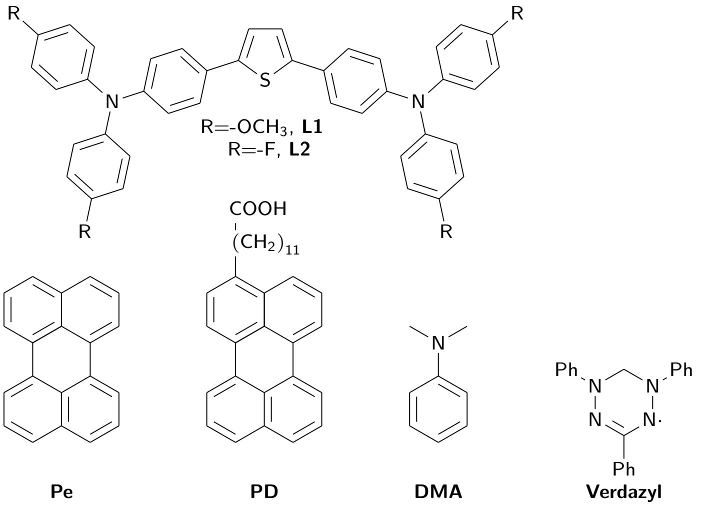


Figure 13. Molecules used as probes of excited-state symmetry breaking, vibrational energy relaxation, electron transfer and fast internal conversion.

state symmetry breaking (ES-SB) in the molecules **L1** (a molecule where ES-SB was anticipated to take place) and **L2** (a molecule where ES-SB was anticipated to not occur) (Figure **??**). ES-SB is a transition from a state where the electronic excitation is delocalised evenly over the whole molecule to a state where it is localised on one side of the molecule. Consequently, the excited state changes from quadrupolar to dipolar. ES-SB had previously been observed in real-time using transient IR spectroscopy,[69] but not in a molecule lacking localised IR marker modes. By analysing the time-resolved fluorescence of **L1** and **L2** using the DHO model, equation **??**, it was possible to decouple the spectral shift dynamics and the population dynamics. This is shown in Figure **??**. This enabled access to the instantaneous transition dipole moment, μ_{em} which was shown to track the excited-state symmetry breaking, with the authors observing a decrease of μ_{em} upon ES-SB.[70] This also enabled the previous assertion [71] that the fluorescence shift tracks excited-state symmetry breaking to be refuted (Figure **??**). This method has since been used by others [72] to track excited-state symmetry breaking in molecules lacking localised IR marker modes, giving further insight into this phenomenon.

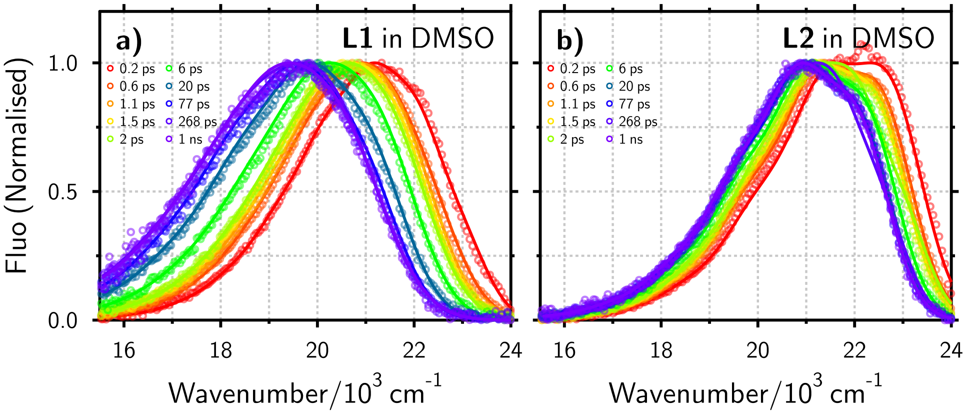


Figure 14. FLUPS data for compared to fitting with equation **??** for **a) L1** (a molecule where ES-SB is anticipated) and for **b) L2** (a molecule where ES-SB is not anticipated) in DMSO . Data and fit peak normalised to ease viewing. Adapted with permission from reference [70]. Copyright 2017 American Chemical Society.

Band-shape analysis has also been used in the study of vibrational energy relaxation. Kasajima *et al.* studied the vibrational relaxation of perylene (**Pe**) and 12-(3-perylenyl)dodecanoic acid (**PD**) (Figure **??**) using fluorescence up-conversion.[73] By taking fluorescence time traces at different wavelengths, and then subtracting the steady-state fluorescence (**1** in Figure **??**) they were able to observe the emission from higher-lying vibrational states populated at early times by the 400 nm excitation pulse. They then decomposed the fluorescence into contributions from $S_1(\nu' = 0,\, 1,\, 2)$ by using a DHO model which takes into account fluorescence from higher-lying vibrational states (equations **??** and **??**) - first defining the Huang-Rhys factor by using the DHO to model the steady-state fluorescence, then using this value to simulate fluorescence from $S_1(\nu' = 1,\, 2)$. Using this, they were able to deduce the contributions of these states to the fluorescence at each time step and thus the vibrational populations, shown in panel 2 of Figure **??**. From this, and the fitting of these populations with a kinetic model, they showed that the relaxation rate of **PD** was higher than **Pe**, attributed to a higher number of vibrational degrees of freedom, and that direct relaxation in the excited state from $|2\rangle \rightarrow |0\rangle$, as well as $|2\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |0\rangle$ were possible (under the assumptions of classical kinetics, which may not necessarily [74] apply to vibrational relaxation).

The **Pe** molecule and the electron donor *N*,*N*-dimethylaniline (**DMA**, Figure **??**) have also been used in conjunction with the spectral decomposition method (Section **??**) in order to study bimolecular electron transfer. In these experiments the decomposition method was used to extract the time evolution of different populations (acceptor, **Pe**, in the S_1 state, ionic product, accceptor in the T_1 state)

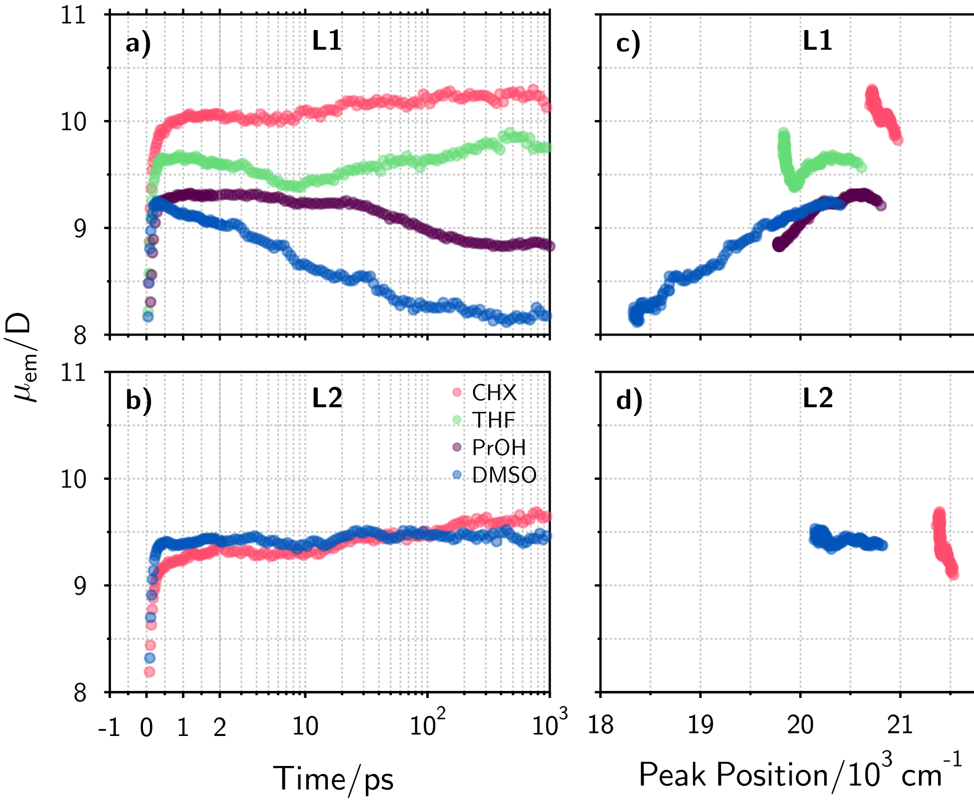


Figure 15. Time dependence of μ_{em} of **a) L1** in cyclohexane (CHX), 1-propanol (PrOH), tetrahydrofuran (THF) and DMSO, and **b)** of **L2** in CHX and DMSO. In **L1**, the μ_{em} decreases as symmetry breaking occurs in PrOH and DMSO. In **L2**, no symmetry breaking occurs in even the most polar DMSO. All traces are normalised to steady state value of μ_{em} at long times. Note the change of axis at 2 ps. μ_{em} vs the peak position of the emission band of **c) L1** in CHX, PrOH, THF and DMSO and **d) L2** in CHX and DMSO. As can be seen, the peak shift does not correlate with the change in the μ_{em} and is thus not a sufficient marker of ES-SB. Adapted with permission from reference [70]. Copyright 2017 American Chemical Society.

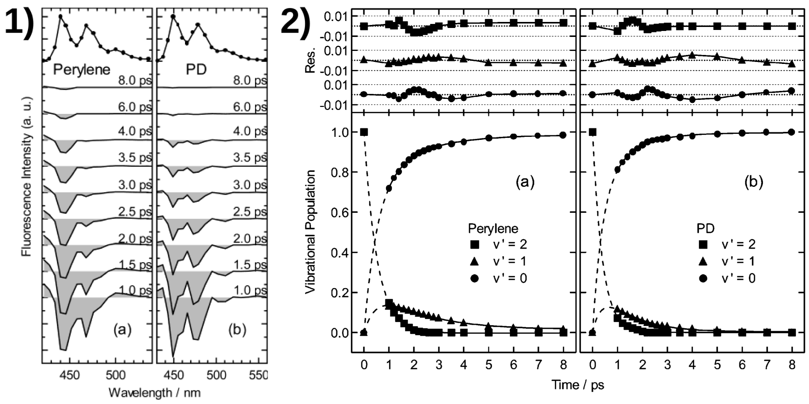


Figure 16. **1)** Time-resolved difference spectra constructed from fluorescence up-conversion measurements of a) **Pe** and **PD** in 2-methyltetrahydrofuran. Steady-state fluorescence spectra shown in top panels. **2)** Vibrational populations of $\nu' = 0$, 1 and 2 in a) **Pe** and b) **PD** as a function of time as deduced from band-shape analysis. Reprinted with permission from reference [73]. Copyright 2004 American Chemical Society.

in both conventional [61] and ionic [75] solvents. In both of these experiments, the populations were compared to modelling using a diffusion reaction equation, which was found to well-describe the reaction dynamics over 5 orders of time, 3 orders of viscosity and at a range of quencher concentrations using a single set of parameters. This was in stark contrast to formal kinetics, which was unable to provide adequate population fits to the data even with a complex system of *ad hoc* states and time-independent rates. This experiment showed the weaknesses of the assumption of the formal kinetics model when the processes are involved are ill-described by well-separated states, and that clear physical insight could be obtained in using a theory that incorporates diffusion. A comparison between the encounter theory results and the formal kinetics results is shown in Figure **??**. This method was also used to understand the same electron transfer reaction in ionic liquids - showing that the charge separation is not significantly affected by the ionic nature of the solvent, but the screening of the Coulomb interaction significantly enhances the generation of free ions.[75] The decomposition method, coupled with the comparison to a solvent of the same viscosity but non-ionic in nature (a DMSO/Glycerol mixture) and the use of a diffusion reaction equation, enabled this insight which is of potential use for a variety of applications where free ions are desired.

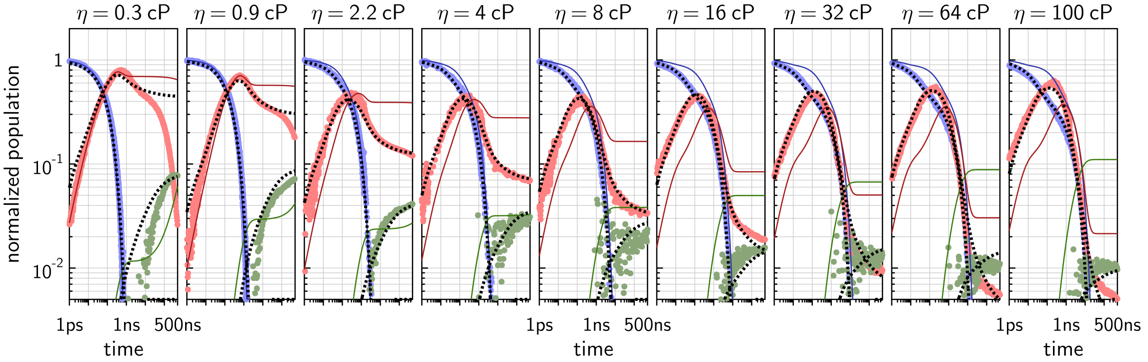


Figure 17. Time evolution of the different populations involved in the photoinduced bimolecular electron transfer reaction between **Pe** and **DMA** in DMSO/Glycerol mixtures of differing viscosity extracted using the spectral decomposition method (blue: **Pe** in the S_1 state, red: **Pe** anion, green: **Pe** in the T_1 state). Coloured lines are a formal kinetic simulation, dashed lines are diffusion reaction simulation). Note the failure of the formal kinetics model. Reproduced from reference [61] with permission from the PCCP Owner Societies.

Band-shape analysis was also used to understand the photophysics of the stable radical 1,3,5-triphenylverdazyl (**Verdazyl**, Figure **??**). By doing a transient absorption experiment on **Verdazyl**, exciting the $D_1 \leftarrow D_0$ transition using 800 nm light, Weinert *et al.* observed that after a c. 500 fs induction period (which they deduce to possibly provide an upper limit on the D_1 lifetime), the transient absorption spectra decay display a time-dependent peak position which shifts to higher energy with time, and that the time traces display non-exponential relaxation back to zero.[76] They thus interpreted this as the signature of a rapid internal conversion back to the ground state, producing a vibrationally hot ground state. Thus, they used the Sulzer-Wieland formalism (equations **??** - **??**) to simulate the steady-state spectrum, then (using those results) the time-dependent spectrum to access the cooling dynamics. By scaling the $D_1 \leftarrow D_0$ transition (found to be necessary due to the presence, deduced using time-resolved anisotropy measurements, of additional transitions that are accessible upon thermal excitation), a good approximation of the time-dependent band shape was simulated for each time step, and thus the temperature at each time step could be accessed and its time behaviour interrogated (the spectra, and the simulated spectra are shown in Figure **??**). Weinert *et al.* found that the cooling could best be described as a biexponential function, and suggested that the shorter time constant could be attributed to intramolecular vibrational redistribution and the longer to intermolecular vibrational relaxation. In this way, they were able to gain a full picture of the relaxation of the **Verdazyl** chromophore upon excitation to its lowest excited state. It should be noted however, that the strict partitioning of VER into intra- and the intermolecular to different time scales in solution has been questioned.[73] Nonetheless, the paper presents a remarkably detailed investigation of the relaxation of the **Verdazyl** chromophore.[76]

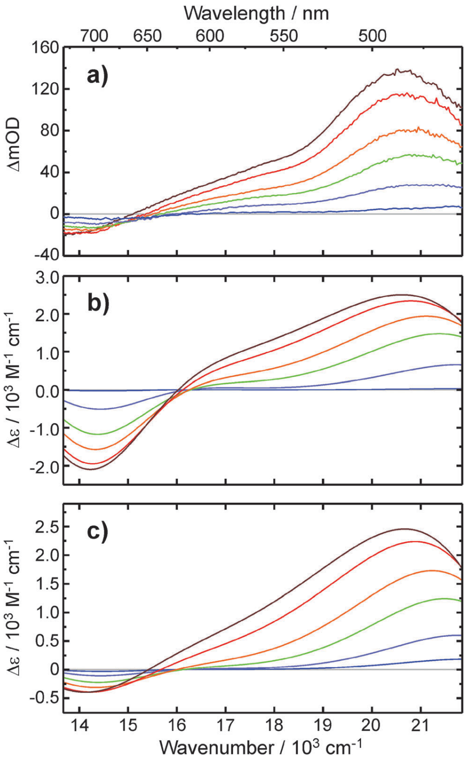


Figure 18. (a) Transient absorption spectra of **Verdazyl** in acetonitrile after 800 nm excitation at time delays specified by colour. (b) Simulation of the spectra using equations **??** - **??**. (c) As in (b) but with the $D_1 \leftarrow D_0$ transition scaled, as was needed for direct comparison. Time delays are 500 fs (brown), 1 ps (red), 2 ps (orange), 3 ps (green), 5 ps (light blue), and 10 ps (dark blue). Reproduced from reference [76] with permission from the PCCP Owner Societies.

2.5. Conclusions

Band-shape analysis methods can provide extremely useful insight into processes where the spectra are anticipated to shift and where the shape of the spectra may directly be related to a physical quantity or process of interest. The monitoring and understanding of solvation and vibrational relaxation are two fields where this is highly relevant, and two fields where band-shape analysis has found much use. More recently the analysis has found application in understanding excited-state symmetry breaking and electron transfer.

The techniques have, historically, found more use in transient fluorescence than in transient absorption, likely due to the complications of non-unipositive data and multiple superimposed band-shapes that transient absorption represents. The recent decomposition method by Angulo *et al.*,[61] as well as the steady-state contribution removal of Kovalenko *et al.*[63] and Wilcken *et al.*[64] does signal progress in this direction. As such, these are useful techniques if one wishes to, without the pre-imposition of a kinetic model, access populations of states or the dynamics of a relaxation process. In this respect, band-shape analysis methods are versatile and useful.

3. Global Kinetic Analysis

Global kinetic models (GKMs) combine the assumptions of classical kinetics with a 'global' view of the data - the simultaneous analysis of all time traces. First, we will discuss both the classical kinetic assumption and the 'global' view of the data, then the common algorithms and common kinetic models. Finally, we will discuss the limitations of these models, and some examples of their use.

The fundamental assumption underlying classical kinetics is that the chemical reaction under consideration can be described as transitions between well-defined states.[3] This leads to the assumption that each chemical transformation can be quatified by a rate constant independent of time. In addition, an assumption that follows is that each well-defined state has a time-invariant electronic spectrum. These assumptions often work well for both the gas phase and solution phase, particularly for processes occurring on timescales much longer than those of vibrational and/or solvent relaxation.[4, 77] For ultrafast processes in liquids that takes place on similar time scale to vibrational relaxation and/or solvent motion, these assumptions can become more problematic, as we will see in this and other sections. For transient spectroscopy, this assumption (along with the common assumption that one may treat the reaction under consideration as a series of exponential steps) results in the comparison between the measured time traces and a sum of exponential decays (which may or may not be convolved with an IRF).

The assumption that underlies a 'global' view of the data is that the data is bilinear. 'Bilinear' is a term that means that the data are separable with respect to two variables - here, wavenumber and time. The assumption specifically is that the data are assumed to be separable in wavenumber and time according to

$$\mathbf{D} = \mathbf{S}(\tilde{\nu}) \cdot \mathbf{C}(t) \tag{37}$$

where $\mathbf{C}(t)$ is a k by n matrix and $\mathbf{S}(\tilde{\nu})$ is a m by k matrix (hereafter the (t) and $(\tilde{\nu})$ are assumed implicit). The number of time points in the experiment is n, m the number of wavenumber points and k is the number spectra and time traces needed to describe the data, $\mathbf{C}$ and $\mathbf{S}$ thus being *concentration* and *spectral* matrices respectively. k should be the smallest number of components needed to fully describe the data above the noise level. Thus, to reiterate, the key assumption is that one may separate the spectral and temporal behaviour of the data (i.e. we may assume bilinearity). This ties into the assumption of classical kinetics - we have well-defined species and thus time-invariant spectra of these species. The assumption of data bilinearity will (obviously) break down when these assumptions are invalid, and thus we see wavenumber-dependent kinetics. In transient spectroscopy, this can be observed when solvent relaxation and/or vibrational cooling occur on the same timescales as the population kinetics, which is particularly common for transient spectroscopy data with sub-ps time resolution. Though various authors only discuss that this is the case for fast population kinetics,[31, 78] the breaking down of the assumption of bilinearity is not restricted to this case - it will occur when vibrational relaxation and solvation compete with the population kinetics. As an example, if one measures a photochemical or photophysical process in high viscosity media such as ionic liquids, the solvation time can reach close to 1 ns.[79] Thus, as has been discussed in more detail previously,[80] any application of global kinetic models to processes involving vibrational relaxation and/or solvation, and interpretation of any extracted time constants, should be done with extreme care. It should also be noted that equation 37 represents the same type of Soft-Modelling decomposition that will be discussed in Section 5 but the difference to methods discussed in that Section being that here we will impose a functional form on the time behaviour.

3.1. Theory

3.1.1. Motivation

What is the motivation to 'go global'? One might say it would be easier to fit time traces at the maxima of various spectral bands and see what time constants one finds in common. Indeed, this is a perfectly reasonable first step in data analysis of transient spectroscopy data. It may quickly show if the kinetics are strongly wavenumber-dependent, and thus if it is worth attempting a global analysis. However, there are clear mathematical limitations to this in terms of the number and proximity of the lifetime components that we may resolve. In an elegant paper discussing exponential analysis of physical phenomena, Istratov and Vyvenko show that even for a signal-to-noise ratio (SNR) of 10^4, the smallest ratio of time constants that one may distinguish in a biexponential fit is 1.2.[81] To give some context, a 'typical' SNR for data taken in our laboratory for step-scan TA is 40 [82] (rapid-scan TA will natrually reach a much higher SNR [15, 83]) and for FLUPS is 90.[70] Thus for these lower SNRs the ratio of time constants that one can distinguish is much larger. Thus, if one simply analyses one time trace and the 'real' kinetics have two time constants that are rather close together in value, one will not be able to distinguish them easily.

We may test this idea using simulations (further details of which are described in Appendix A) where we repeatedly (with a re-drawn noise each time) simulate fitting a biexponential (i.e. N=2 in equation 3) curve with various ratios between the two lifetimes. We may then get an error on the recovered parameters, calculated as before (equation 4) and observe how well the parameters are recovered. The results, as well as some example fits, are shown in Figure 19. What Figure 19 demonstrates is that, even for very high SNRs, as the ratio of the time constants decreases, the error on the parameters increases to levels such that no confidence in the recovered parameters may be had. In fact, the error on the longer time constant rises to above 1000% below a ratio of 1.25 for SNRs below 1000.

In addition, the Istratov/Vyvenko paper provides a thorough discussion of the limits of distinguishability of a double or triple exponential fit, and provides an exceptionally useful discussion of the various algorithms that have historically been used to analyse these data.[81]

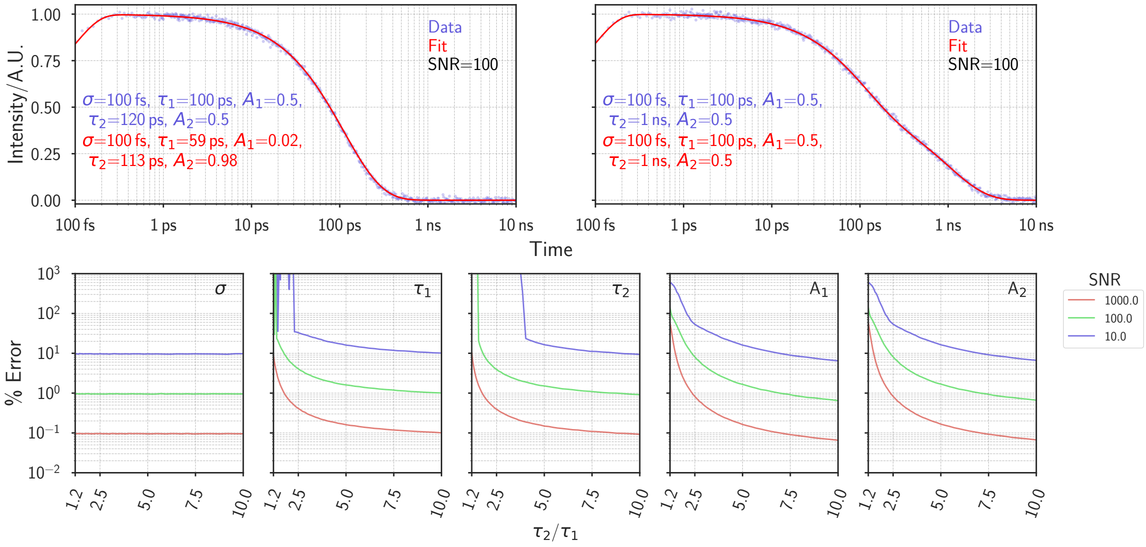


Figure 19. Upper panels: example fits (red) to data (blue) generated in these simulations. The examples are of a less accurate fit where the $\tau_1 \simeq \tau_2$ (left panel) and an accurate fit where $\tau_1 \ll \tau_2$ (right panel). Lower panels: Percentage error on recovered parameters from fitting the equation 3 with N=2 to data simulated using the same equation. Data points are averaged accuracy of 10,000 fits. $A_1 = A_2 = 0.5$, $\tau_1 = 100\,\text{ps}$, $\sigma = 100\,\text{fs}$. Further details on all simulations may be found in appendix A.

Fortunately, there is a combined experimental and analytical workaround to the problem shown in Figure 19 - analysing multiple traces simultaneously. This was demonstrated by Beechem *et al.*[84–86] in an instructive example that will be recapitulated here. The example shows that by simultaneously analysing more than one trace where one may assume that the time constants are the same (or share a known relationship) the error surface one optimises over becomes much more defined and thus the time constants are more easily distinguished.

One typically optimises with respect to a χ^2_ν metric, a goodness-of-fit metric that is defined as

$$\chi^2_\nu = \frac{1}{\nu}\sum_{i=1}^{N}\frac{[D(t_i) - D_{fit}(t_i)]^2}{\sigma_i^2} \tag{38}$$

where the data is $D(t)$, $D_{fit}(t)$ is a fitting function, and σ_i is the error on data point i. ν is the degree of freedom

$$\nu = n_{data} - p_{fit}. \tag{39}$$

where n_{data} is the number of data points and p_{fit} is the number of fitted parameters. A $\chi^2_\nu \gg 1$ indicates a poor model fit, $\chi^2_\nu > 1$ indicates the fit has not fully captured the data and in principle $\chi^2_\nu = 1$ indicates a fit to within experimental uncertainties.

It is possible to simulate such a χ^2_ν surface for the analysis of a single biexponential decay, where the decay is described by equation 3 and the parameters contained in

Table 1. The time grid is linear from -2 to 50 ns, with 8192 data points describing the decay, 'typical' of a time-resolved fluorescence experiment performed using the time-correlated single photon counting (TCSPC) technique.[23]

Trace	τ_1/ns	A_1	τ_2/ns	A_2	σ/ns
1	4	0.25	5	0.75	0.1

Table 1. Parameters of equation 3 used to simulate trace for Figure 20 **a)**.

The surface (Figure 20**a**) shows several interesting features. First, it is very flat around the minimum region around the parameter space of 4 and 5 ns (dark blue areas correspond to χ^2_ν values ≤ 1.2, values where one would typically consider a fit reasonable). Thus, even if one's algorithm successfully finds this minimum, the error analysis will reveal that one may have no confidence in the parameters recovered. The alleys along each parameter axis also indicate that the one lifetime value may wander far from the true value and the χ^2_ν value may be almost completely compensated by the other lifetime's amplitude. Additionally, there is a point along the principal diagonal where $\tau_1 = \tau_2$, i.e. the analysis of the biexponential decay converges to a monoexponential fit with a τ of 4.25 (the weighted mean of the two lifetimes). The decrease in χ^2_ν in proceeding from this monoexponential fit to the biexponential one is relatively minor (<10%) and thus the justification to attempt a biexponential fit (to a curve that is biexponential by construction) is statistically marginal.

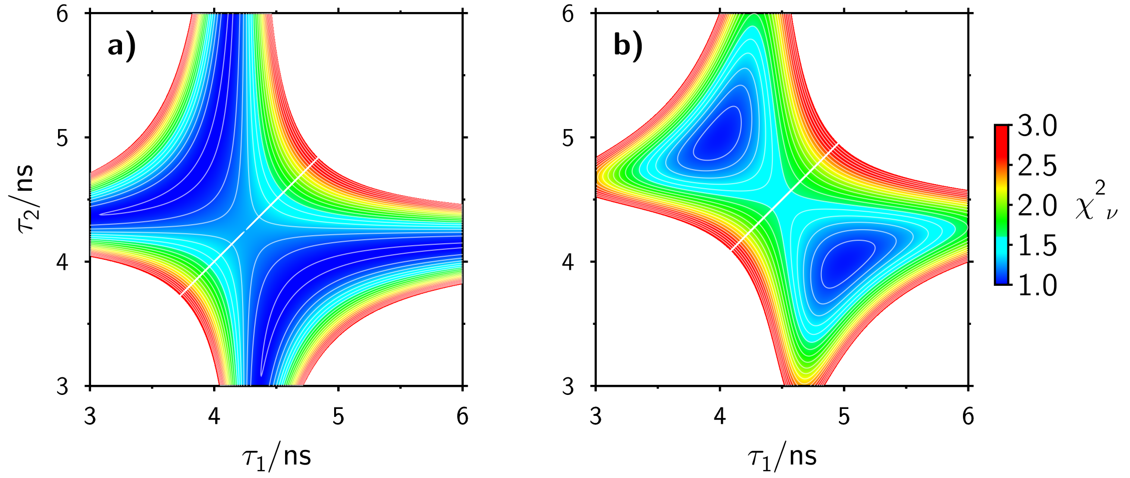


Figure 20. Surface constructed [86] from χ^2_ν values for fitting a biexponential model to a single trace (**a**) or two traces (**b**) simultaneously. White areas outside of coloured surface plot correspond to χ^2_ν values > 3. In plot **a**, the model simulated corresponds to equation 3 with the parameters of Table 1, whereas in plot **b**, to the parameters contained in Table 2. NB that dark blue areas correspond to a $\chi^2_\nu \leq 1.2$. The SNR for these simulations was 1000.

If however, one is fortunate enough to have more than one decay trace and is able to assume that the time constants are identical (or have some known relationship), it is possible to analyse them simultaneously and this combination of multiple traces results in an error surface with a considerably better defined minimum. If one examines the error surface obtained upon the simultaneous analysis of the two decays, both described by equation 3 and simulated using the parameters contained in Table 2, it is clear from Figure 20**b** that the error surface minimum is better defined and near the correct values of 4 and 5 ns.

Trace	τ_1/ns	A_1	τ_2/ns	A_2	σ/ns
1	4	0.25	5	0.75	0.1
2	4	0.75	5	0.25	

Table 2. Parameters of equation 3 used to simulate traces for Figure 20 **b**).

The covariance valleys (i.e. the large areas where we may change the lifetime by a large amount with very little difference to the χ^2_ν parameter) along each parameter axis have been almost completely eliminated, and thus the justification of the use of a biexponential is now clear. This considerable improvement arises from the fact that multiple experiments may be viewed as sets of rotated error surfaces, and that in the global analysis the union of these surfaces is minimised. The larger the degree of variation in the amplitudes is, the larger the rotation between the individual error surfaces, and the greater the ease of minimisation to the correct answer.

Here, we have extended the simulations of Beechem *et al.* to also show the effect of noise on this χ^2_ν surface (note the range of lifetimes that define the surface has been increased). This is shown in Figure 21, where we may clearly see that as the noise level increases the area where the $\chi^2_\nu \leq 1.2$ may occupy considerable parameter space (in the case of an SNR of 10, almost the whole surface). Thus whilst the simultaneous analysis of two traces does lead to a more defined error surface, enabling us to find a more statistically meaningful minimum, it is still imperative to (if one wishes to report parameters with confidence) explore the error surface and to know the experimental noise level.

In addition, an exploration of how the SNR effects the error on the recovered parameters is shown in Figure 22 as a function of the ratio between the two lifetimes. It is clear that the error on the parameters is considerably less than for the single trace experiments shown in Figure 19 in all cases. Thus the surfaces (shown in Figures 20 and 21) and parameter error simulations (shown in figures 19 and 22) clearly highlight one of our rationales to ‘go global’ - we will be able to find more statistically meaningful minima and more accurately extract parameters from our data.

Another reason for using global analysis is that the speed of fitting is considerably increased - if one individually fits n_t traces with n_{comp} time constants (thus n_{comp} amplitudes), one has $2(n_t{\cdot}n_{comp})$ free parameters to optimise. If one simply links the time constants the number of free parameters is lessened to $(n_t{\cdot}n_{comp}) + n_{comp}$. This can produce a formidable increase in efficiency - Beechem *et al.*[85] discuss an

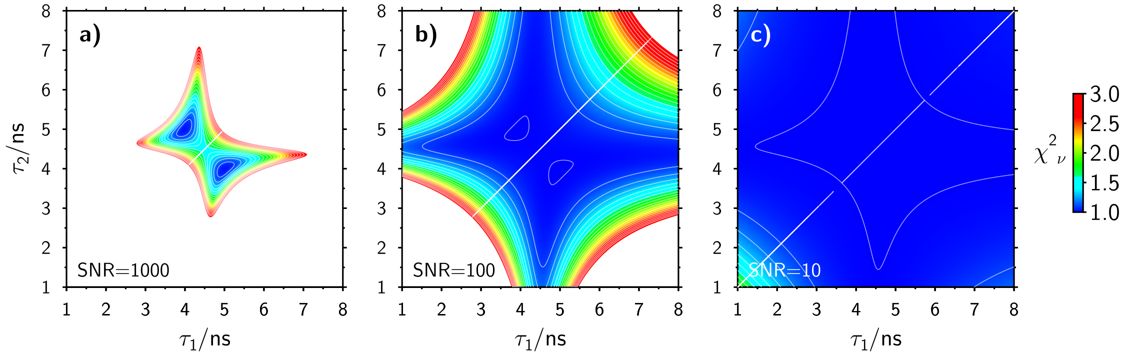


Figure 21. Surface constructed [86] from χ^2_ν values for fitting a biexponential model to two traces simultaneously with different levels of relative noise added. Specifically, **a)** corresponds to an SNR of 1000, **b)** to an SNR of 100 and **c)** to an SNR of 10. White areas outside of coloured surface plot correspond to χ^2_ν values > 3. The model simulated corresponds to equation 3 with the parameters of Table 2. NB that dark blue areas correspond to a $\chi^2_\nu \leq 1.2$.

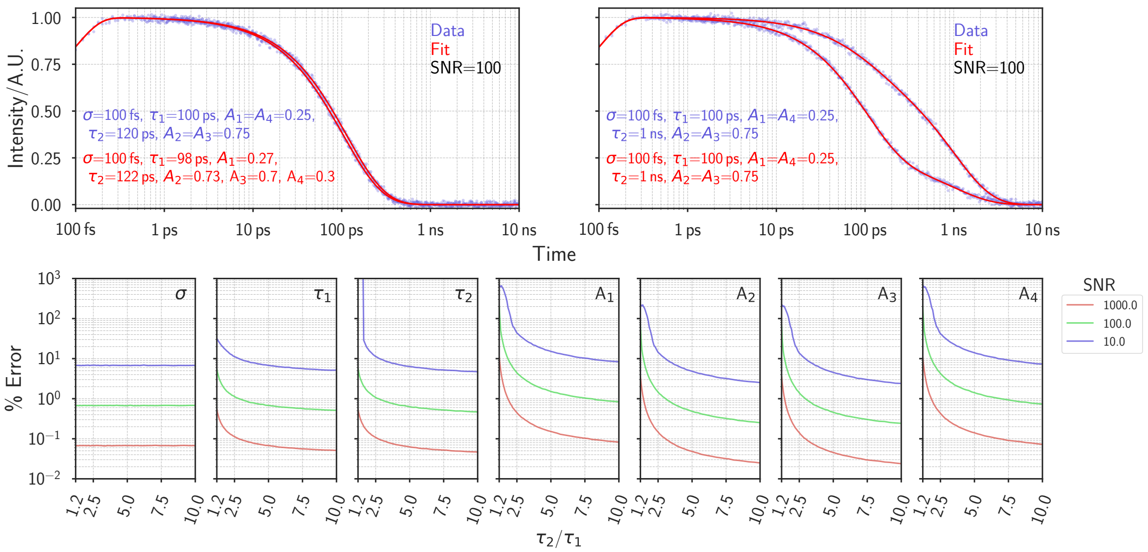


Figure 22. Upper panels: example fits (red) to data (blue) generated in these simulations. The examples are of a less accurate fit where the $\tau_1 \simeq \tau_2$ (left panel) and an accurate fit where $\tau_1 \ll \tau_2$ (right panel). Lower panels: Percentage error on recovered parameters from fitting the equation 3 and linked time constants to data simulated using the same equation. Data points are averaged accuracy of 10,000 fits. $\tau_1 = 100\,\text{ps}$, $\sigma = 100\,\text{fs}$, A_1=0.25, A_2=0.75, A_3=0.75, A_4=0.25. Further details on all simulations may be found in Appendix A.

example of a multi-wavelength fluorescence anisotropy experiment where the single-wavelength analysis took months to perform, whereas the globally linked optimisation simply took a matter of hours (one may safely assume that these timescales would be considerably shortened today - nonetheless, the increase in efficiency is striking). Recasting of the problem using matrices enables the number of free parameters in the optimisation to be only n_{comp} free parameters, increasing the speed much further.

3.1.2. Matrix Method of Global Kinetic Analysis

The final rationale behind global analysis is an outcome that simplifies the assignment of the time constants to associated spectra. When one links the time constants across all wavenumbers in the experiment, the spectral matrix in equation 37 will contain the wavenumber-dependence of the amplitude of the time constant in question. It is, in general, considerably easier to assign a contribution by looking at a spectrum than by looking at the time constant alone. If one simply takes the spectral shape associated with the amplitudes in **S** from equation 37, these are referred to as the decay associated spectra (**DAS**),[78] or **DADS**, where the second **D** stands for difference in the case of TA measurements, as they are directly associated with the decay profiles in **C** from equation 37.[78] The fact that they are directly associated to the decay profiles enables one to remove these as fitting parameters - one simply guesses the time constants, then uses these to compute the **DAS** (from which the calculation of the fit matrix $\mathbf{D}_{\mathrm{fit}}$ follows). For guessing the **DAS**, one guesses the **C** matrix and then computes the **S** matrix via:

$$\hat{\mathbf{S}} = \mathbf{D}(\mathbf{C}^{\mathrm{T}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}} \tag{40}$$

which may be simplified to

$$\hat{\mathbf{S}} = \mathbf{D}{\cdot}\mathbf{C}^{+}. \tag{41}$$

$\mathbf{A}^{+}$ is the Moore-Penrose pseudoinverse,

$$\mathbf{A}^{+} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}, \tag{42}$$

of the matrix **A**, and is used due to it being both rapidly and stably able to compute the best-fit solution to a system of linear equations. The circumflex above **S** denotes ‘estimator of’.[78]

The fit matrix may then be computed by

$$\mathbf{D}_{\mathrm{fit}} = \mathbf{S}{\cdot}\mathbf{C}(\theta). \tag{43}$$

i.e.

$$\mathbf{D}_{\mathrm{fit}} = \mathbf{D}{\cdot}\mathbf{C}^{+}{\cdot}\mathbf{C}(\theta) \tag{44}$$

Where **S** may be referred to as conditionally linear [87] parameters, as they are solved for conditionally on the unknown parameters θ, using equation 40. This method of computation of the fit matrix and subsequent fitting is referred to informally as the ‘matrix method’ of global kinetic analysis (GKA). The model for the **C** matrix is dependent upon θ in a non-linear fashion, thus iterative numerical methods must be used to solve the problem. It should also be noted that this method **works perfectly for single traces**, as these are formally matrices with one of the axes simply being of length one (thus one may also use this method to speed up fitting of single traces). The optimal algorithms for the solution of this problem will not be discussed here, and are well described by van Stokkum and co-workers, who have done much to profligate Global Kinetic Analysis in the transient spectroscopy field - the review in *Biochimica et Biophysica Acta* is exceptionally thorough in the discussion of the mathematics, implementation and usage of Global Kinetic Models.[78, 87, 88]

The species associated spectra or **SAS** (sometimes referred to as **SADS** in TA spectroscopy, where the second **D** again stands for difference) are spectra that are said to reflect the exact spectra of the specific species. Thus the **DAS** are equivalent to these only if the **C** matrix has been constructed using a kinetic equation that explicitly takes into account the assumed decay pathways - these can be solved using matrix methods, as has been well described by Berberan-Santos and Martinho.[89] If one however simply uses the most simple sum-of-exponentials approach, a shortcut exists if one wishes to convert between the **DAS** and something that many use as an 'approximation' to the **SAS** - the evolution associated spectra (**EAS** or **EADS**). Evolution associated spectra describe how time-resolved spectra evolve over time, but can be assigned to species or states only under the assumptions of classical kinetics and bilinearity. Under the same reasoning, the associated time constants will not always correspond to a well-defined process. However, the determination of **EAS** can be useful for visualizing the dominant spectral changes and their associated timescales, and thus be useful in the selection of a model. The **SAS** on the other hand correspond to genuine species spectra (if the kinetic model used is correct). The **DAS** correspond to the **SAS** if and only if the kinetic process under investigation is that of n_{comp} parallel decay components - i.e. for a ground state A and 2 photoexcited components, B and C, the model one would have in mind is shown in Figure 23.

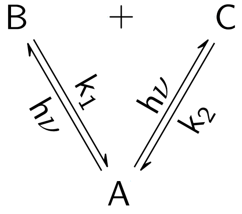


Figure 23. Parallel decay scheme.

This is a model that is only likely for mixtures of molecules which do not interact in any way. The spectra associated with this model are easily transformed to the **SAS** of the model of a sequential decay, i.e. for the scheme in Figure 23, the corresponding scheme of a sequential decay is shown in Figure 24.

The results of transforming the spectra assuming a sequential decay model are only valid to refer to as the **SAS** if the sequential model is indeed appropriate. Otherwise the assumption of a sequential model results in the recovery of the **EAS** from the **DAS**. The **DAS** are transformed to the **EAS** by the use of the so-called **B** matrix, a lower triangular matrix which expresses that the i-th **EAS** is a linear combination of the i-th and following **DAS**. The **B** matrix is constructed using [78]:

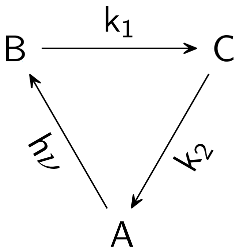


Figure 24. Sequential decay scheme.

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

where $b_{l1} = 1$, $j \leq l$, k_i is the rate constant associated with spectral component i. The **EAS** are then constructed from the **DAS** using:

$$\mathbf{EAS} = \mathbf{DAS} \cdot \mathbf{B}. \tag{46}$$

The Matrix Method of Global Kinetic Analysis is able to recover time constants that are close together in ratio even at relatively low SNR, as well as being remarkably fast to compute. Its performance is encapsulated in Figure 25, and as can be seen it is exceptionally accurate for almost all SNRs, as well as very accurately recovering the **EAS** used in simulating the data.

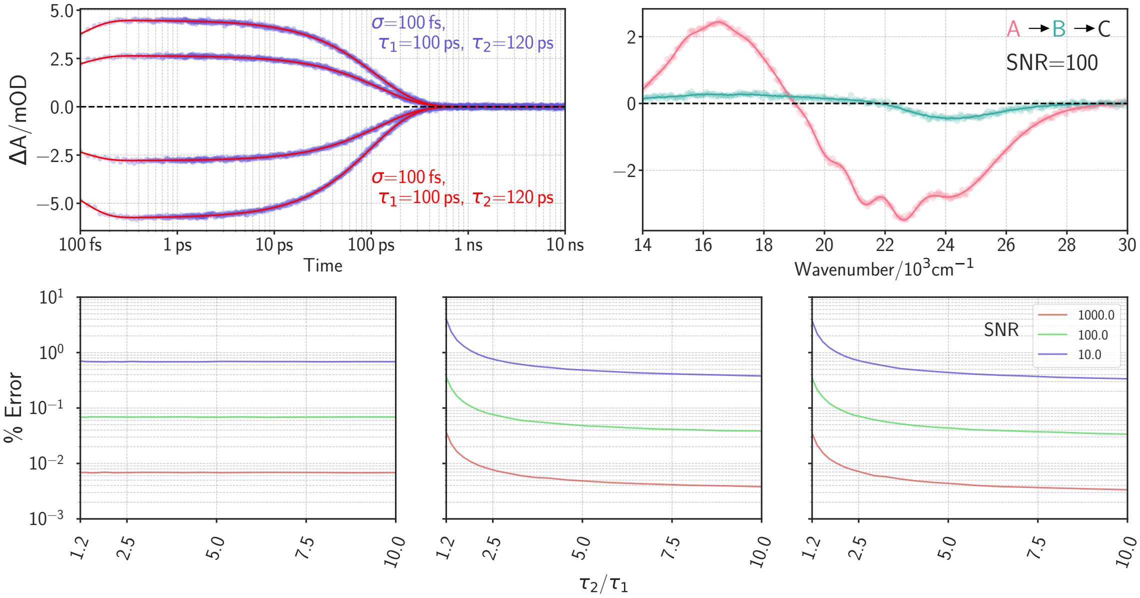


Figure 25. Upper panels: example fits (solid lines) to data (coloured points) generated in these simulations. The example is of an accurate fit where the $\tau_1 \simeq \tau_2$ (time traces in left panel, **EAS** in right panel). Lower panels: Percentage error on recovered parameters from fitting the equation 43 to data simulated using the same equation. Data points are averaged accuracy of 10,000 fits. $\tau_1 = 100\,\text{ps}$, $\sigma = 100\,\text{fs}$. Further details on all simulations may be found in Appendix A.

3.1.3. Wavenumber-dependent method of Global Kinetic Analysis

As was pointed out by Fita *et al.*,[90] the Matrix Method of fitting the transient spectroscopy data with IRF deconvolution requires that there is no wavenumber-dependence of the IRF width, which is not a valid assumption for many experiments, as the assumption is that the number of components in the **C** matrix is the same as the number of independent time traces that may be multiplied by the spectral matrix **S** to construct the dataset. Multiple methods have been developed to circumvent this problem. The simplest is to truncate the data to after the IRF at all wavenumbers, then analyse the resulting matrix. This does preclude the analysis of components comparable to one's time resolution however, and thus may lead to a loss of information.

The method used by Fita *et al.*[90] was to use Tikhonov regularisation [91] (discussed in relation to lifetime density methods in Section 4) to deconvolve the time traces and then fit them using the Matrix Method. This is in some ways not an ideal strategy - the Tikhonov regularisation deconvolves the data and returns a 'noise free' matrix - thus one loses the statistical information on the measurement if one fits this returned matrix, and so an error in the Tikhonov regularisation could cause a large and unknown error in the Global Kinetic Analysis.

A superior method is to implement an analysis that is able to take into account the wavenumber dependence of the IRF. This was discussed glancingly by Mullen and van Stokkum [92] as well somewhat more thoroughly by Slavov *et al.*,[93] though Slavov *et al.*'s primary goal was to use a method to fit data without first chirp correcting the data matrix. The paper itself is, from a programming perspective, scant on actual detail, simply stating that the construction of the fit matrix occurs in the same way as by the Matrix Method used by van Stokkum, Fita and others.[78, 90] This is untrue if one wishes to implement a wavenumber dependence of the IRF or the time zero position (i.e. the chirp). In fact, one no longer has a concentration guess matrix **C**, but a concentration guess tensor, $\boldsymbol{\mathsf{C}}$ which has dimensions of m by n by n_{comp}, where m is number of wavenumbers, n is number of time points and n_{comp} is number of fit components. Thus (if one simply is fitting with a multiexponential model),

$$\boldsymbol{\mathsf{C}}(t,\tilde{\nu},n_{comp}) = \frac{1}{2}\exp\{-\tau^{-1}(t-t_0(\tilde{\nu}))\}\exp\{0.5\cdot\tau^{-2}\cdot\sigma(\tilde{\nu})^2\}\left[1+\mathrm{erf}\left(\frac{t-t_0(\tilde{\nu})-\cdot\tau^{-1}\cdot\sigma(\tilde{\nu})^2}{\sqrt{2}\sigma(\tilde{\nu})}\right)\right] \tag{47}$$

where $t_0(\tilde{\nu})$ and $\sigma(\tilde{\nu})$ are wavenumber dependent time-zero and IRF width parameters (i.e. this tensor is generated iteratively over the wavenumbers). A spectral amplitude matrix **S** is then constructed by the use of the Moore-Penrose pseudoinverse as before:

$$S(\tilde{\nu}) = \mathbf{D}\cdot\boldsymbol{\mathsf{C}}^{+} \tag{48}$$

Then, the fit matrix is generated by the summation over the n_{comp} axis of the tensor:

$$\mathbf{D_{fit}} = \sum_{i=1}^{n_{comp}} \mathbf{SC}(t, \tilde{\nu}, i) \tag{49}$$

and calculation of the **EAS** may proceed in the same manner as in equation 46. Clearly, as the construction of the tensors is a wavenumber-dependent process, the method is considerably slower than the standard Matrix Method. Thus, if one has sufficiently thin sample cells or sufficiently slow kinetics that one does not need a wavenumber-dependent IRF or does not need to fit the IRF (respectively) the simpler Matrix implementation of equation 43 is preferred.

3.2. Published Implementations

Multiple groups have published programs and packages for the transient spectroscopy community's use. Wilderen *et al.* [94] and Slavov *et al.* [93] have both published MATLAB based implementations, largely differing in the support for multi-pulse experiments (the Wilderen *et al.* package) and wavelength-dependent analysis (the Slavov *et al.* package). The Slavov *et al.* package also supports Lifetime Density Analysis, which will be discussed in Section 4. The Wilderen *et al.* package has been used in the study of gold and silver clusters, as well in the study of a variety of biomolecules by a number of groups.[95–99] The Slavov *et al.* package has also been used for analysis of transient spectroscopy of biomolecules and photoswitches.[100]

The group of van Stokkum, as well as publishing extensively detailed papers on the theory [78] and algorithms [87, 88] of Global Kinetic Analysis, has published the R-based software package TIMP [92], as well as the Java-based GUI for said package, Glotaran [101]. This package is the most widely used by the transient spectroscopy community, with thus far (according to Google Scholar in early 2020) 593 citations to its name.

3.3. Applications

Global Kinetic Analysis has been applied to an extremely wide variety of problems in transient spectroscopy. As such, given the goal of this literature review is to provide a broad overview, not all uses of GKA have been elucidated here. Given the number of uses (the van Stokkum review on Global Analysis has 1135 citations in early 2020 according to Google Scholar), this would likely be impossible. This section simply gives a few, significant examples to demonstrate the successes and failures of the technique.

The first applications where GKA made a significant mark were largely in transient fluorescence experiments with ~ 100s of ps time resolution. The various molecules used in these experiments are shown in Figure 26. Thus, the assumption of bilinearity in equation 37 holds for all of these experiments, given that vibrational relaxation and solvation have sub-ps timescales in the solvents used. GKA quickly showed its use in enabling researchers to disentangle complex mixtures of fluorophores. In an experiment where a mixture of fluorophores was prepared in ethanol, the method was able to accurately recover the lifetimes and spectra of **POPOP** (5-Phenyl-2-[4-(5-phenyl-1,3-oxazol-2-yl)phenyl]-1,3-oxazole, τ_{fl} of 1.25 ns), anthracene (τ_{fl} of 4.27 ns), and 9,10-diphenylanthracene (τ_{fl} of 6.21 ns), even in the regions where the anthracene spectrum is very faint.[102] In addition, this study showed that the algorithm was able to accurately describe the excited-state proton-transfer of a pH 3 study of 2-naphthol, using the kinetic scheme in Figure 27.

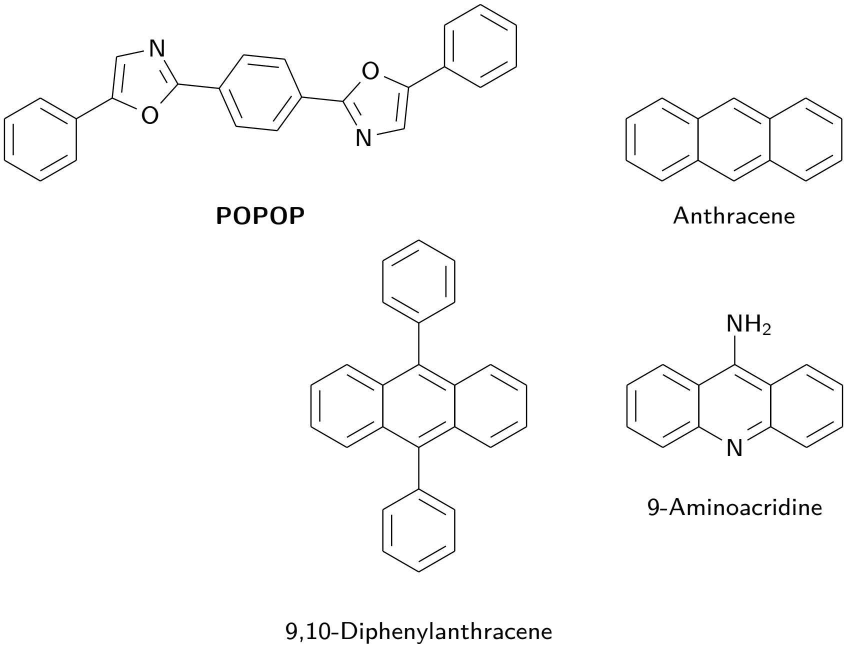


Figure 26. Fluorophores used to demonstrate the versatility of global analysis in time-resolved fluorescence.

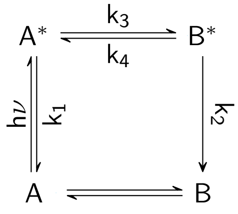


Figure 27. Parallel decay scheme.

The GKA well-describes the reaction, providing well-separated **SAS** (the small ‘hump’ at 355 nm was ascribed to Raman scatter) that sum to the steady-state fluorescence spectrum (Figure 28), with rate constants that agreed reasonably well with the predicted pH-dependence of the rate constant k_B from a previous study on the same molecule at different pH. Thus, here, GKA was able to show the reaction to be well described by the proposed kinetic scheme and to provide good estimates for the fluorescence spectra of the individual species.

GKA was also used to greatly speed up and improve the analysis of fluorescence anisotropy experiments. Beechem *et al.* pointed out that Global Kinetic Methods

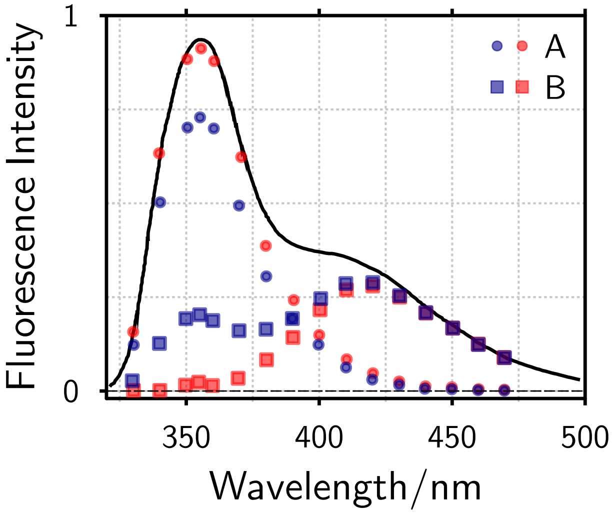


Figure 28. Analysis of the excited-state proton transfer of 2-napthol. 16 wavelengths were used in the GKA. The **SAS** assuming the values of k_B $(k_2 + k_4)$ and k_A $(k_1 + k_3)$ extracted (red symbols symbols) and for the lowest possible value of k_B (blue symbols) illustrating that this does not well describe the data. The black line is the steady-state spectrum. Replotted with permission from reference [102]. Copyright 1986 American Chemical Society.

were able to discriminate between the anisotropic and isotropic rotor model for 9-aminoacridine (Figure 26) - analysis in a single-wavelength manner enabled a perfect description in terms of a simple isotropic rotor, which a global analysis revealed to be inaccurate - the molecule is best described as an anisotropic rotor, but as the time constants only differ by c. 20%, single-wavelength analysis is unable to discriminate, whereas Global Kinetic methods justified the use of the anisotropic rotor model statistically.[85] The method was also used in the study of the binding of epidermal growth factor to its receptor, and its simultaneous use on the stopped-flow fluroescence anisotropy data at multiple different concentrations allowed the unambiguous resolution to the question of if there were multiple affinity classes in the receptor populations (there were), which the non-global studies did not have enough statistical information to properly interrogate.[103]

GKA has also been shown to be very effective for ns time resolution transient absorption experiments - as an example, the complete cycle of the motor inversion of the Hemithioindigo 1 (**HTI1**) from **A** to **D** shown in Figure 29 was studied by the Wilcken *et al.* using transient absorption experiments from the sub-ps to the 18 ms time regime.[64] The combination of this data and the use of global analysis enabled the disentanglement of the triplet excited state of **A** and the product **B** (a product that had not been directly observed), as well as the observation of product **D** (equally, never before directly observed). This enabled a complete picture of the rotation scheme of **1**, as shown in Figure 30. In addition, the global spectra were able to be 'filled in' with the correct amount of ground state bleach to give the 'pure' absorption spectra of **B** and **D**. The deduction of the full motor cycle (using the GKA) also enabled the authors to provide stronger conclusions about how to improve molecular motors - using a simplified rate model they were able to deduce that the most logical parameters to concentrate upon to increase the 'easily accessible' rotation rate (i.e. the rotation rate under relatively undemanding ambient light) would be to increase the quantum yields of the inversion steps (i.e. **A** to **B**, **B** to **C**) and to reduce the thermal barriers, rather than to reducing the parasitic triplet processes. Here, GKA enabled physical insight, a mechanistic picture to be developed and to reveal clear and improved design principles for future chemistry.

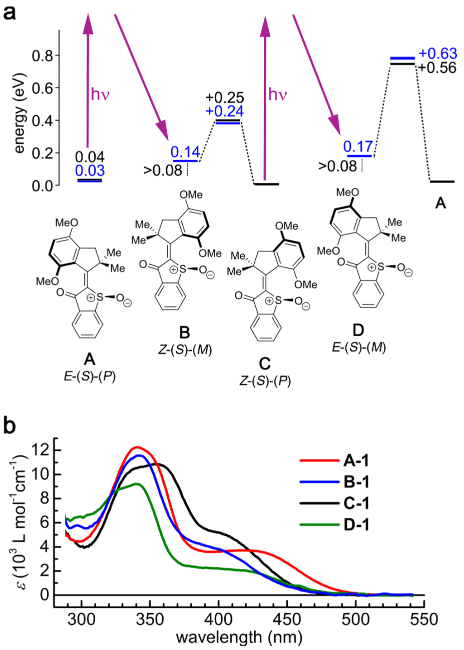


Figure 29. (a) Comparison of the theoretical (blue) and experimentally determined (black) ground state energy profile of **HTI1**. Only (S)-enantiomers are shown. (b) Molar absorption coefficients of all four isomers A–D of **HTI1**. The spectra of the B and D isomers are extracted from the transient absorption measurements. Reprinted with permission from reference [64]. Copyright 2018 American Chemical Society.

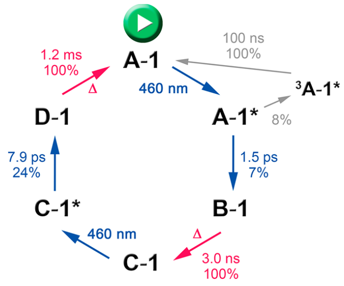


Figure 30. Mechanism of **HTI1** motor operation under irradiation conditions at ambient temperature. All measurements were conducted at 22 °C in dichloromethane. Processes involving excited states are shown in blue or gray arrows; thermal processes in the ground state are shown with pink arrows. Reprinted with permission from reference [64]. Copyright 2018 American Chemical Society.

GKA is less reliable for experiments where effects such as vibrational relaxation and/or solvation cause the data to deviate from bilinearity. These conditions most often manifest themselves on very short timescales and thus interpretation of transient experiments with ∼ 10s-100s of femtoseconds time resolution requires more care and can lead to considerably more errors. In their paper on the excited-state behaviour of the Schiff bases **HNAQ** and **HNAN** (Figure 31), Fita *et al.*[90] thus compared their single-wavelength fits at 5 wavelengths in their experiment to the time constants shown by the global analysis. As they found high coincidence, they deduced that the global analysis technique was acceptable to use in this case, though they noted that they did not analyse the **DADS** of **HNAQ** as nonexponential processes distorted their shape severely.[104] They were able to interpret the various decay processes (Figure 33) involved using said spectral shapes (Figure 32) to attempt to describe fully the excited-state proton-transfer and other processes of the Schiff bases. Due to the 'dip' in the 4 ps **DAS** coinciding with the shape of the fluorescence spectrum, this **DAS** was assigned to the decay of the relaxed keto S_1 state. The ∼ 20-30 fs component was assigned as the initial proton transfer and the 0.5 ps to the vibrational relaxation by comparison with the **HNAQ** time constants, though it seems more likely (given the short nature of the time constants) that these are occurring simultaneously and in a nonexponential manner, and thus the separation of the processes described here is simply due to the stepwise model imposed. The 13 ps component is thus assigned to the relaxation back to the enol form, and the residual to a photoproduct, speculated to be the E-enol form. As pointed out, the nonexponential nature of vibrational relaxation [74] and the short time constants make assignment speculative at best but here GKA was able to at least identify the fluorescent state of the **HNAN** and its photochromism, and thus enabled a clearer photophysical picture of the molecule.

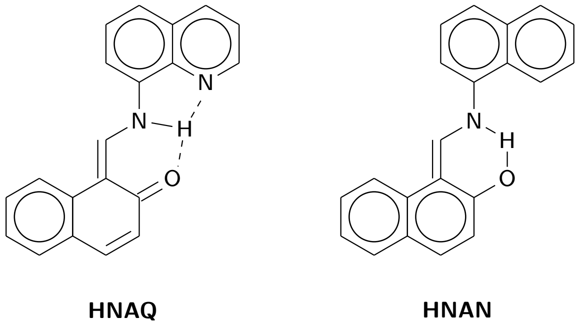


Figure 31. **HNAQ** and **HNAN**.

One area where global kinetic analysis finds an at first glance surprising use is in the analysis of wavepackets arising from the coherent excitation of vibrations. There are two ways in which it has been used for this purpose.

The older usage of GKA is in essence a 'sophisticated subtraction technique' - one uses enough exponentials until the residuals contain only oscillatory components and background noise.[105] These residual maps can then be Fourier transformed to reveal the power spectrum of the excited vibrations. A noteable use of this was in

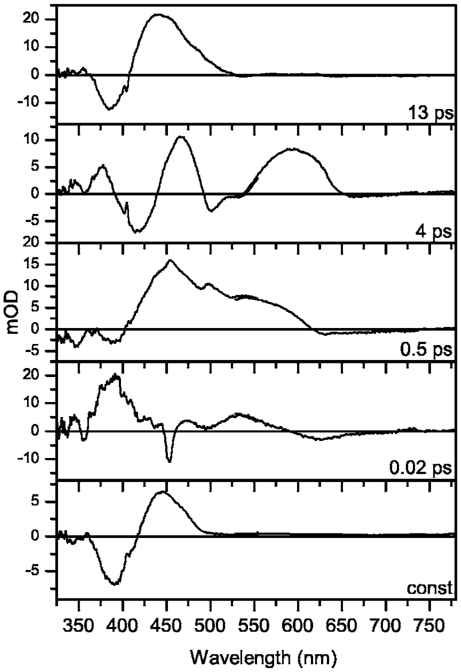


Figure 32. Decay associated difference spectra (**DADS**) of **HNAN**, excited at 400 nm, after deconvolution and GKA. Reprinted with permission from Ref [90]. Copyright 2006 American Institute of Physics.

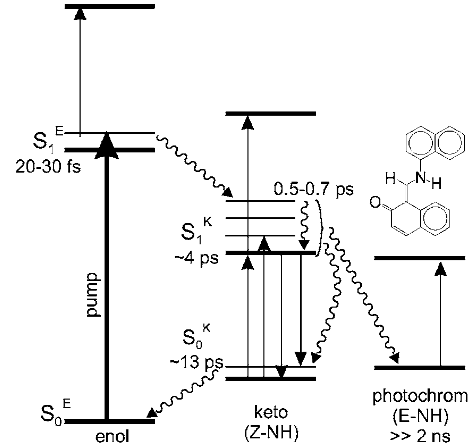
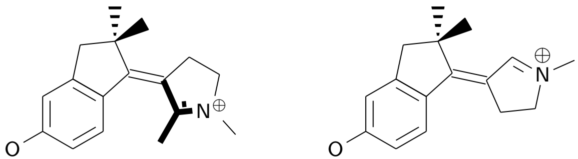


Figure 33. Jablonski diagram illustrating the relaxation routes in **HNAN** proposed by Fita *et al.* Straight and wavy arrows denote radiative and non-radiative transitions, respectively. Reprinted with permission from Ref [90]. Copyright 2006 American Institute of Physics.

the study of a synthetic switch molecule by Gueye *et al.*[106] In their comparison of the synthetic Rhodopsin analogues Z-**1** and E-**2**, (Figure 34) where the Z and E refer to the most stable ground state isomer, which undergo EZ isomerisation upon excitation at 400 nm.



Z-1 *E*-2

Figure 34. Z-**1** and E-**2**.

An example of these data and analysis are shown in Figure 35. The high quality of the data and the comparison between on- (excitation centred at 400 nm) and off-resonance (excitation centred at 800 nm) allowed them to discriminate between vibrational coherences on the S_0 surface and the S_1 surface. Using this, they were able to assign a low-frequency mode of 80 cm^{-1} which is propagated from the S_1 surface to the S_0 surface through a conical intersection. This enabled them to draw conclusions as to the motion of how Z-**1** and E-**2** approach the conical intersection: Z-**1** ballistically, E-**2** diffusively, and showed the single chemical modification necessary to change this behaviour to preserve the vibrational coherence across the conical intersection.

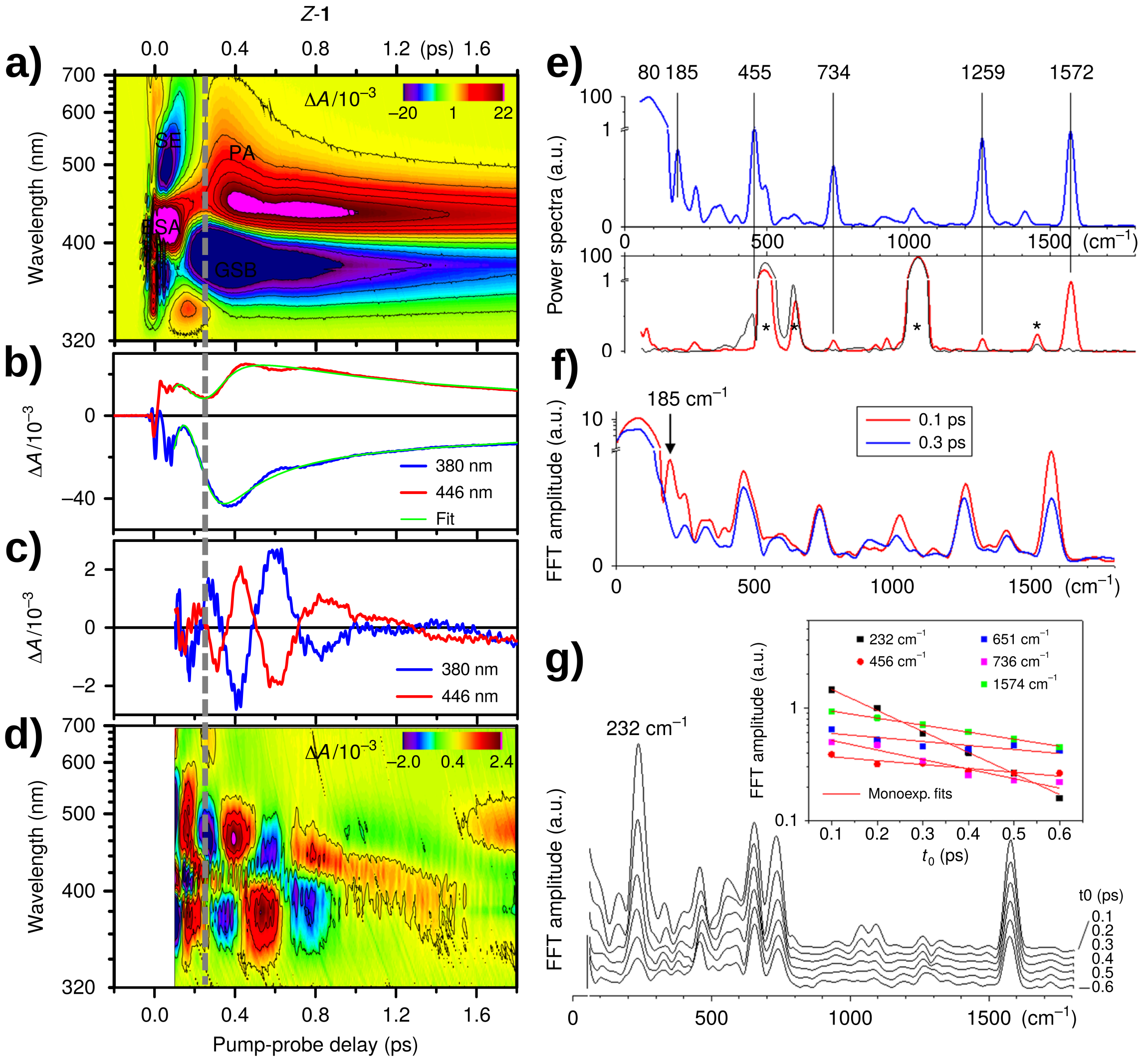


Figure 35. Data and analysis from the paper of Gueye *et al.* **a)** Transient Absorption Data of *Z*-**1** upon excitation using a 8.5 fs pulse centred at 400 nm., showing a clear 'node' at 250 fs, taken to be the moment of impulsive S_1 decay of **1** at the conical intersection. **b)** Selected time traces at 380 nm and 446 nm, as well as the global fit used to subtract the population dynamics. **c)** Residuals at time traces of **b**. **d)** Residuals at all wavelengths of the global fit of the TA data. **e)** Power spectra of the residuals in **d** (blue), and from off-resonant excitation centred at 800 nm (red). **f)** The same Fourier analysis of the results in **d**, performed on a sliding 1-ps long time window starting at 0.1 ps (red) or 0.3 ps (blue), i.e. before or after the impulsive decay. **g)** Same as **f**, for *E*-**2**, with an inset displaying the five dominant mode intensities as a function of the initial position of the Fourier window. All experiments were done in methanol. Reproduced with permission from reference [106].

The second (and newer) method of using GKA to study wavepackets was published independently by van Stokkum *et al.*[107] and Schott *et al.*[108] The essence of the approach is to simultaneously model the population dynamics and the oscillations, using the equation

$$S(t,\tilde{\nu}) = \mathbf{C}(\theta)\cdot\mathbf{S} + \sum_{n=1}^{N_{osc}} DOAS_n(\tilde{\nu}) \times \cos(\omega_n t' - \phi_n(\tilde{\nu}))\exp(-\gamma_n t') \qquad (50)$$

where the first term is a parametrised global kinetic fit as in equation 2.10 or 2.15, and the second term represents a sum of n oscillations ω_n, with a wavenumber-dependent phase ($\Phi_n(\tilde{\nu})$), which dephase in an exponential manner with rate constant γ_n and have amplitudes across the wavenumber range contained in the 'Damped Oscillation Associated Spectrum' $DOAS_n(\tilde{\nu})$. In the framework of this equation, the pump-related artefacts (i.e. XPM, RAs, the coherent artefact and the perturbed polarisation decay) are able to be considered and separated by their damping behaviour (i.e. damping that is within the IRF), and thus this enables the entire dataset to be analysed - as opposed to e.g. the paper on Z-**1** and E-**2**, which only considers data after 100 fs to avoid this. There are a number of potential advantages and disadvantages to this method of wavepacket analysis - a clear advantage is that one is able to extract the frequencies of all vibrations involved up to the noise level of one's data - not necessarily possible by using a Fourier transform, as minor vibrations are able to be obscured by larger amplitude and/or broad ones. Additionally, one gets the phase information, which is also not achieved by the simple Fourier transform. However, these come with a significant caveat (noted in the paper of van Stokkum *et al.*[107] and not mentioned in the paper of Schott *et al.*[108]) that the fitting assumes a perfect Lorentzian line shape (i.e. a single exponential dephasing time [47]). As noted by van Stokkum *et al.* this is not necessarily an accurate assumption - bath memory and coherence transfer may perturb this, and so the number of resolved oscillations can not be said to correspond to the number of coherences present. It was noted that the agreement between the extracted parameters plotted as Lorentzian line shapes and the Fourier transform were in good agreement however, indicating that the vibrational frequency and damping rates were at least well described (though it was noted that in multiple cases likely the same vibration is described by more than one $DOAS$ due to the non-exponential nature of the damping of the vibration in question). The other clear disadvantage is that the fitting of a considerable number of parameters (e.g. 35 oscillations for the complex biomolecule measured in the van Stokkum study)[107] does increase computational time, in direct contrast to the Fourier transform (a computationally extremely quick calculation).

The method of oscillation modelling was used by the Schott *et al.* to distinguish between the various possibilities of products after the photodissociation of the $D_{\infty h}$ trihalide ions I_3^-, Br_3^-, IBr_2^- and ICl_2^-, finding no neutral products present immediately after excitation and only a single diatomic product. They were also able to observe that in the $C_{\infty h}$ point group trihalides I_2Br^- and I_2Cl^-, only the homonuclear product is observed, not easily identifiable from the absorption spectra alone. Thus, the method of modelling the oscillations is able to give additional useful chemical insight to this problem and, while the assumptions of the model must make one wary of over-interpretation of the 'individual' vibrations, can assist in understanding wavepackets when there are vibrations of considerably different intensity involved in the problem.

3.4. Conclusions

Global Kinetic Models have enabled considerably clearer insight into a variety of chemical problems, particularly in sub-ns transient fluorescence and ns transient absorption experiments, where solvation and vibrational relaxation are not generally observed and so the key assumption of data bilinearity holds. In these domains, these models have enabled a full kinetic model of an archetypal rotor molecule to be elucidated and for guidelines for rotor design to be deduced,[64] for the unambiguous assignation of 9-aminoacridine as an anisotropic rotor [85] and for an excellent description of an excited-state proton transfer reaction [102], amongst others. When non-bilinear contributions are present in the data and/or the population dynamics under investigation are highly nonexponential, the separation of the single exponential components into genuine ‘states’ should be done with extreme care, as it can prove actively unhelpful in the comprehension of the real physics of the system. Comparison with steady-state data (as in the Fita paper)[90] is a useful strategy in this case, as is a ‘holistic’ view of any very short time constants observed. These models have also found use as a background-subtraction technique in the analysis of wavepackets, with insight into the approach to a conical intersection being recently gained.[106] The models have also been extended to explicitly account for vibrations, a method that is very useful for the extraction of weak vibrations and phase information. Overall the use of Global Kinetic Models can be useful - if one keeps in mind the wide selection of assumptions that go into these models, and avoids the spectre of over-interpretation.

4. Lifetime Distribution Models

Lifetime distribution models (LDMs) are data analysis techniques which attempt to describe the data by a continuous distribution of exponential decays. LDMs can be divided into three classes. These are: 1, fitting the underlying distribution directly (i.e. assuming a distribution to simulate the decay and optimising the parameters of the chosen distribution vs. the data); 2, a regularised deduction of the lifetime distribution from the data; and 3, non-parametric approaches that 'straddle the line' between LDMs and Soft-modelling (Section 5) methods, though these are are less common. The regularisation methods (which will be defined and discussed in more detail in Section 4.4) typically differ from how the global kinetic models in Section 3 worked in that regularisation methods are often applied to each wavelength separately and then a lifetime 'map' of the data is plotted. Additionally, both regularisation methods and non-parametric approaches lie at the interface of model-dependent and model-independent data analysis. This is due to the fact that the model of classical kinetics (and thus exponentials) underlies these analyses but no bias is imposed upon the shape of the distribution. Thus these methods do not in general make the assumption of bilinearity, though they do make the assumption of classical kinetics. Direct fitting may or may not assume bilinearity depending on the particular method employed.

4.1. Theory

4.1.1. Motivation

Why should we care about large distributions of rates? Are they common? It is entirely possible that they are more common than one would imagine - as shown in 1985 by James and Ware,[109] it is possible to achieve satisfactory fits using a biexponential function when the underlying data have in fact been generated from distributions. This is pictorially illustrated in Figure 36, which shows that a biexponential model may fit data generated using a Gaussian distribution of rates very well, despite these two functions implying very different things about the underlying process.

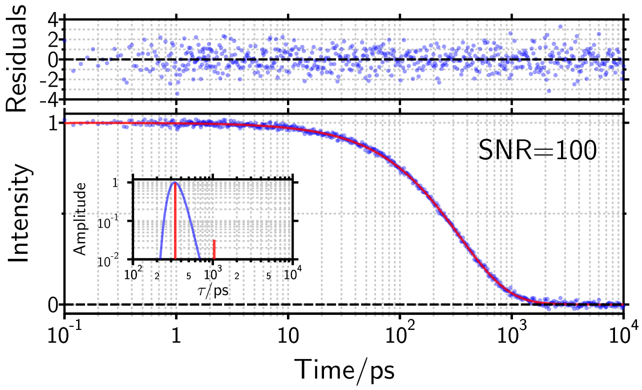


Figure 36. Data generated using a single Gaussian distribution (blue line in inset in lower panel) and fit using a biexponential function (red sticks in inset in lower panel). As may be seen, the biexponential function well describes the data and gives unstructured residuals.

James and Ware showed this more thoroughly by generating data from multiple rate distributions: two delta functions (i.e. a genuine biexponential function), from one or two Gaussians, and from two triangular distributions. In each of these cases 'realistic' noise (not described) was added to the dataset and a biexponential function was sufficient to describe all of the simulated data sets.

The question of 'how well does a biexponential function fit data simulated using a distribution' has here been re-tested for single traces generated using equations 51 - 52,

$$\Gamma(k) = \frac{1}{\sigma_1\sqrt{2\pi}}\mathrm{exp}\left\{-\frac{(k-\mu_1)^2}{2\sigma_1^2}\right\} + \frac{1}{\sigma_2\sqrt{2\pi}}\mathrm{exp}\left\{-\frac{(k-\mu_2)^2}{2\sigma_2^2}\right\} \tag{51}$$

$$\begin{aligned} I(t) &= \int_0^\infty \Gamma(k)\mathrm{exp}\{-kt\}\mathrm{d}k \otimes \mathrm{IRF} \\ &\simeq \sum_{k=10^{-6}}^{k=10} \Gamma(k)\mathrm{exp}\{-k(t-t_0)\}\mathrm{exp}\{0.5\cdot k^2\cdot\sigma^2\}\left[1+\mathrm{erf}\left(\frac{t-t_0-k\cdot\sigma^2}{\sqrt{2}\sigma}\right)\right], \end{aligned} \tag{52}$$

which are decays generated from a rate distribution $\Gamma(k)$ of two Gaussians, centred at 0.3 and 0.03 ps^{-1} (Figure 37), convolved with a Gaussian IRF of 100 fs width. More detail on these simulations is provided in appendix A. These data then had noise added and were then fit using a biexponential trial function (with an initial guess provided by an exploration of a large range of guess lifetimes using a simulated annealing procedure) and the χ^2_ν recorded. The procedure of the addition of noise and fitting was done 10,000 times and and an average χ^2_ν calculated. These results show that even when the distributions are remarkably broad (Gaussian FWHM being 25% of the centre) one must have a SNR of greater than 100 to distinguish that a biexponential is not an accurate description of the data. This is an important observation - unless one has exceptional data quality, *a priori* one cannot distinguish between a distribution of rates and a biexponential. Thus, it would be of use to be able to analyse the data without the assumption of a specific rate model (or to trial fitting with an assumed distribution), ideally using an algorithm that is as unbiased as possible toward a specific description of the data.

Additionally, there are a variety of physical and chemical problems where one should anticipate a distribution of rates. For example, where one has a chromophore in multiple conformations and/or environments - chromophores in multiple protein-induced conformations [110] adsorbed to surfaces,[111] or heterogeneous and slowly relaxing solvent environments (e.g. ionic liquids or glasses)[112] among others.

4.2. Rate Distributions

If one presumes that one's kinetics may be described by a distribution, it then becomes reasonable to ask what distributions may we expect to see. These will be more relevant for the direct fitting method, as this method requires the assumption of a functional form of the rate distribution. Here, two common rate distribution

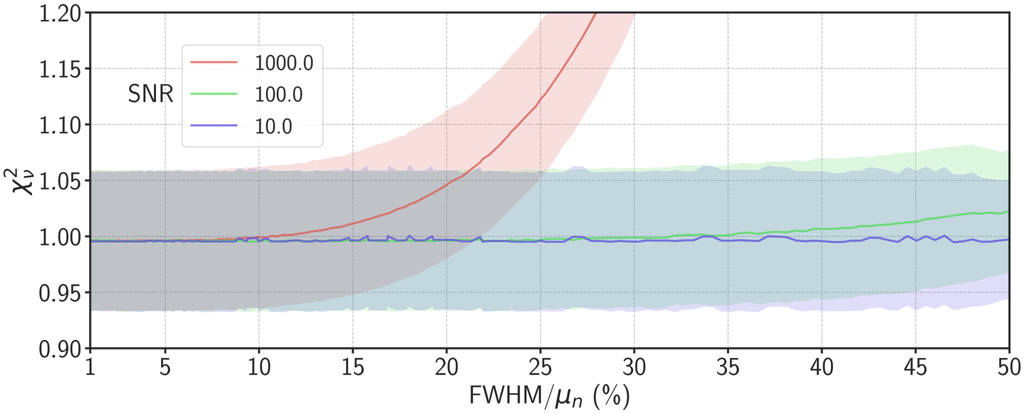


Figure 37. Reduced Chi Squared values of decay data generated using equations 51 and 52, and then fit assuming a biexponential decay. Data points are averaged Reduced Chi Squared values of 10,000 fits, shading is standard deviation. $\mu_1 = 0.3\,\text{ps}^{-1}$, $\mu_2 = 0.03\,\text{ps}^{-1}$.

forms will be discussed and their properties elucidated, as well as one uncommon, but potentially useful distribution.

4.2.1. Gaussian Distribution

The Gaussian distribution is used to represent the probability density function of a normally distributed random variable - thus, if one anticipates a normal distribution of a molecule about a certain parameter the Gaussian distribution is the natural distribution to model this. In addition, the central limit theorem states that if one considers any N number of random variables (with not necessarily Gaussian probability density functions), as N increases the distribution of this sum approaches the Gaussian distribution.[113] Thus, even if one has multiple sources of heterogeneity that are non-Gaussian, a Gaussian may be the result. As such, this is an exceptionally useful and pervasive function. The Gaussian distribution of rates, $\Gamma_g(k)$, is defined as

$$\Gamma_g(k) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(k-\mu)^2}{2\sigma^2}\right\} \tag{53}$$

where μ is the expected value and σ is the standard deviation of the distribution.

4.2.2. Skewed-Gaussian Distribution

The Skewed Gaussian [114] distribution of rates, $\Gamma_{ln}(k)$,

$$\Gamma_{sk}(k) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(k-\mu)^2}{2(\sigma+\beta)^2}\right\} \text{if } k < \mu \\ \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(k-\mu)^2}{2(\sigma-\beta)^2}\right\} \text{if } k \geq \mu \end{cases} \tag{54}$$

where μ and σ are defined as in the Gaussian Distribution and β is an additional skewness factor. This distribution has been described as useful when minor asymmetry is present in a distribution, that is when the distribution may be considered ‘near-normal’. The derivation of such a class of distributions arose from the doubts about the unqualified use of normality, which may be summed up in a quote by Geary

[114]: ‘Normality is a myth; there never was, and never will be, a normal distribution’. To our knowledge, the Skewed-Gaussian distribution has not been widely applied to rate distributions, but could be a function of use if one’s rate distribution is indeed near-normal.

4.2.3. Stretched and Compressed Exponential

A function that has been used to describe various relaxation phenomena that corresponds to an underlying rate distribution is that of the stretched exponential function,

$$I(t) = \exp\big[-(t/\tau_0)^{\beta}\big] \tag{55}$$

which is often used as a purely empirical decay law, with $0 < \beta < 1$. τ_0 is the lifetime parameter (corresponding to the exponential lifetime when $\beta = 1$). This equation (approximately, the distribution may only be represented by elementary functions for $\beta = 1/2$) corresponds to a rate distribution ($\Gamma(k)$) of the form:

$$\Gamma_\beta(k) = \tau_0 \frac{B}{(k\tau_0)^{(1-\beta/2)/(1-\beta)}} \times \exp\left[-\frac{(1-\beta)\beta^{\beta/(1-\beta)}}{(k\tau_0)^{\beta/(1-\beta)}}\right] f(k) \tag{56}$$

where the auxiliary function *f(k)* is

$$f(k) = \begin{cases} 1/[1 + C(k\tau_0)^{\delta}], \ \delta = \beta(0.5-\beta)/(1-\beta), \ \text{if } \beta \le 0.5 \\ 1 + C(k\tau_0)^{\delta}, \ \delta = \beta(\beta-0.5)/(1-\beta), \ \text{if } \beta > 0.5 \end{cases} \tag{57}$$

and the parameters B and C, arising from fitting the approximate analytical equation 56 to the numerically integrated exact solution, corresponding to different β values are shown in Table 3 (values of β in between these are accessible by interpolation).

β	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
δ	2/45	3/40	3/35	1/15	0	3/20	7/15	6/5	18/5
B	0.145	0.197	0.243	0.285	$1/(2\pi^{1/2})$	0.306	0.360	0.435	0.700
C	0.89	0.50	0.35	0.25	0	0.13	0.22	0.4015	0.33

Table 3. Exponent δ and parameters B and C in Equations 56 and 57.

This numerical equation for the distribution was presented by Berberan-Santos *et al.*,[115] and the derivation (as well as an extensive discussion of the mathematical properties of the stretched exponential function) may be found in this paper. This numerical equation, whilst useful, does illustrate a shortcoming of the stretched exponential - it is only for very few values of β that analytical solutions for the rate distribution exist, which complicates interpretation if one uses it as the fitting function. A clear advantage however is that the implementation is simple - one uses it as a replacement for an exponential in making the **C** matrix as defined in equation 37, and then one may use it in a global or non-global spectral model to analyse the data as in Section 3. The other disadvantage of the stretched exponential model is that the rate constant at time zero is infinite - not a behaviour that would appear to correspond to a genuine physical picture. The problems associated with this undesirable short-time behaviour were ameliorated by Berberan-Santos *et al.* with the introduction of the Modified Stretched Exponential,[115]

$$I(t) = \exp\left[1 - \left(1 + \frac{t}{\tau_0}\right)^{\beta}\right] \tag{58}$$

which was derived by setting $t = \tau_0$ as the time origin, and renormalising. The distribution of rate constants may be easily obtained as

$$\Gamma_{\beta}^{m}(k) = \exp(1 - k\tau_0)\Gamma_{\beta}(k). \tag{59}$$

A further modification was introduced, enabling any time to be selected as the time origin by defining the dimensionless parameter α

$$\alpha = \frac{t_0}{\tau_0} \tag{60}$$

and the further modified stretched exponential

$$I(t) = \exp\left[\alpha^{\beta} - \left(\alpha + \frac{t}{\tau_0}\right)^{\beta}\right] \tag{61}$$

and distribution

$$\Gamma_{\beta}^{\alpha m}(k) = \exp(\alpha^{\beta} - \alpha k\tau_0)\Gamma_{\beta}(k). \tag{62}$$

Here, for a large α, the decay becomes a stretched exponential only for very long times. For further discussion, the reader is again referred to the original paper.[115] The modified versions of the stretched exponentials thus may be useful distributions to be used in the analysis of transient spectroscopy data where a distribution is presumed. In general however the stretched exponential is not presumed to have an underlying physical basis and the distribution is not interpreted. Typically the value of β is taken to give a degree of the amount of heterogeneity in the rate, but not interpreted further.

The compressed exponential is defined as equation 55 where $\beta > 1$. This may be interpreted as that the rate increases with time, starting from zero, and the decay is super-exponential. This has received considerably less attention and use than the stretched exponential, however it has found some use in the literature - it was previously found to be relevant in protein folding kinetics, with the interpretation that they were observed due to a free diffusion down a potential energy surface.[116] In addition, it was found to be relevant in non-equilibrium charge recombination, with the interpretation being that the speeding up of this process with time was due to the excited state being initially populated far from the crossing region between the reactant and product states.[117]

4.3. Direct Simulation of the Rate Distribution

The method of directly simulating a rate distribution and then comparing the time decay that would arise from this distribution is conceptually the simplest of the methods discussed in this section. The method is, if done in a 'non-global' manner, a matter of simulating the rate distribution $\Gamma(k)$, dependent on parameters θ and then using equation 52 to simulate the time trace. This time trace is then compared to the data and the parameters θ are optimised to achieve the best agreement. If done in a global manner, the initial step is the same but simply the result of equation 52 is used as the time behaviour of one 'species' in the **C** matrix, and then the optimisation of the parameters θ proceed as in equations 41-43.

4.4. Regularisation Methods

If one wishes to directly access the rate distribution $\Gamma(k)$, which enters into the 'true' intensity decay of one's system via:

$$f(t) = \int_0^\infty \Gamma(k)\exp\{-kt\}\mathrm{d}k \tag{63}$$

then one would, formally, have to perform an inverse Laplace transform to access the form of $\Gamma(k)$. This problem is ill-conditioned [118] (i.e. a small error in the initial data may cause a large error in the answers) and thus to attempt to solve it one must make use of two tools: *discretisation* and *regularisation*. *Discretisation* enables us to approximate the continuous distribution by a tractable number of points in the presumed distribution of rates/lifetimes. Thus, one can describe the time evolution at a single wavenumber as:

$$\mathrm{D}(\tilde{\nu}) = \sum_{j=1}^{n} \theta_j(k_j,\tilde{\nu})\exp\{-k_j t\}\otimes\mathrm{IRF}(t,\tilde{\nu}) \tag{64}$$

where $\theta(k_j,\tilde{\nu})$ are the pre-exponential amplitudes that represent the probability of the process with rate k. In the equivalent matrix form this equation is expressed as

$$\mathrm{D}_{\mathrm{fit}}(\tilde{\nu}) = \mathbf{E}\theta_{\tilde{\nu}} \tag{65}$$

where $\mathrm{D}_{\mathrm{fit}}$ is a vector of length n, i.e. one time trace, $\theta_{\tilde{\nu}}$ is a vector of length n_{rates} containing the pre-exponential amplitudes associated with the number of rates. **E** is an $m \times n$ matrix of the IRF convolved exponential decays.

Discretising does improve the conditioning of the problem (and thus the stability and tractability). However, this number can however still be very large (possibly comparable to the number of data points in the experiment) and thus whilst the conditioning is improved, it is still not sufficiently improved. If one simply analyses the data without regularisation, the most mathematically optimal solution will be found - with the large number of fitting parameters, this will be a result that fits to both the data and the noise, an undesirable occurrence. Regularisation, is, in essence, a process whereby the fitting function is modified with a 'penalty' term to prevent undesirable outcomes of the fitting procedure. This penalty term often effectively smooths the data in some way, though this is not necessarily the case. Regularisation introduces a bias into the fitting procedure to attempt to make the deduction of the underlying rate distribution less sensitive to noise. These problems generally take the form of the equation

$$\underset{\theta}{\text{minimize}} \quad \chi^2_N + \alpha(\mathrm{R}(\theta)) \tag{66}$$

where

$$\chi^2_N = \frac{1}{n_{data}} \sum_{i=1}^{N} \frac{[D(t_i) - D_{fit}(t_i)]^2}{\sigma_i^2} \tag{67}$$

where n_{data} is the number of data points and σ_i is the noise variance associated with the ith point. The χ^2_N differs from the χ^2_ν defined in equation 38 by the lack of division by degrees of freedom. Instead, one divides simply by the number of data points as in principle the number of "fitted parameters" could be as large as the number of data points here. This definition of χ^2_N follows that of Livesey and Brochon.[119] In essence, one minimises the difference between the fit and the data, with the fit constructed using the parameters θ, with the compensating regulariser R reducing the bias in the solution. The regulariser is therefore also dependent upon the fit parameters θ in some way. The hyper-parameter (a positive value) α controls the trade-off between the residual minimisation and the regularisation. As α tends to zero one will get a very accurate, but exceptionally sensitive to perturbation, result - by contrast, if one has a high α, one anticipates that the resulting distribution of lifetimes will be less noisy.

It should be noted that here the χ^2_N statistic has been chosen for convenience and readability - in actual fact, the χ^2_N statistic may be replaced by a simple sum of squared residuals, a log-likelihood function, or in essence any fit statistic one may think of.

Various forms of regularisation exist in the literature - common ones that will be discussed here are: Tikhonov regularisation (also known as ℓ^2 regularisation or ridge regression), LASSO (or ℓ^1-regularisation), elastic net regularisation (a combination of LASSO and Tikhonov regularisation) and the maximum entropy method.

4.4.1. Tikhonov regularisation

Tikhonov regularisation [91, 93, 118, 120] is the method of regularisation that has likely been in use the longest, having been introduced by Tikhonov in 1964.[121] It places a direct penalty upon the amplitudes of the fit parameters (θ) using an ℓ^2-norm, which we define for a vector $\mathbf{x}$ as

$$||\mathbf{x}||_2^2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}. \tag{68}$$

In short, the minimisation problem 66 becomes

$$\underset{\theta}{\text{minimize}} \;\; \chi_N^2 + \alpha||L\theta||_2^2 \tag{69}$$

where L is a matrix which is most frequently the identity matrix, though a number of alternative matrices may be used (1st and 2nd derivative approximations, as well as a fusion of the identity and the 1st derivative matrix)[120]. This problem may be elegantly optimised by using augmented matrices - this is well explained in the paper by Dorlhiac *et al.*[120] The use of the ℓ^2-norm presumes that the noise in the data and the parameters (i.e. the rate distribution) are normally distributed, thus will produce broader kinetic distributions than may be genuinely present - it may thus also prove difficult to separate out closely spaced kinetic contributions.

4.4.2. LASSO

The LASSO (least absolute shrinkage and selection operator)[122] is similar to Tikhonov regularisation - it places a penalty upon the amplitudes of the fit parameters, it is simply that this penalty is the ℓ^1-norm instead of the ℓ^2-norm. The ℓ^1-norm is defined for a vector $\mathbf{x}$ as

$$||\mathbf{x}||_1 = \sum_{k=1}^{n} |x_k|. \tag{70}$$

This does have the advantage that as whilst in Tikhonov regularisation no parameter will be shrunk to zero, the use of the ℓ^1-norm enables such behaviour. As such, the LASSO produces sparser solutions,[122] with the downside that the problem

$$\underset{\theta}{\text{minimize}} \;\; \chi_N^2 + \alpha||L\theta||_1 \tag{71}$$

does not have a closed form solution (unless it is the rare case where the columns of the $\mathbf{E}$ matrix are orthogonal). Fortunately, relatively efficient algorithms have been developed to solve this regularisation problem, with an efficient one, replicated here, being [123]

Algorithm 1.

```
1. Set initial guess for $\theta_{\tilde{\nu}}$ to the Tikhonov regularised solution
2. Compute $\gamma_1$, $\gamma_1$ is the largest eigenvalue of $\mathbf{E}^{\mathrm{T}}\mathbf{E}$
3. Set $\mathbf{B} = \gamma_1\mathbf{I_p} - \mathbf{E}^{\mathrm{T}}\mathbf{E}$ where $\mathbf{I_p}$ is the $p \times p$ identity matrix, where $p$ is the number of parameters
4. Set $K$ to 0, $K$ being an iteration counter
5. For each coefficient $j$ in $\theta_{\tilde{\nu}}$
While $\frac{\theta_j^K - \theta_j^{K-1}}{\theta_j^{K-1}} > 10^{-12} \wedge \theta_j^{K-1} \neq 0$ a) $K = K + 1$
b) $\mathbf{U}^{(\mathbf{K})} = \mathbf{E}^{\mathrm{T}}\mathbf{E_j}\theta_j + \mathbf{B_j}\theta^{(K-1)}$
c) $\theta^{(K)} = \mathrm{sgn}(\mathbf{U})(\frac{|\mathbf{U}|}{\gamma_1} - \alpha)^+$
```

with more detail being provided in the original paper by Xiong *et al.*, as well as the paper by Dorlhiac *et al.*.[120, 123]

4.4.3. Elastic Net

Elastic net regularisation is simply a combination of Tikhonov regularisation and LASSO - it has been seen to outperform the LASSO whilst maintaining the ability to provide sparser solutions than the pure Tikhonov regularisation. The elastic net minimisation is

$$\underset{\theta}{\text{minimize}} \;\; \chi^2_N + \alpha(\rho||L\theta||_1 + (1-\rho)||L\theta||_2^2) \tag{72}$$

with $0 \leq \rho \leq 1$ being an additional hyper-parameter that determines the weight of the two different norms. The use of augmented matrices is also useable here, as again discussed in the paper by Dorlhiac *et al.*[120]

4.4.4. Maximum Entropy

The maximum entropy method [124–126] (MEM) has a somewhat different form of the regulariser. The equation that is to be minimised in this case is

$$\underset{\theta}{\text{minimize}} \;\; \chi^2_N - \alpha S(\theta) \tag{73}$$

where S is the Shannon-Jaynes entropy,[127] defined as

$$S = -\sum_{j=1}^{M} \theta_j \ln\left(\frac{\theta_j}{\Theta_j}\right) \tag{74}$$

where θ is, as before, the rate distribution and Θ is the so-called prior distribution,[128] which incorporates any prior knowledge about the distribution. Often this, if no knowledge is available, is simply set as flat (as is the initial guess).[129] The MEM has several advantages over the other forms of regularisation discussed here - the primary one being that it, unlike the other forms of regularisation, is strictly justified in information theory terms.[124, 125, 128, 129] Thus the MEM seeks to find the distribution that best describes the data with the least amount of bias in the solution (thus, the *maximum* entropy) - the solution that is strictly justified by the data. The disadvantage relative to the previous forms of regularisation discussed is that, at least in the form described in equation 74, the rate distribution may only have positive amplitudes - not necessarily true if one has e.g. a transient absorption experiment or rising terms in a transient fluorescence experiment. Fortunately this problem was addressed by simply expressing the rates as a difference of two positive sets of amplitudes and computing the Shannon-Jaynes entropies of these terms separately then summing them,[128, 130] overcoming this and enabling rate distributions of any sign to be analysed using this method.

4.4.5. The Hyper-Parameter

Some attention deserves to be paid to the hyper-parameter. It is readily apparent that the manner in which all of these regularisation techniques will succeed or fail in converging in a manner which we would consider accurate is dependent upon the value of the hyper parameter(s) α (and ρ). Too large a value and the χ^2_N statistic will be virtually irrelevant and the parameters will have nothing to do with the data, too small a value and we will be fitting to miniscule amounts of noise. Choosing the value of the hyper-parameter in a non-automated manner would be both deeply biased and deeply tedious, and so is not recommended. There are a variety of automated methods for picking the best hyper-parameter, based on the data and the regulariser. For Tikhonov regularisation, LASSO and elastic net there are multiple popular methods to select the best hyper-parameter. These include the generalised-[93, 120] and k-fold [131] cross validation, the C_p statistic,[120, 122] the L-curve [93, 120] and the minimal product method.[93] Their use and a comparison of their effectiveness is more thoroughly described in the paper of Dorlhiac *et al.*,[120] and the paper by Slavov *et al.*[93] For the MEM, the subject of the regularisation parameter is often ill-discussed (it is hardly mentioned in the paper by Kumar *et al.*)[128] however this was remedied in a paper by Lórenz-Fonfría and Kandori [132] where they discuss the effectiveness and use of the generalised cross validation, L-curve, Bayesian inference methods (also discussed previously by Skilling and Gull [133]) and the so-called ‘discrepancy’ criterion (optimising the α parameter until the χ^2_N value reaches 1).

4.4.6. Comparison of Regularisation Methods

All of the regularisation methods mentioned above are similar in that they aim to numerically stabilise the inverse Laplace transform to extract the genuine underlying rate distribution from experimental data. They differ in the regularisation term and thus will give different answers for identical data. As such, a (brief) comparison between the methods and a guide of when they are useful to use is presented here. A more detailed discussion has recently been presented by Smith *et al.*[134] The Tikhonov regulariser, as discussed, assumes that the noise and fit parameters are normally distributed. If this assumption is known to hold it may be advantageous to use this method as due to the augmented matrix implementation it is both fast and easy - however it will have difficulty separating out closely spaced kinetic contributions and so if one does not anticipate a broad distribution of rates it may not be of use. The LASSO, having changed the ℓ^2-norm for the ℓ^1-norm, will provide sparser solutions. The disadvantages of this method are that it may provide overly sparse solutions and that there is no closed form solution, making implementation harder than Tikhonov regularisation. A potential solution to this is that of the elastic net, which provides a mixture between the Tikhonov and the LASSO regularisers, and so may be considered to be a better pick when nothing may be anticipated about the underlying distribution. A clear disadvantage of the elastic net is that as there are two hyper-parameters (one to control the amount of regulariser used in the optimisation and one to control the trade-off between Tikhonov regularisation and LASSO regularisation) and as such selection of these optimal parameters is more time-consuming. A significant problem is that none of the preceding regularisers are unbiased - the Tikhonov regulariser will tend to a solution of a normal distribution, the LASSO a sparse distribution. This is due to the form of the specific regularisers - chosen to stabilise the solution, not chosen to pick the correct solution. The maximum entropy regulariser is different, in that the regulariser is chosen to specifically arrive at the least biased solution from an information theory perspective.[124, 125, 128, 129] This means that is has the clear advantage that one will arrive at the least biased solution and the solution that is explicitly justified by the data. The clear disadvantages of this method are the difficulty of implementation, though there is recent work in this direction.[135] Overall, the maximum entropy method is, if able to be implemented, clearly superior in terms of the robustness of the solutions.

4.5. Non-Parametric Methods

In addition, multiple approaches exist that 'straddle the line' between LDMs and Soft-modelling (Section 5) methods. These methods re-represent the kinetics by use of a numerical transform in order to provide a characterisation of the data. Stazger and Zinth [136] suggested the use of differentiation of the data on a logarithmic time scale (calling their method Logarithmically differentiated absorption changes, LDAC) to visualise the lifetime components in the data, and pointed out that this could aid in the identification of distributions in lifetimes. They noted that numerical differentiation of the data could run into problems due to noise, but suggested various options of smoothing the data to overcome this (though it should be noted that this smoothing does represent a loss of statistical information).

Berg and Kaur recently suggested the use of log-moments [137] to characterise kinetics in one or many dimensions, thoroughly describing the mathematics behind the method and illustrating its use on a variety of different types of decay function.[137] They show that the lower log-moments are relatively insensitive to data fluctuations (making them likely more attractive than the LDAC process), and that the number of stable, reproducible, log-moments provides information on the information content of the data. They also noted that log-moments do require convergence of the data at short and long times, though noted that this is not necessarily a weakness in the method, but a legitimate data concern: measurements lacking sufficient time range and resolution may misrepresent the system kinetics.[137] For a full discussion the original paper is much recommended.

4.6. Published Implementations

For Tikhonov Regularisation, LASSO and Elastic Net Regularisation free software exists in Python (open source)[120] though it should be noted that upon testing by us, various elements of the Program 'PyLDM' crashed, and no option to save the analysis currently exists in the program. A closed source option programmed in the Matlab programming language 'OPTIMUS' also exists, which also can perform global kinetic analysis as described in Section 3.[93] Recently Đorđević *et al.* also published an open-source implementation in Python, validated using data taken by (some of) the authors of this review.[131] For the MEM, Steinbach *et al.*[138] has published a freeware program that is still currently available and was last updated in late 2017 (it is also closed source). It is noted by us that all of these do (at present) appear to be significantly less used by the transient spectroscopy community than the implementations of Global Kinetic Models discussed in Section 3 - 89 citations for the

Steinbach freeware as of early 2020 according to Google Scholar, 5 for the 'PyLDM' program and 51 for the OPTIMUS program. For simulations of rate distributions to fit the underlying the distribution directly, as well as the non-parametric methods, there exists no published software explicitly for this purpose (to our knowledge).

4.7. Applications

Similar to the situation with GKMs, LDMs have been applied to a wide range of problems. As such only a select few examples of their use, in addition to their strength and weaknesses, will be given here.

The analysis of assuming a rate distribution and optimising its parameters to reach closest agreement with the data has been applied to study the spatial distribution of hydroxyl groups on γ- and δ-alumina surfaces by Métivier *et al.*[139] Specifically, they studied the time-resolved fluorescence of a pyrene derivative **PPTEOS** (Figure 38) which was able to be covalently grafted to the hydroxyl groups of alumina using TCSPC.

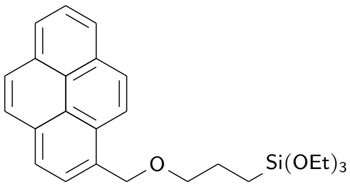


Figure 38. **PPTEOS**.

As pyrene molecules may form excimers when close in space, the distribution of the fluorescence lifetimes (as well as the relative intensity of the excimer fluorescence) enabled the clear deduction that γ and δ-alumina surfaces are very different. This was shown by separately analysing the excimer and monomer fluorescence decays using a sum of two stretched exponentials, one corresponding to a rising component in the excimer fluorescence case. The distributions of rise times of the excimer were narrower for the δ-alumina, and the centre of the decaying lifetime distribution was also lower, providing evidence that there was more interactions amongst the excimers in this alumina (Figure 39). This enabled Métivier *et al.* to gain the clear physical insight that δ-alumina has a larger number of Lewis acid sites and much larger (or more frequent) clustered zones than the γ-alumina, due to this narrower and shifted distribution.

Though this gave additional physical insight, it did have the disadvantage that (as discussed earlier) that the stretched exponential is a phenomenological function (in general) and that the Laplace inversion of it is not analytical except for a few values of β. Another example of analysis of assuming a rate distribution and optimising its parameters to reach closest agreement with the data is a study of tryptophan (**Trp**, Figure 40) residues in a variety of proteins.[140] As compared to the fluorescence

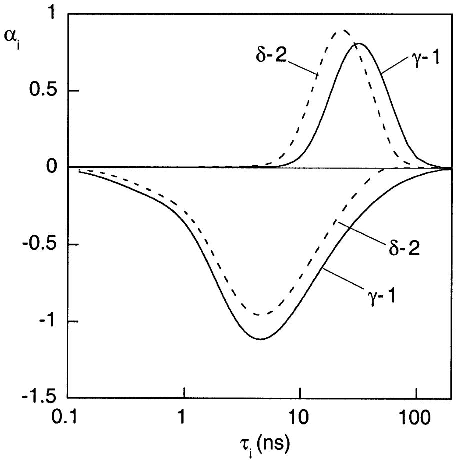


Figure 39. Lifetime distributions of rising components (negative peak) and decaying components (positive peak) from the fit of two stretched exponentials to the excimer fluorescence of Pyrene attached to γ-1 or δ-2 alumina surfaces. Reproduced from reference [139] with permission from the PCCP Owner Societies.

of **Trp** in buffer, which gave a single exponential decay, the fluorescence decays from the **Trp** in proteins were non-exponential. These were analysed in terms of distributions, which were interpreted as arising from the protein fluctuating around a set of conformational coordinates.

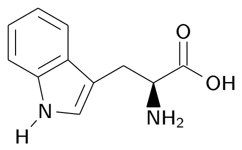
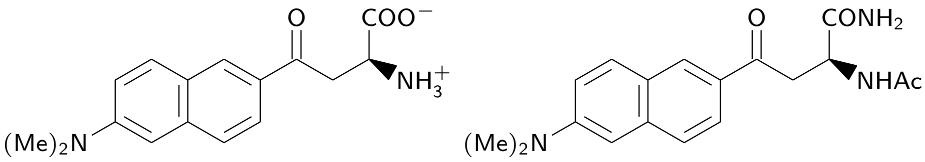


Figure 40. Tryptophan, **Trp**.

A significant achievement of this paper was that the distributions were able to be explicitly related to a physical model of the protein,[110] wherein a single potential well of conformations with a width related to the motional degree of freedom is related to the lifetimes. This enabled Alcala *et al.* to conclude that a single potential well of conformational distributions is appropriate in the protein ribonuclease T1 (at physiological temperatures) but that multiple potential wells are needed to explain the conformational distributions of neurotoxin variant 3.[110, 140] A disadvantage of their analysis is that, for some proteins, there was little significant difference when analysing the distributions with a uniform model (a model that we would not anticipate to correspond to a genuine physical picture) and a Gaussian model. This is a disadvantage as compared to the regularisation approach.

The regularisation approach has also been used to study protein heterogeneity, by incorporating a synthetic amino acid **Aladan** (Figure 41) into streptococcal protein G at 7 different sites, thus creating 7 mutants which may report on solvation of a separate protein area.[141] In addition, the derivative **NAAA** (Figure 41) was measured alone as a control comparison. 2 sites buried in the protein were replaced

(F30, L7), 2 sites that were partially exposed to the solvent were replaced (W43, Y3) and 3 sites fully exposed to the solvent were replaced (A24, V21, T16). Fluorescence up-conversion was used to extract the time-resolved Stokes shift of the different proteins, and this multiexponential Stokes shift was analysed using the MEM. This is shown in Figure 42.



Aladan N-acetyl-Aladanamide (**NAAA**)

Figure 41. The synthetic amino acid and free dye control from the paper of Abbyad *et al.*.

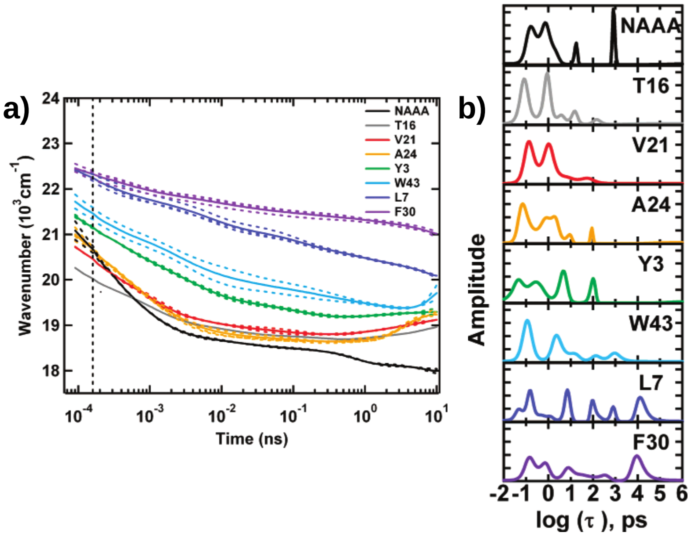


Figure 42. **a)** The time-dependent Stokes shifts of **Aladan** in the 7 protein environments specified as well as that of **NAAA** in a pH 5.4 buffer. The IRF is indicated using the vertical dashed line. Solid curves are site averages, dotted lines individual data sets. **b)** Rate distributions of the dynamic Stokes shifts in **a** using the MEM. Reprinted with permission from reference [141]. Copyright 2007 American Chemical Society.

The MEM analysis reveals that the rate distributions of the Stokes shifts in the three exposed residues (T16, V21, A24) are qualitatively very similar to that of the free dye derivative **NAAA**. However, the distributions in the partially and completely buried residues differ significantly, with rate distributions at longer lifetimes in the buried residues of F30 and L7. Using this analysis Abbyad *et al.* were able to clearly show that the solvation response in a protein is not most controlled by the local secondary structure - the V21 and A24 are in very different secondary structures but have very similar rate distributions. In fact, the solvation response in a protein is simply controlled by the level of exposure to the solvent. A key disadvantage of the MEM is simply put that it is complicated to implement - Abbyad *et al.* used the software of Steinbach [138] and almost all papers using the

MEM also function as a presentation of a particular group's implementation of the MEM, suggesting that the effort involved in programming the method is large.

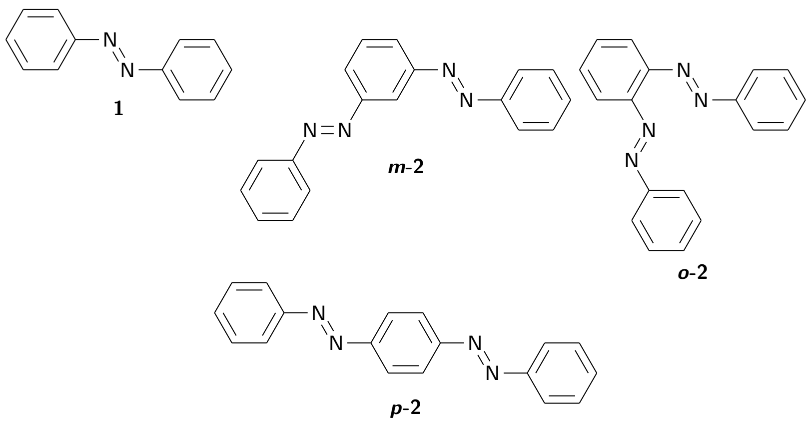


Figure 43. The various azobenzene derivatives investigated by Slavov *et al.*[142]

Another example of the use of regularisation methods was in the study of azobenzene derivatives (Figure 43) by Slavov *et al.*[142] The specific regulariser used was not specified in the paper nor the supplementary information, but the software used was the OPTIMUS program published by the same group,[93] so it can be assumed that it was Tikhonov, LASSO or elastic net regularisation. The lifetime maps (Figure 44) were able to give these authors insight into the physics of the azobenzene derivatives that was not available using GKA methods. For the E → Z isomerisation reaction, the transient spectroscopy of which is shown in Figure 44, they were able to clearly identify the hot ground states by the tilted nature of the lifetime distribution visible at c. 10 ps in all lifetime maps. Additionally, the authors were able to show that in fact the previously observed 'biexponential' decay to the conical intersection was more accurately described as a distribution - a considerably different picture of the physics as a probability distribution in the search for the conical intersection region caused by differences in the ground-state distribution. This shows a clear advantage of LDMs over GKMs - the non-exponential vibrational energy dissipation and/or solvation is implicitly taken into account by the use of a distribution of rates. Thus here the LDM was advantageous for clarifying how these processes may be viewed.

It must be noted that a large amount of the applications of LDMs are in biophysical problems involving proteins (likely due to their recognised dynamically heterogeneous nature), with seemingly less uptake in the transient spectroscopy of 'small' molecules. As the example of Slavov *et al.* demonstrates, these models can also offer useful clarification for their chemistry and physics - enabling in their case a clearer picture of how the molecule reaches the conical intersection to isomerise.

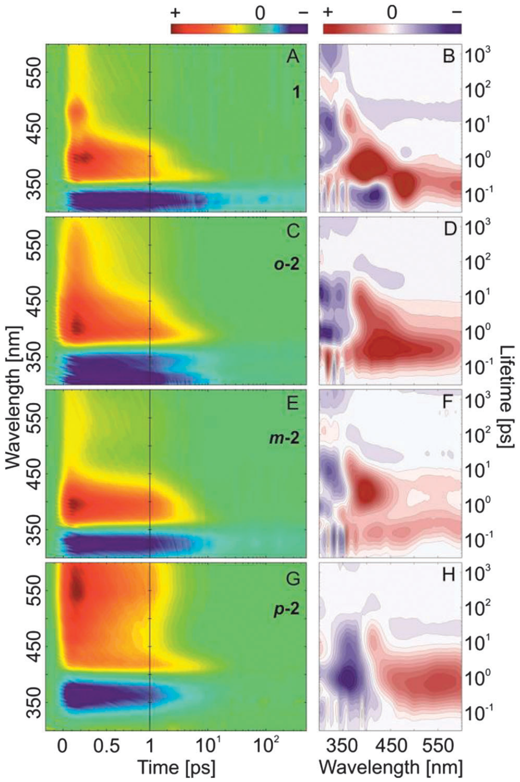


Figure 44. Transient absorption data for **1** (A), ***m*-2** (C), ***o*-2** (E), ***p*-2** (G) and the corresponding lifetime density maps (B-H). Excitation was to the π-π^* transition for the E → Z isomerisation reaction. Reproduced from reference [142] with permission from the PCCP Owner Societies.

4.8. Conclusions

Lifetime density methods can provide considerably clearer pictures of decay processes in heterogeneous environments than a global kinetic analysis that assumes clearly separated states. This is particularly relevant for proteins,[110, 141] molecules in confined spaces (e.g. micelles) and adsorbed on heterogeneous surfaces. Distributions are not only relevant to these situations however, and may provide clearer explanations for phenomena in 'small' molecules.[142] One must keep in mind that this does still assume that formal kinetics may be used to model the system, and thus that one has a distribution of rate constants. If differing physics enter one's system (e.g. diffusion) then these models may be considered inappropriate. A key success of these techniques is also that they are not exclusive to systems that have broad distributions of lifetimes - they may also be used to analyse systems that have discrete states. Indeed, given that LDMs do not require the assumption of bilinearity, they would seem well suited as both a 'first look' analysis and as an analysis that can be used to provide insight into the chemical and physical processes in the investigated molecular system. If one is able to equally well describe the data using a biexponential where the two lifetimes coincide with the centre of two peaks in the lifetime distribution map, then this simply provides more evidence that the data are 'truly' described by a biexponential and not a distribution.[109] The significant downside to these methods is that they are more complex to implement than the GKMs, and less functioning software is available. With the recent paper discussing their positives in fluorescence by Smith *et al.*[134] it is hoped that these may be overcome and the use of LDMs may become more commonplace in transient spectroscopy.

5. Soft-Modelling Methods

Soft-modelling methods are methods that do not assume a rate distribution, as the GKMs or LDMs do. They also do not assume that particular spectral shapes are to be expected in the data, as band-shape analysis models do. Soft-modelling methods assume that the data are bilinear and then endeavour to decompose the data into the smallest 'necessary' number of components to describe the data. As they do not assume explicit functions for the time or spectral behaviour, they are sometimes considered as preliminary 'exploratory' methods to assist in the postulation of a reasonable model.

5.1. Theory & Motivation

How can one say how many 'components' one has in a data set? To do so in a minimally biased way is of use both for learning about the system under investigation and for clarifying any future modelling that may occur (e.g. the number of rates anticipated in a GKM). This is the objective of soft-modelling or factor analysis methods. The term soft-modelling arises from its comparison to 'hard-modelling', i.e. parametric model fitting such as GKM. An additional useful property of these methods is that they may be applied to so-called 'multi-variate' or 'multi-way' data - i.e. one may simultaneously analyse multiple experiments using a connected set of parameters. This is, in theory, possible with the methods discussed in the previous sections, however it is done with considerable more ease in soft-modelling. Soft-modelling methods assume (in general) only that the data is bilinear and then endeavour to decompose the matrix into the significant components. In essence, all attempt to find a set of L basis vectors and scores that minimise the reconstruction error of the data. The main approaches that have been used in transient spectroscopy are singular value decomposition (SVD), and multivariate curve resolution-alternating least squares (MCR-ALS, of which SVD also forms the basis of).

5.1.1. Singular Value Decomposition

In a sense, the SVD may be thought about as a generalisation of an eigendecomposition to non-square matrices.[122] The fundamental equation of SVD is

$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} \tag{75}$$

thus any real or complex matrix $\mathbf{D}$ (of dimensions $m \times n$) may be factorised in the form $\mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}$. This is represented pictorially in Figure 45.

$\mathbf{U}$ is an $m \times n$ matrix as well, whilst $\mathbf{S}$ and $\mathbf{V}^{\mathrm{T}}$ are $n \times n$ matrices. The matrices $\mathbf{U}$ and $\mathbf{V}$ are orthonormal vectors, i.e.

$$\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I} \tag{76}$$

and are usually referred to as the *left-singular vectors* and *right-singular vectors* of $\mathbf{D}$, respectively. $\mathbf{U}$ are a set of orthonormal eigenvectors of $\mathbf{D}\mathbf{D}^{\mathrm{T}}$ and $\mathbf{V}$ are a set of orthonormal eigenvectors of $\mathbf{D}^{\mathrm{T}}\mathbf{D}$. $\mathbf{S}$ is a diagonal matrix and the diagonal contains the *singular values* s_i in descending order. The singular values of a real matrix

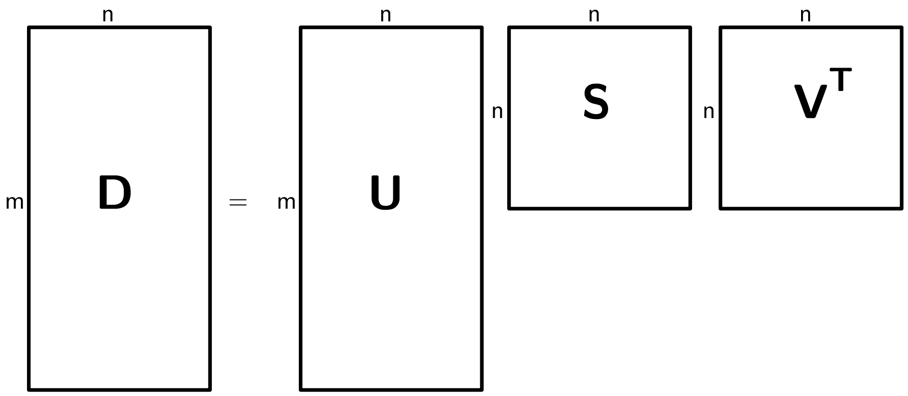


Figure 45. A pictorial representation of SVD.

are always positive and real, and the non-zero singular values are the square roots of the non-zero eigenvalues of both $\mathbf{DD}^{\mathrm{T}}$ and $\mathbf{D}^{\mathrm{T}}\mathbf{D}$. An extremely large amount of information may be extracted from the SVD - for a full description of a large number of its uses in chemistry, Chapter 5 of the Maeder and Neuhold book is recommended.[65] We will simply discuss using SVD to deduce the *rank* of a matrix, and to reconstruct a matrix.

The *rank* of a matrix is the number of linearly independent rows or columns in said matrix. The rank of a matrix may thus be useful as a way of estimating the number of components in a data set for further analysis or indeed for identifying the number of chemical/physical states involved in the photochemical/photophysical process. As the singular values are ordered in decreasing magnitude the eigenvectors continually lose importance and once the singular values are small enough, their contribution may be safely ignored. Thus in the ideal case of noiseless data, the 'small' singular values are zero and thus the number of singular values of importance gives us an estimation of the rank. Hence, a powerful result of SVD is that it enables us to estimate the matrix rank, as well as the influence of the noise. As a motivating example, let us examine a simulated dataset of transient absorption data and the use of SVD to estimate its rank. The simulation is of an intersystem crossing process using spectral shapes for the fluorescence, singlet excited state absorption and triplet excited state absorption described using equations **??** - **??**, and is shown in Figure 46. Notably, this is a simulated data set containing two spectral components where bilinearity holds perfectly.

If one decomposes this data matrix using SVD assuming different SNR, one sees that the singular values (Figure 47) are less and less easily separated as the noise on the data is increased - thus, the SVD may also tell us about the noise on our data. The eigenvectors also become similarly 'noisier' as the noise on the data is increased.[65] Thus, the SVD has given us information about the structure of our data - the *rank*, which we can use to further refine models, and the noise level.

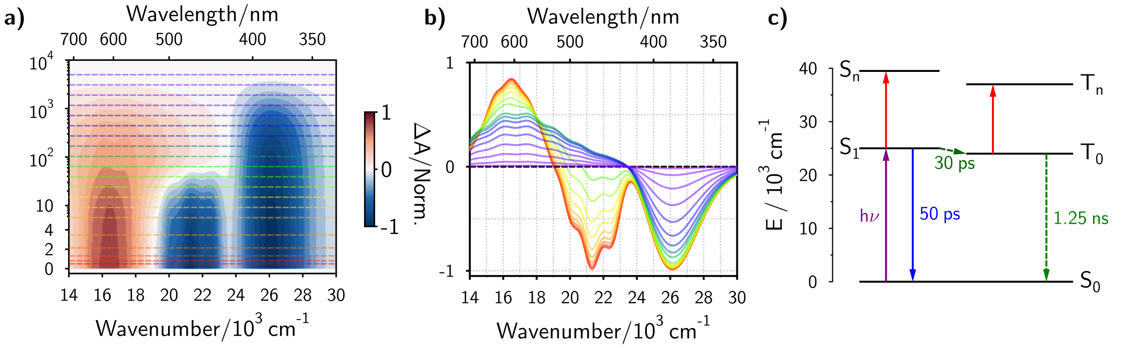


Figure 46. **a)** Contour plot of simulated TA data. **b)** Spectra at times indicated as cuts in **a**. **c)** model used to simulate the dynamics of the data. Spectra are simulated using equations **??** and **??** using the parameters in Table A1.

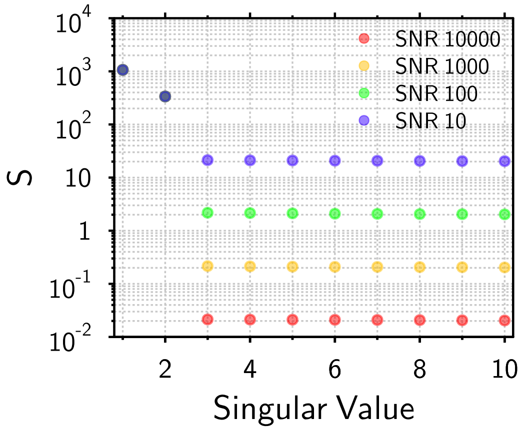


Figure 47. Singular values from SVD of data from matrix shown in Figure 46, with different levels of noise added.

This, it must be noted, is predicated on the assumption that the noise in the data is uniform and normal in character, and that the assumption of bilinearity holds. If this is not the case, structured noise may appear at a larger singular value than a 'genuine' contribution to the data.[94] A demonstration of this may be seen below - if one performs SVD on a system which shifts with a $\Delta\nu$ of $1600\,\mathrm{cm}^{-1}$ and single time constant of 2 ps, (an oversimplification of the complex time behaviour of solvent relaxation [21]) from an unrelaxed fluorescent state to a relaxed one (simulated data shown in Figure 48) the number of components found by the SVD is highly dependent on the noise level, with less of a clear separation than in the case where bilinearity holds (Figure 49). Thus its use as a method to detect the number of independent components in data should be done only if one is confident that bilinearity applies to the area where one performs the SVD.

This brings us to the other common use of SVD - reconstruction. If one can clearly see (for example, in the case of the data without noise in Figure 47) that one has a

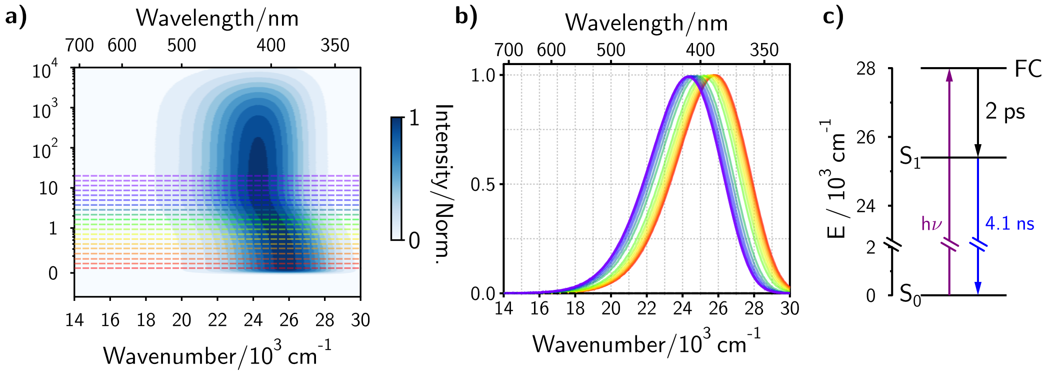


Figure 48. **a)** Contour plot of simulated time-resolved fluorescence data. **b)** Spectra at times indicated as cuts in **a**. **c)** model used to simulate the dynamics of the data. Spectra are simulated using equation **??** using the parameters in Table A2.

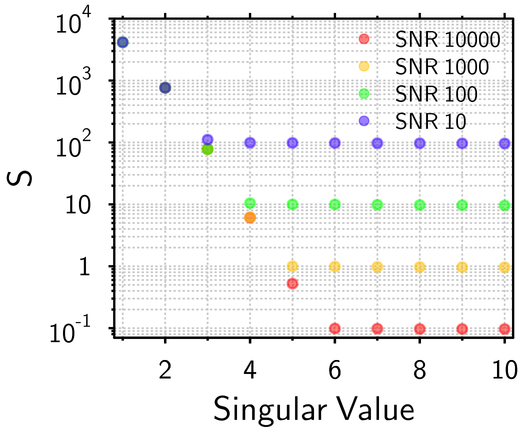


Figure 49. Singular values from SVD of data from matrix shown in Figure 48, with different levels of noise added.

number of significant singular values in a data set that are below the total number of singular values, then it is possible to ignore the contribution of these singular values, thus

$$\tilde{\mathbf{D}} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^{\mathrm{T}} \tag{77}$$

where the tildes represent that a reduced number of singular values have been used in the reconstruction. This is demonstrated graphically in Figure 50.

This is an attractive feature for the ‘removal’ of noise, at first glance. However, one loses the statistical information that one has gained by putting the effort into doing an experiment (and, as stated before, if the noise on the experiment is non-normal, one may be discarding ‘real’ contributions). Thus, whilst this is possibly an acceptable way of ‘cleaning up’ data before presenting it in a Figure (though we would argue that this is still not a good thing to do) to reconstruct a matrix using SVD and then to analyse this data (as has been done by some groups [143]) is not recommended.

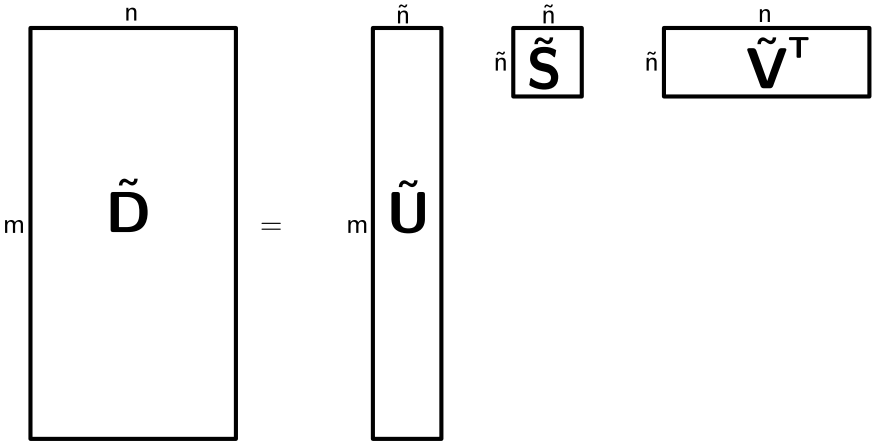


Figure 50. A pictorial representation of SVD using the reduced number of dimensions.

It should be noted that the rank one obtains from SVD is an **estimate** - any linear dependencies between the concentrations/spectra of the components will cause the matrix to be so-called *rank-deficient*, i.e. that the rank of the matrix will be lower than the 'true' number of components in the system. In addition, as difference spectra always contain the initial state spectrum, a rank-deficient data set is automatically obtained. Thus, in transient spectroscopy where difference spectra are obtained the matrix will be *rank-deficient* irrespective of the process being *full-rank* or *rank-deficient*.[144] Maeder and Ruckebusch have suggested the use of evolving factor analysis (EFA), essentially a technique where SVD is used iteratively as one increases the number of time points of one's matrix, as a confirmatory tool.[31, 65]

5.1.2. MCR-ALS

Multivariate curve resolution (MCR) methods make the assumption that the data may be described by a bilinear model (as in 37) but that no specific mathematical function is known for the $\mathbf{C}(\mathrm{t})$ or $\mathbf{S}(\tilde{\nu})$ matrices.[31] Generally the MCR model is written as

$$\mathbf{D} = \mathbf{S}(\tilde{\nu}) \cdot \mathbf{C}(t) + \mathbf{E} \tag{78}$$

where $\mathbf{C}(\mathrm{t})$ or $\mathbf{S}(\tilde{\nu})$ are defined as in 37, and $\mathbf{E}$ is the variation unexplained by the bilinear model. This can be seen as being related to SVD by e.g. the $\mathbf{C}(\mathrm{t})$ effectively being represented by $\mathbf{S}\cdot\mathbf{V}^{\mathrm{T}}$. The alternating least squares algorithm implementation of MCR (MCR-ALS) is very widely used due to its flexibility - one may incorporate a wide variety of prior knowledge in one's analysis, and one may apply the algorithm to multiple sets of data simultaneously with ease.[31, 65] The prior knowledge is incorporated as constraints (e.g. concentrations/spectra must be non-negative, unimodal and/or exactly equal to a known profile) and constraints may be imposed on the spectral, concentration or both/neither domain. Under these constraints, the MCR-ALS approach alternately updates the spectral and the concentration profiles to reduce the **E** matrix to some convergence condition.[65]

5.2. Published Implementations

SVD is implemented in virtually every programming language, and generally in a highly computationally efficient manner. The most widely-used program for MCR-ALS (321 references according to Google Scholar at time of writing) has been published by de Juan *et al.* and is written in MATLAB, with the code freely available.[145] Due to the wide applicability of MCR-ALS in chemometrics, the software is relatively developed, a graphical user interface provided and tutorials available.[146]

5.3. Applications

The use of SVD as a rank estimator prior to GKA is recommended by van Stokkum, among others [78, 94] and thus its usage in this fashion is sufficiently pervasive that to report on all instances of this would be virtually impossible. A significant example of the usage of SVD in the analysis of transient spectroscopy data comes from Ernsting *et al.*,[105] where SVD was used in order to separate the kinetic process from the wavepacket-induced oscillations of the spectrum. In short, this was accomplished by constructing a $K \times K$ rotation matrix $\mathbf{R}$ such that the data $\mathbf{D}$ may be described by

$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{R}^{-1}\mathbf{R}\mathbf{V}^{\mathrm{T}}. \tag{79}$$

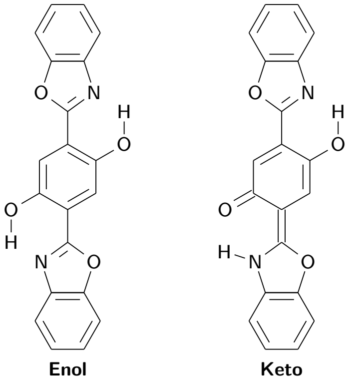


Figure 51. Enol and Keto form of **BBXHQ**.

Using this, they demonstrated that by using Fourier transforms on the rows of $\mathbf{V}^{\mathrm{T}}$, they could iteratively fill in the rotation matrix $\mathbf{R}$ to 'sweep' the oscillations arising from both coherent excitation of vibrations and from pump-induced artefacts into row K of $\mathbf{R}\mathbf{V}^{\mathrm{T}}$. This enabled them to separate out the kinetic modelling and discussion of the oscillations. By separating these out, they showed that the excited state proton transfer of the molecule 2,5-bis(2'-benzoxazolyl)hydroquinone (**BBXHQ**, Figure 51) is irreversible and occurs during the first half-period of an in-plane bending motion of the molecule.[105] This insight was only able to be accomplished by the comprehensive modelling of the spectra i.e. the simultaneous (but separated) analysis of the kinetics and the oscillations. Notably, here, the SVD kinetics were fit (leading to, as discussed earlier, a loss of statistical information) but the transformations involved in the fit could be applied also to the experimental data and compared, leading to a comparison with the 'real' data (Figure 52), which Ernsting *et al.* stated showed the approximations taken were reasonable.

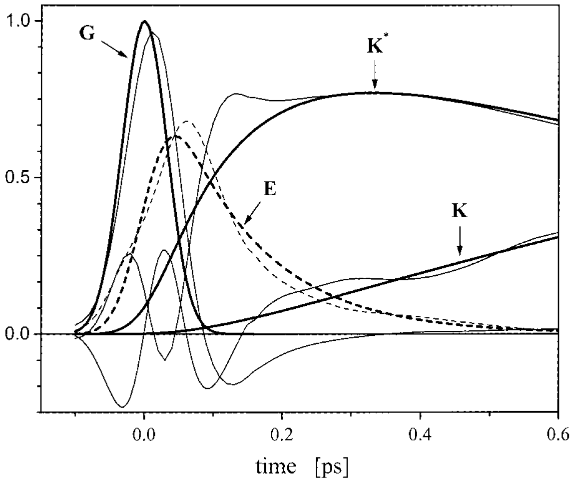


Figure 52. Population dynamics for the fit of an irreversible proton transfer model Enol → Keto* →. Fit using SVD is thick lines, transformation applied to experimental data is thin lines. **E** represents enol, **K*** excited keto immediately after reaction, and **K** the relaxed keto form. **G** represents instrument response function. Oscillations in thin lines between **K** and **K*** is due to the similarity in their spectra. Reprinted with permission from reference [105]. Copyright 2001 American Chemical Society.

The use of MCR-ALS in transient spectroscopy is not widespread - to our knowledge, it is confined to the group of Ruckebusch and collaborators. Nonetheless, they have applied it to a number of spectroscopic problems, some of which will be discussed here. For a more thorough overview of their work, the review in reference [31] is instructive. Here, a number of molecules where the group has applied MCR-ALS will be discussed (the molecules are detailed in Figure 53).

Aloïse *et al.* applied MCR-ALS to benzophenone and 4-methoxybenzophenone in an attempt to clarify the thorny issue of how the molecule benzophenone undergoes such a fast ISC ($\sim 10\,\text{ps}$ characteristic time) from an $S_1(n,\pi^*)$ state to an $T_1(n,\pi^*)$ state, in clear violation of El-Sayed's rule.[147] With the imposition of a kinetic model as a constraint on the MCR, they observed 3 states to be necessary to describe the photophysics, with the kinetic equation applied being shown in Figure 54.

When they compared the results of the 4-methoxybenzophenone (which has a solvent-dependent ordering of the $T_2(\pi,\pi^*)$ and $T_1(n,\pi^*)$ states) they observed striking similarities of the spectral shapes between the intermediate state in of 4-methoxybenzophenone in cyclohexane and of the benzophenone in all solvents. As the ordering of the triplet states is similar in cyclohexane, this pointed to a direct comparison being valid. When the solvent polarity was changed to bring the π,π^* triplet state lower than the n,π^* triplet (Figure 55), the authors observed

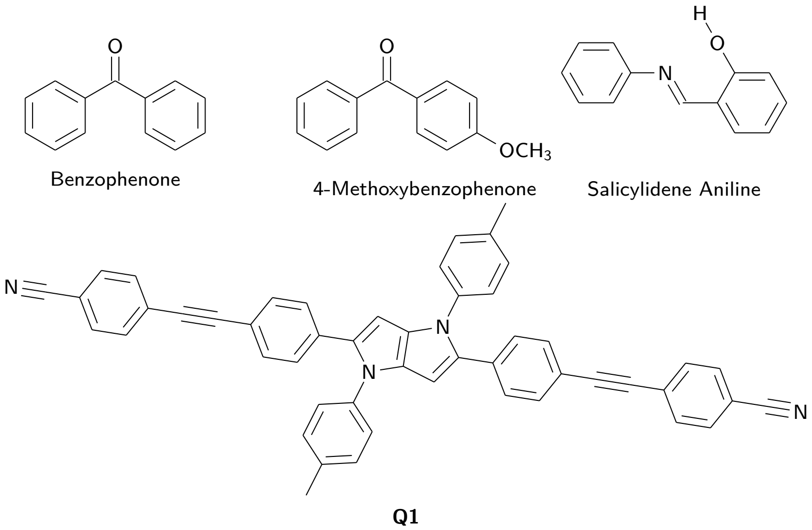


Figure 53. Molecules where MCR-ALS has been used to assist in the understanding of their photochemistry/photophysics.

$$\mathsf{S}_1(\mathsf{n},\pi^*) \xrightarrow{\qquad\quad} \mathsf{IS} \xrightarrow{\qquad\quad} \mathsf{T}_1(\mathsf{n},\pi^*).$$

Figure 54. Kinetic scheme used by Aloïse *et al.*[147] IS is defined as an unknown intermediate state.

no necessary intermediate state and thus concluded that the $T_2(\pi,\pi^*)$, either by vibronic interaction with the $T_1(n,\pi^*)$ state or by direct population (the simple kinetic model was not able to distinguish between these two possibilities), played a crucial role in the photophysics of benzophenone,[147] as anticipated previously.

The combination of MCR-ALS and a formal kinetic model is perhaps less impressive if the process under investigation is significantly obscured by the IRF, and occurs in an ultrafast manner. This is the case for salicylidene aniline (Figure 53), which Mouton *et al.* investigated. The authors themselves point out that a sum of exponential functions is likely inappropriate for processes occurring on such a fast timescale,[148] but do not discuss the fact that their deconvolution procedure assumes a single Gaussian width for the IRF, which (as discussed by Fita *et al.*[90] and shown for data measured in our laboratory in Figure 4) is inaccurate and likely leads to distortion of the ensuing 'intrinsic' deconvolved time profile.[148]

A power of MCR-ALS (as discussed above) is the ease of which one may simultaneously model data from multiple experiments. This was used to further understand excited-state symmetry breaking by Dereka *et al.*[149] Time-resolved infrared (TRIR) spectroscopy [69] and time-resolved fluorescence spectroscopy [70] had previously been used to study these systems, with the quadrupolar, intermediate and dipolar states showing clear and distinct spectral signatures in the TRIR [69] and the fluorescence transition dipole moment decreasing upon ES-SB. Transient UV-Visible Absorption however, had not been shown to be unambiguously sensitive to the symmetry breaking process. By simultaneously analysing TRIR and Transient UV-Visible Absorption data on the same system (**Q1**, Figure 53)

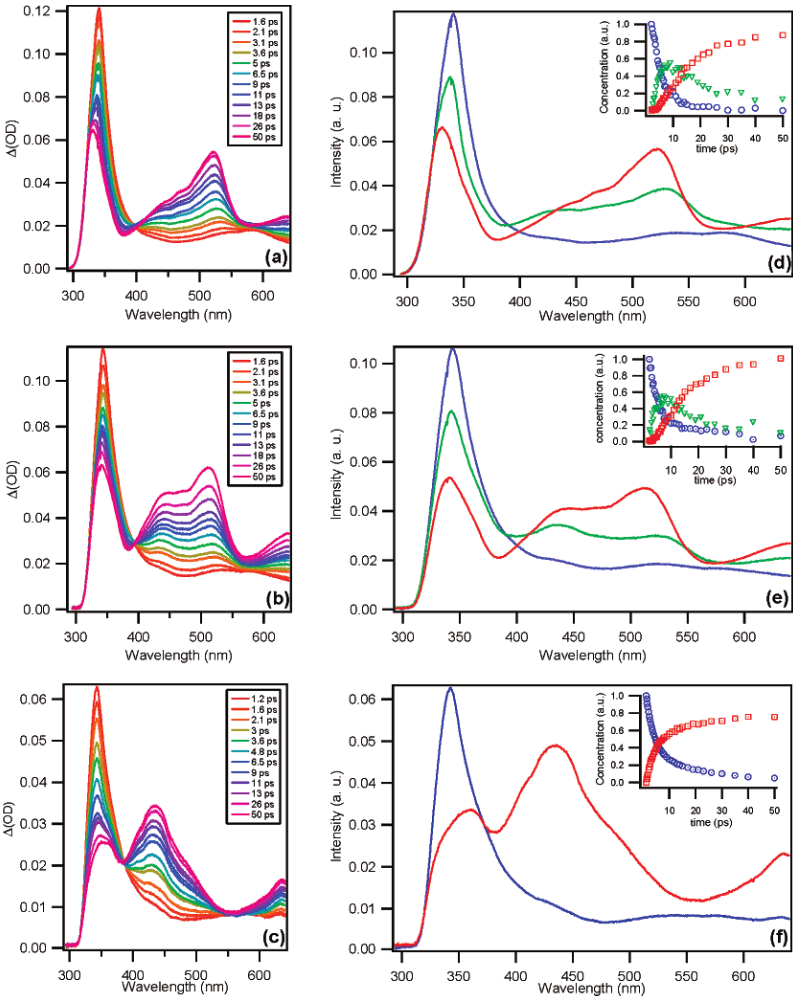


Figure 55. TA spectra, decomposed spectra from MCR-ALS and concentration profiles of decomposed states for 4-Methoxybenzophenone in **a)** and **d)** cyclohexane, **b)** and **e)** acetonitrile, and **a)** and **d)** 3:1 water/acetonitrile. Reprinted with permission from reference [147]. Copyright 2008 American Chemical Society.

distinct spectral signatures were able to be extracted from the visible TA data that corresponded to the timescale of symmetry breaking. They showed that the main spectral changes upon symmetry breaking are red-shifts of the SE that may not be unambiguously attributed to symmetry breaking without the multi-set analysis. Thus, here the MCR-ALS assisted in the assignment of symmetry breaking to a transient spectroscopy technique where it had not been hitherto unambiguously observed (Figure 56).

5.4. Conclusions

SVD and MCR-ALS are powerful approaches for e.g. rank estimation, error estimation and decomposition of the data where *a priori* nothing is known about the chemical system under investigation. These may feed into information for a more detailed and specific model (e.g. the models of Section 3 or 4) or may themselves be

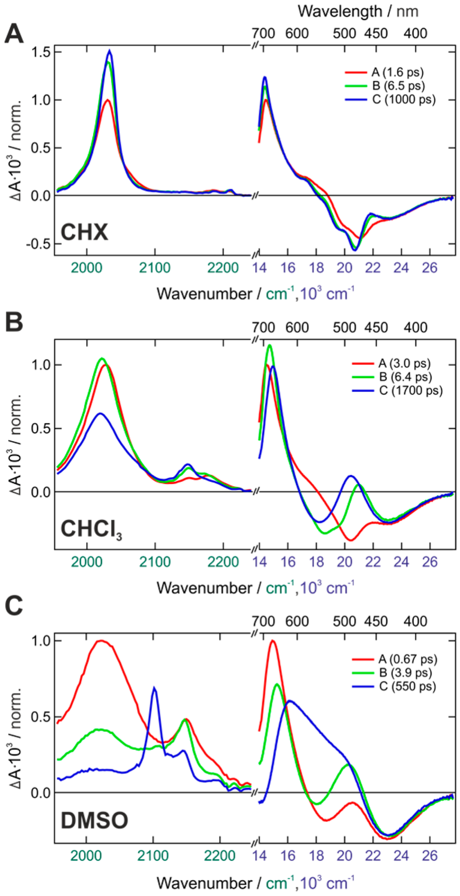


Figure 56. Spectra from the TRIR (left) and UV–vis TA (right), along with corresponding time constants obtained from the multiset analysis using the MCR-ALS approach with a kinetic constraint (A → B → C → GS) applied on the time-dependent concentration profiles. **A** is in cyclohexane (no ES-SB), **B** in chloroform (partial ES-SB), **C** in DMSO (complete ES-SB). Reprinted with permission from reference [149]. Copyright 2017 American Chemical Society.

sufficient to fully describe the data. It should be noted here however that as the MCR-ALS model makes the explicit assumption of bilinearity, this approach will provide an inaccurate representation of the number of processes involved in a photophysical/photochemical process if vibrational relaxation and/or solvation are observed in the data to a significant degree. This can be taken explicitly into account, as shown by the work of Ernsting *et al.*,[105] but requires detailed analysis and care to be taken.

6. Overall Conclusions

It is hoped that this review has been an informative jaunt through the many methods of data analysis available to the experimentalist in the quest to extract more information from one's data, and as to how to use these methods. All of the methods have their strengths and weaknesses, and thus a direct comparison is not necessarily applicable. Nonetheless, for general guidelines an attempt will be made.

Band-shape analysis models (Section **??**) have been shown to be highly useful in the description of non-exponential processes (e.g. vibrational relaxation, solvation) and in fluorescence more generally. As no bilinearity is assumed, these methods can be directly applied to these problems and (if the appropriate model is employed) give some physical insight into the origin of the absorption/emission.[73] A difficulty of these methods is that if one has a large number of overlapping bands, separation may be difficult. With the advent of the empirical decomposition scheme of Angulo *et al.*[61] and a proposal for the easy removal of the steady-state contributions [63, 64] these methods will hopefully become of more use in transient absorption spectroscopy. This would be a boon to this field, as this would enable extraction of clear populations of excited states, photoproducts and the decoupling of these from solvation and vibrational relaxation.

Global kinetic models (Section 3) were shown to be of great use if one may assume bilinearity and a classical kinetic process - the global description makes a great improvement over single-wavelength analysis in extracting close-by rate components (Figure 25) and has been shown to greatly assist in e.g. unambiguously discriminating between anisotropy models in TCSPC [85] and in deducing the full photocycle of a molecular motor.[64] However, when the assumption of bilinearity may not be made and non-exponential dynamics interfere with the data matrix, the results may be severely perturbed (Figure 49) - thus one should be sure that a classical kinetic model is an accurate representation of the process and that vibrational relaxation and/or solvation do not contribute before staking one's life on the time constants extracted.

Lifetime density models (Section 4) can be powerful if one anticipates that one will have a rate distribution in the process under investigation, and as a model that makes 'minimal' assumptions other than that the process may be described in a classical kinetic (i.e. sum of exponentials) manner. They have been shown to enable the discrimination of protein environments and show that the solvation in proteins is likely solvent-exposure related, not secondary-structure related (Figure 42).[141] As the method does not fit the data 'globally' it does not suffer from the bilinearity problem however this does mean close-by kinetic contributions would be difficult to distinguish. In addition, these models are computationally demanding and more intensive from a programming perspective. However, with more published implementations and discussions on their speed,[120, 134] improvement in this direction is expected.

Soft-modelling methods (Section 5) are powerful as an ‘initial glance’ at the data, without any explicit assumptions of classical kinetics. Their assumption of bilinearity may cause significant problems if one endeavours to use them as a method to assign a number of anticipated ‘states’ to data where this assumption does not hold. Nonetheless, they can be shown to be of use in e.g. the separation of ‘pure’ kinetic contributions from oscillatory contributions [105] and also for the ease of the analysis of multiple datasets simultaneously.[149] In addition, if the band-shape is explicitly known, accurate populations may be recovered.

Acknowledgements

Dr Arnulf Rosspeintner, Dr Alexandre Fürstenberg and Dr Bernhard Lang are warmly thanked for discussions and proof-reading. Simulations were done on the Baobab cluster at the University of Geneva.

References

[1] R. G. W. Norrish and G. Porter, Nature **164** (4172), 658 (1949).
[2] G. Porter, Proc. R. Soc. Lond. A **200** (1061), 284–300 (1950).
[3] K. J. Laidler, *Chemical Kinetics*, 3rd ed. (Harper & Row, New York, 1987).
[4] P. L. Houston, *Chemical Kinetics and Reaction Dynamics* (Dover Publications, Mineola, N. Y, 2006; OCLC: ocm70122619), OCLC: ocm70122619.
[5] S. A. Kovalenko, A. L. Dobryakov, J. Ruthmann, and N. P. Ernsting, Phys. Rev. A **59** (3), 2369–2384 (1999).
[6] M. Gerecke, G. Bierhance, M. Gutmann, N. P. Ernsting, and A. Rosspeintner, Rev. Sci. Instrum. **87** (5), 053115 (2016).
[7] K. Tominaga, G. C. Walker, W. Jarzeba, and P. F. Barbara, J. Phys. Chem. **95** (25), 10475–10485 (1991).
[8] S. F. Swallen, K. Weidemaier, H. L. Tavernier, and M. D. Fayer, J. Phys. Chem. **100** (20), 8106–8117 (1996).
[9] K. Weidemaier, H. L. Tavernier, S. F. Swallen, and M. D. Fayer, J. Phys. Chem. A **101** (10), 1887–1902 (1997).
[10] G. Angulo, J. Jedrak, A. Ochab-Marcinek, P. Pasitsuparoad, C. Radzewicz, P. Wnuk, and A. Rosspeintner, J. Chem. Phys. **146** (24), 244505 (2017).
[11] A. G. Pour, C. N. Lincoln, V. Perlík, F. Šanda, and J. Hauer, Phys. Chem. Chem. Phys. **19** (36), 24752–24760 (2017).
[12] V. Balevičius Jr, T. Wei, D. D. Tommaso, D. Abramavicius, J. Hauer, T. Polívka, and C. D. P. Duffy, Chem. Sci. (2019).
[13] R. G. Fedunov, A. V. Plotnikova, A. I. Ivanov, and E. Vauthey, J. Phys. Chem. A **121** (2), 471–481 (2017).
[14] C. A. Rumble and E. Vauthey, Phys. Chem. Chem. Phys. **21** (22), 11797–11809 (2019).
[15] B. Lang, Rev. Sci. Instrum. **89** (9), 093112 (2018).
[16] U. Megerle, I. Pugliesi, C. Schriever, C. F. Sailer, and E. Riedle, Appl. Phys. B **96** (2-3), 215–231 (2009).
[17] X. X. Zhang, C. Würth, L. Zhao, U. Resch-Genger, N. P. Ernsting, and M. Sajadi, Rev. Sci. Instrum. **82** (6), 063108 (2011).
[18] E. Tokunaga, A. Terasaki, and T. Kobayashi, J. Opt. Soc. Am. B **13** (3), 496 (1996).
[19] N. Boens, W. Qin, N. Basarić, J. Hofkens, M. Ameloot, J. Pouget, J. P. Lefèvre, B. Valeur, E. Gratton, M. vandeVen, N. D. Silva, Y. Engelborghs, *et al.*, Anal. Chem. **79** (5), 2137–2149 (2007).
[20] K. Ekvall, P. van der Meulen, C. Dhollande, L. E. Berg, S. Pommeret, R. Naskrecki, and J. C. Mialocq, J. Appl. Phys. **87** (5), 2340–2352 (2000).
[21] M. L. Horng, J. A. Gardecki, A. Papazyan, and M. Maroncelli, J. Phys. Chem. **99** (48), 17311–17337 (1995).

[22] B. Lang, G. Angulo, and E. Vauthey, J. Phys. Chem. A **110** (22), 7028–7034 (2006).

[23] W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques* (Springer, Berlin; New York, 2005; OCLC: 262681319), OCLC: 262681319.

[24] M. Lorenc, M. Ziolek, R. Naskrecki, J. Karolczak, J. Kubicki, and A. Maciejewski, Appl. Phys. B **74** (1), 19–27 (2002).

[25] M. Rasmusson, A. N. Tarnovsky, E. Akesson, and V. Sundström, Chem. Phys. Lett. **335** (3), 201–208 (2001).

[26] Z. Vardeny and J. Tauc, Opt. Commun. **39** (6), 396–400 (1981).

[27] M. Joffre, D. Hulin, A. Migus, A. Antonetti, C. B. à la Guillaume, N. Peyghambarian, M. Lindberg, and S. W. Koch, Opt. Lett. **13** (4), 276–278 (1988).

[28] C. Rullière, editor, *Femtosecond Laser Pulses: Principles and Experiments*, 2nd ed. Advanced Texts in Physics (Springer, New York, 2005).

[29] J. P. Likforman, M. Joffre, G. Chériaux, and D. Hulin, Opt. Lett. **20** (19), 2006–2008 (1995).

[30] O. Devos, N. Mouton, M. Sliwa, and C. Ruckebusch, Anal. Chim. Acta **705** (1), 64–71 (2011).

[31] C. Ruckebusch, M. Sliwa, P. Pernot, A. de Juan, and R. Tauler, J. Photochem. Photobiol. C **13** (1), 1–27 (2012).

[32] E. T. Whittaker, P. Edinburgh. Math. Soc. **41**, 63–75 (1922).

[33] P. H. C. Eilers, Anal. Chem. **75** (14), 3631–3636 (2003).

[34] L. Zhao, J. L. P. Lustres, V. Farztdinov, and N. P. Ernsting, Phys. Chem. Chem. Phys. **7** (8), 1716–1725 (2005).

[35] D. A. Turton and K. Wynne, J. Chem. Phys. **128** (15), 154516 (2008).

[36] S. Yamaguchi and H. O. Hamaguchi, Appl. Spectrosc. **49** (10), 1513–1515 (1995).

[37] M. Kubista, R. Sjöback, S. Eriksson, and B. Albinsson, Analyst **119** (3), 417–419 (1994).

[38] J. A. Gardecki and M. Maroncelli, Appl. Spectrosc. **52** (9), 1179–1189 (1998).

[39] G. Eppeldauer, J. Res. Natl. Inst. Stand. Technol. **103** (6), 615–619 (1998).

[40] U. Resch-Genger, D. Pfeifer, C. Monte, W. Pilz, A. Hoffmann, M. Spieles, K. Rurack, J. Hollandt, D. Taubert, B. Schönenberger, and P. Nording, J. Fluoresc. **15** (3), 315–336 (2005).

[41] D. Pfeifer, K. Hoffmann, A. Hoffmann, C. Monte, and U. Resch-Genger, J. Fluoresc. **16** (4), 581–587 (2006).

[42] A. Ishizaki and G. R. Fleming, Proc. Natl. Acad. Sci. U.S.A. **106** (41), 17255–17260 (2009).

[43] M. A. C. Saller, A. Kelly, and J. O. Richardson, Faraday Discuss. **221** (0), 150–167 (2019).

[44] J. B. Birks and D. J. Dyson, Proc. Royal Soc. A **275** (1360), 135–148 (1963).

[45] G. Angulo, G. Grampp, and A. Rosspeintner, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. **65** (3-4), 727–731 (2006).

[46] C. A. Rumble and M. Maroncelli, J. Chem. Phys. **148** (19), 193801 (2018).

[47] F. Bloch, Phys. Rev. **70** (7-8), 460–474 (1946).

[48] R. D. B. Fraser and E. Suzuki, Anal. Chem. **41** (1), 37–39 (1969).

[49] X. Li, M. Liang, A. Chakraborty, M. Kondo, and M. Maroncelli, J. Phys. Chem. B **115** (20), 6592–6607 (2011).

[50] C. M. Metzler, A. E. Cahill, S. Petty, D. E. Metzler, and L. Láng, Appl. Spectrosc. **39** (2), 333–339 (1985).

[51] D. B. Siano and D. E. Metzler, J. Chem. Phys. **51** (5), 1856–1861 (1969).

[52] D. B. Siano, J. Chem. Educ. **49** (11), 755 (1972).

[53] R. A. Marcus, J. Phys. Chem. **93** (8), 3078–3086 (1989).

[54] Y. Fujimura, K. Sato, and T. Nakajima, Bull. Chem. Soc. Jpn. **55** (1), 277–281 (1982).

[55] S. Arzhantsev, K. A. Zachariasse, and M. Maroncelli, J. Phys. Chem. A **110** (10), 3454–3470 (2006).

[56] R. G. Fedunov, I. P. Yermolenko, A. E. Nazarov, A. I. Ivanov, A. Rosspeintner, and G. Angulo, J. Mol. Liq. p. 112016 (2019).

[57] P. Sulzer and K. Wieland, Helv. Phys. Acta **25**, 653–676 (1952).

[58] O. Schalk, J. P. Yang, A. Hertwig, H. Hippler, and A. N. Unterreiner, Mol. Phys. **107** (20), 2159–2167 (2009).

[59] M. Sajadi and N. P. Ernsting, J. Phys. Chem. B **117** (25), 7675–7684 (2013).

[60] T. Kumpulainen, A. Rosspeintner, and E. Vauthey, Phys. Chem. Chem. Phys. **19** (13), 8815–8825 (2017).

[61] G. Angulo, A. Rosspeintner, B. Lang, and E. Vauthey, Phys. Chem. Chem. Phys. **20** (39), 25531–25546 (2018).

[62] A. V. Deshpande, A. Beidoun, A. Penzkofer, and G. Wagenblast, Chem. Phys. **142** (1), 123–131 (1990).

[63] S. A. Kovalenko, R. Schanz, V. M. Farztdinov, H. Hennig, and N. P. Ernsting, Chem. Phys. Lett. **323** (3), 312–322 (2000).

[64] R. Wilcken, M. Schildhauer, F. Rott, L. A. Huber, M. Guentner, S. Thumser, K. Hoffmann, S. Oesterling, R. de Vivie-Riedle, E. Riedle, and H. Dube, J. Am. Chem. Soc. **140** (15), 5311–5318 (2018).

[65] M. Maeder and Y. M. Neuhold, *Practical Data Analysis in Chemistry*, 1st ed. No. 26 in Data Handling in Science and Technology (Elsevier, Amsterdam, 2007; OCLC: 255942082), OCLC: 255942082.

[66] M. P. Grubb, A. J. Orr-Ewing, and M. N. R. Ashfold, Rev. Sci. Instrum. **85** (6), 064104 (2014).

[67] M. Gerecke, C. Richter, M. Quick, I. N. Ioffe, R. Mahrwald, S. A. Kovalenko, and N. P. Ernsting, J. Phys. Chem. B **121** (41), 9631–9638 (2017).

[68] E. Heid and C. Schröder, J. Phys. Chem. B **121** (41), 9639–9646 (2017).

[69] B. Dereka, A. Rosspeintner, Z. Li, R. Liska, and E. Vauthey, J. Am. Chem. Soc. **138** (13), 4643–4649 (2016).

[70] J. S. Beckwith, A. Rosspeintner, G. Licari, M. Lunzer, B. Holzer, J. Fröhlich, and E. Vauthey, J. Phys. Chem. Lett. **8** (23), 5878–5883 (2017).

[71] W. Kim, J. Sung, M. Grzybowski, D. T. Gryko, and D. Kim, J. Phys. Chem. Lett. **7** (15), 3060–3066 (2016).

[72] T. Kim, W. Kim, H. Mori, A. Osuka, and D. Kim, J. Phys. Chem. C **122** (34), 19409–19415 (2018).

[73] T. Kasajima, S. Akimoto, S. i. Sato, and I. Yamazaki, J. Phys. Chem. A **108** (16), 3268–3275 (2004).

[74] A. Pigliucci, G. Duvanel, L. M. L. Daku, and E. Vauthey, J. Phys. Chem. A **111** (28), 6135–6145 (2007).

[75] A. Rosspeintner, M. Koch, G. Angulo, and E. Vauthey, J. Phys. Chem. Lett. pp. 7015–7020 (2018).

[76] C. Weinert, B. Wezisla, J. Lindner, and P. Vöhringer, Phys. Chem. Chem. Phys. **17** (20), 13659–13671 (2015).

[77] R. D. Levine, *Molecular Reaction Dynamics*, paperback edition ed. (Cambridge University Press, Cambridge, 2009; OCLC: 699260402), OCLC: 699260402.

[78] I. H. M. van Stokkum, D. S. Larsen, and R. van Grondelle, Biochim. Biophys. Acta **1657** (2–3), 82–104 (2004).

[79] X. X. Zhang, M. Liang, N. P. Ernsting, and M. Maroncelli, J. Phys. Chem. Lett. **4** (7), 1205–1210 (2013).

[80] H. Marciniak and S. Lochbrunner, Chem. Phys. Lett. **609**, 184–188 (2014).

[81] A. A. Istratov and O. F. Vyvenko, Rev. Sci. Instrum. **70** (2), 1233–1257 (1999).

[82] J. S. Beckwith, B. Lang, J. Grilj, and E. Vauthey, J. Phys. Chem. Lett. **10** (13), 3688–3693 (2019).

[83] M. J. Feldstein, P. Vöhringer, and N. F. Scherer, J. Opt. Soc. Am. B **12** (8), 1500–1510 (1995).

[84] J. R. Knutson, J. M. Beechem, and L. Brand, Chem. Phys. Lett. **102** (6), 501–507 (1983).

[85] J. M. Beechem, M. Ameloot, and L. Brand, Instrum. Sci. Technol. **14** (3-4), 379–402 (1985).

[86] J. M. Beechem, in *Methods in Enzymology*, Vol. 210 (, , 1992), pp. 37–54.

[87] K. M. Mullen, M. Vengris, and I. H. M. van Stokkum, J. Glob. Optim. **38** (2), 201–213

(2007).

[88] K. M. Mullen and I. H. M. van Stokkum, Numer. Algorithms **51** (3), 319–340 (2009).

[89] M. N. Berberan-Santos and J. M. G. Martinho, J. Chem. Educ. **67** (5), 375 (1990).

[90] P. Fita, E. Luzina, T. Dziembowska, C. Radzewicz, and A. Grabowska, J. Chem. Phys. **125** (18), 184508 (2006).

[91] G. M. Petrov, J. Quant. Spectrosc. Radiat. Transfer **72** (3), 281–287 (2002).

[92] K. M. Mullen and I. H. M. van Stokkum, J. Stat. Softw. **18** (3) (2007).

[93] C. Slavov, H. Hartmann, and J. Wachtveitl, Anal. Chem. **87** (4), 2328–2336 (2015).

[94] L. J. G. W. van Wilderen, C. N. Lincoln, and J. J. van Thor, PLOS ONE **6** (3), e17373 (2011).

[95] S. Jin, S. Wang, Y. Song, M. Zhou, J. Zhong, J. Zhang, A. Xia, Y. Pei, M. Chen, P. Li, and M. Zhu, J. Am. Chem. Soc. **136** (44), 15559–15565 (2014).

[96] M. Zhou, S. Vdović, S. Long, M. Zhu, L. Yan, Y. Wang, Y. Niu, X. Wang, Q. Guo, R. Jin, and A. Xia, J. Phys. Chem. A **117** (40), 10294–10303 (2013).

[97] M. M. Warren, M. Kaucikas, A. Fitzpatrick, P. Champion, J. Timothy Sage, and J. J. van Thor, Nat. Commun. **4**, 1461 (2013).

[98] J. P. Bustamante, S. Abbruzzetti, A. Marcelli, D. Gauto, L. Boechi, A. Bonamore, A. Boffi, S. Bruno, A. Feis, P. Foggi, D. A. Estrin, and C. Viappiani, J. Phys. Chem. B **118** (5), 1234–1245 (2014).

[99] J. Hu, Y. Li, H. Zhu, S. Qiu, G. He, X. Zhu, and A. Xia, ChemPhysChem **16** (11), 2357–2365 (2015).

[100] C. Slavov, N. Bellakbil, J. Wahl, K. Mayer, K. Rück-Braun, I. Burghardt, J. Wachtveitl, and M. Braun, Phys. Chem. Chem. Phys. **17** (21), 14045–14053 (2015).

[101] J. Snellenburg, S. Laptenok, R. Seger, K. Mullen, and I. H. M. van Stokkum, J. Stat. Softw. **49** (3) (2012).

[102] J. E. Löfroth, J. Phys. Chem. **90** (6), 1160–1168 (1986).

[103] D. L. Rousseau, J. V. Staros, and J. M. Beechem, Biochemistry **34** (44), 14508–14518 (1995).

[104] P. Fita, E. Luzina, T. Dziembowska, D. Kopeć, P. Piątkowski, C. Radzewicz, and A. Grabowska, Chem. Phys. Lett. **416** (4), 305–310 (2005).

[105] N. P. Ernsting, S. A. Kovalenko, T. Senyushkina, J. Saam, and V. Farztdinov, J. Phys. Chem. A **105** (14), 3443–3453 (2001).

[106] M. Gueye, M. Manathunga, D. Agathangelou, Y. Orozco, M. Paolino, S. Fusi, S. Haacke, M. Olivucci, and J. Léonard, Nat. Commun. **9** (1), 313 (2018).

[107] I. H. M. van Stokkum, C. C. Jumper, J. J. Snellenburg, G. D. Scholes, R. van Grondelle, and P. Malý, J. Chem. Phys. **145** (17), 174201 (2016).

[108] S. Schott, L. Ress, J. Hrušák, P. Nuernberger, and T. Brixner, Phys. Chem. Chem. Phys. **18** (48), 33287–33302 (2016).

[109] D. R. James and W. R. Ware, Chem. Phys. Lett. **120** (4), 455–459 (1985).

[110] J. R. Alcala, E. Gratton, and F. G. Prendergast, Biophys. J. **51** (4), 597–604 (1987).

[111] D. R. James, Y. S. Liu, P. De Mayo, and W. R. Ware, Chem. Phys. Lett. **120** (4), 460–465 (1985).

[112] J. C. Brochon, A. K. Livesey, J. Pouget, and B. Valeur, Chem. Phys. Lett. **174** (5), 517–522 (1990).

[113] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to Probability Theory* The Houghton-Mifflin Series in Statistics (Houghton Mifflin, Boston, 1971).

[114] G. S. Mudholkar and A. D. Hutson, J. Stat. Plan. Inference **83** (2), 291–309 (2000).

[115] M. N. Berberan-Santos, E. N. Bodunov, and B. Valeur, Chem. Phys. **315** (1–2), 171–182 (2005).

[116] H. K. Nakamura, M. Sasai, and M. Takano, Chem. Phys. **307** (2), 259–267 (2004).

[117] O. Nicolet, N. Banerji, S. Pagès, and E. Vauthey, J. Phys. Chem. A **109** (37), 8236–8245 (2005).

[118] G. Landl, T. Langthaler, H. W. Englt, and H. F. Kauffmann, J. Comput. Phys. **95** (1), 1–28 (1991).

[119] A. K. Livesey and J. C. Brochon, Biophys J **52** (5), 693–706 (1987).

[120] G. F. Dorlhiac, C. Fare, and J. J. van Thor, PLOS Comput. Biol. **13** (5), e1005528

(2017).
[121] C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind* No. 105 in Research Notes in Mathematics (Pitman Advanced Pub. Program, Boston, 1984).
[122] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* Adaptive Computation and Machine Learning Series (MIT Press, Cambridge, MA, 2012).
[123] S. Xiong, B. Dai, J. Huling, and P. Z. G. Qian, Technometrics **58** (3), 285–293 (2016).
[124] J. Skilling and R. K. Bryan, Mon. Not. R. Astron. Soc. **211** (1), 111–124 (1984).
[125] A. K. Livesey and J. Skilling, Acta Crystallogr. A **41** (2), 113–122 (1985).
[126] J. C. Brochon, in *Methods in Enzymology, Part B: Numerical Computer Methods*, Vol. 240 (, , 1994), pp. 262–311.
[127] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, Hoboken, N.J, 2006; OCLC: ocm59879802), OCLC: ocm59879802.
[128] A. T. N. Kumar, L. Zhu, J. F. Christian, A. A. Demidov, and P. M. Champion, J. Phys. Chem. B **105** (32), 7847–7856 (2001).
[129] S. A. Vinogradov and D. F. Wilson, Appl. Spectrosc. **54** (6), 849–855 (2000).
[130] V. A. Lórenz-Fonfría and H. Kandori, Appl. Spectrosc. **60** (4), 407–417 (2006).
[131] N. Đorđević, J. S. Beckwith, M. Yarema, O. Yarema, A. Rosspeintner, N. Yazdani, J. Leuthold, E. Vauthey, and V. Wood, ACS Photonics **5** (12), 4888–4895 (2018).
[132] V. A. Lórenz-Fonfría and H. Kandori, Appl. Spectrosc. **61** (1), 74–84 (2007).
[133] J. Skilling and S. F. Gull, Lecture Notes-Monograph Series **20**, 341–367 (1991).
[134] D. A. Smith, G. McKenzie, A. C. Jones, and T. A. Smith, Methods Appl. Fluoresc. **5** (4), 042001 (2017).
[135] R. Esposito, C. Altucci, and R. Velotta, J. Fluoresc. **23** (1), 203–211 (2013).
[136] H. Satzger and W. Zinth, Chem. Phys. **295** (3), 287–295 (2003).
[137] M. A. Berg and H. Kaur, J. Chem. Phys. **146** (5), 054104 (2017).
[138] P. J. Steinbach, J. Chem. Inf. Comput. Sci. **42** (6), 1476–1478 (2002).
[139] R. Métivier, I. Leray, J. P. Lefèvre, M. Roy-Auberger, N. Zanier-Szydlowski, and B. Valeur, Phys. Chem. Chem. Phys. **5** (4), 758–766 (2003).
[140] J. R. Alcala, E. Gratton, and F. G. Prendergast, Biophys. J. **51** (6), 925–936 (1987).
[141] P. Abbyad, X. Shi, W. Childs, T. B. McAnaney, B. E. Cohen, and S. G. Boxer, J. Phys. Chem. B **111** (28), 8269–8276 (2007).
[142] C. Slavov, C. Yang, L. Schweighauser, C. Boumrifak, A. Dreuw, H. A. Wegner, and J. Wachtveitl, Phys. Chem. Chem. Phys. **18** (22), 14795–14804 (2016).
[143] S. M. Dyar, A. L. Smeigh, S. D. Karlen, R. M. Young, and M. R. Wasielewski, Chem. Phys. Lett. **629**, 23–28 (2015).
[144] L. Blanchet, C. Ruckebusch, J. P. Huvenne, and A. de Juan, Chemom. Intell. Lab. Syst. **89** (1), 26–35 (2007).
[145] J. Jaumot, A. de Juan, and R. Tauler, Chemom. Intell. Lab. Syst. **140**, 1–12 (2015).
[146] A. de Juan, J. Jaumot, and R. Tauler, Anal. Methods **6** (14), 4964–4976 (2014).
[147] S. Aloïse, C. Ruckebusch, L. Blanchet, J. Réhault, G. Buntinx, and J. P. Huvenne, J. Phys. Chem. A **112** (2), 224–231 (2008).
[148] N. Mouton, M. Sliwa, G. Buntinx, and C. Ruckebusch, J. Chemom. **24** (7-8), 424–433 (2010).
[149] B. Dereka, A. Rosspeintner, R. Stężycki, C. Ruckebusch, D. T. Gryko, and E. Vauthey, J. Phys. Chem. Lett. **8** (24), 6029–6034 (2017).
[150] S. van der Walt, S. C. Colbert, and G. Varoquaux, Comput. Sci. Eng. **13** (2), 22–30 (2011).
[151] W. McKinney, in *Proceedings of the 9th Python in Science Conference*, Austin, TX, pp. 51–56.
[152] S. Hoyer, J. Hamman, C. Fitzgerald, F. Maussion, K. Fujii, M. Roos, A. Kleeman, T. Kluyver, D. Cherian, J. Munroe, S. Clark, Zac Hatfield-Dodds, *et al.*, Pydata/Xarray v0.11.1 Zenodo 2019, Jan.
[153] J. D. Hunter, Comput. Sci. Eng. **9** (3), 90–95 (2007).
[154] M. Waskom, O. Botvinnik, P. Hobson, J. B. Cole, Y. Halchenko, S. Hoyer, A. Miles, T. Augspurger, T. Yarkoni, T. Megies, L. P. Coelho, D. Wehner, *et al.*, Seaborn: V0.5.0

(November 2014) Zenodo 2014, Nov.

[155] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, *et al.*, Nat. Methods **17** (3), 261–272 (2020).

Appendix A. Simulation Details

All simulations of data and fitting were done in Python 3.6.7 code written by the authors. The numPy,[150] pandas,[151] xarray,[152] matplotlib,[153] seaborn,[154] and sciPy [155] libraries were used. All code is available upon a request to the authors.

For simulations in which the accuracy of repeated fitting events was tested, the following procedure was used. The signal corresponding to the e.g. biexponential decays (either single decays or decays across a spectrum) was first simulated and then for each different fitting event normally distributed random noise was added corresponding to the overall SNR level. A fit to the data using the corresponding (e.g. biexponential) equation was then attempted, using the scipy.optimize.least_squares optimizer and thus optimising to the lowest

$$\left(\frac{signal - fit}{error}\right)^2.$$

The initial guesses of all parameters were set to be 10% away from the genuine parameters used in the simulation of the data. In the case of Figure 25, the spectra were simulated on a 320-point grid (320 points corresponding to the number of pixels typically filled by white light in TA experiments in our laboratory) which spanned from 14,000 to 30,000 cm^{-1} in steps of 50 cm^{-1}. The spectral shapes were simulated using equations **??** and **??** and the parameters specified in Table A1. The shorter component was made up of the Ground State Bleach, Fluorescence and ESA_1 line shapes (i.e. was computed as ESA_1 - (Ground State Bleach + Fluorescence) and the longer component was made up of the Ground State Bleach and ESA_2 (i.e. was computed as ESA_2 - Ground State Bleach).

Component	Lineshape Equation	ν^0	σ	S	ω
Ground State Bleach	**??**	25	0.8	1.5	1
Fluorescence	**??**	22.5	0.5	1.25	1.2
ESA_1	**??**	14.5	0.5	2.5	1
ESA_2	**??**	13	0.7	3	1.5

Table A1. Parameters of equations **??** and **??** used simulate absorption and emission spectra for testing fitting using equation 44 in Figure 25. All parameters reported in $10^3\,\mathrm{cm}^{-1}$.

For simulations testing the accuracy of fitting data generated from distributions of rates to a biexponential equation, the rate grid was formed of 10,000 points from 10^{-6} to 10 ps on a logarithmic grid. These were then used to simulate traces in accordance with equations 51 - 52. The initial guess was found using the scipy.differential_evolution algorithm to find the best parameters on the global fitting surface (lifetime search from 100 fs to 10 ns on both of the lifetime fitting parameters). These guesses were then given to the scipy.optimize.least_squares optimizer.

For simulations of applying SVD to shifting spectra, the spectra were also simulated on a 320-point grid spanning 14,000 to 30,000 cm^{-1}. The initial time-zero spectrum was simulated using the parameters shown in Table A2, with the ν^0 shifting with a

2 ps time constant to $26.4 \times 10^3\,\text{cm}^{-1}$. SVD was performed using the numpy.linalg.svd package.

Lineshape Equation	ν^0	σ	S	ω
??	28	1	2	1.2

Table A2. Parameters of equation **??** used simulate emission spectra for testing how many SVD-detected components were found. All parameters reported in $10^3\,\text{cm}^{-1}$.

Appendix B. Abbreviations and Acronyms

2PA	Two-**P**hoton **A**bsorption
ΔA	Change in **A**bsorption
ALS	**A**lternating **L**east **S**quares
CT	**C**harge **T**ransfer
D$_n$	**D**oublet n state
DADS	**D**ecay **A**ssociated **D**ecay **S**pectra
DAS	**D**ecay **A**ssociated **S**pectra
DHO	**D**isplaced **H**armonic **O**scillator
DMA	*N*-*N*-**Dim**ethyl**a**niline
DOAS	**D**amped **O**scillation **A**ssociated **S**pectrum
EADS	**E**volution **A**ssociated **D**ecay **S**pectra
EAS	**E**volution **A**ssociated **S**pectra
EFA	**E**volving **F**actor **A**nalysis
ESA	**E**xcited-**S**tate **A**bsorption
ES-SB	**E**xcited **S**tate **S**ymmetry **B**reaking
ET	**E**lectron **T**ransfer
FLUPS	**Fl**uorescence **Up**-conver**s**ion
FWHM	**F**ull **W**idth at **H**alf **M**aximum
GDD	**G**roup **D**elay **D**ispersion
GKA	**G**lobal **K**inetic **A**nalysis
GKM	**G**lobal **K**inetic **M**odel
GS	**G**round **S**tate
GSB	**G**round **S**tate **B**leach
GUI	**G**raphical **U**ser **I**nterface
IC	**I**nternal **C**onversion
IR	**I**nfra**r**ed
IRF	**I**nstrument **R**esponse **F**unction
ISC	**I**nter**s**ystem **C**rossing
IVR	**I**ntramolecular **V**ibrational **R**elaxation
LASSO	**L**east **A**bsolute **S**hrinkage and **S**election **O**perator
LDAC	**L**ogarithmically **D**ifferentiated **A**bsorption **C**hanges
LDM	**L**ifetime **D**istribution **M**odels
MCR	**M**ultivariate **C**urve **R**esolution
MEM	**M**aximum **E**ntropy **M**ethod
OD	**O**ptical **D**ensity
OKE	**O**ptical **K**err **E**ffect
RA	**R**aman **A**rtefact
S$_n$	**S**inglet n state
SADS	**S**pecies **A**ssociated **D**ecay **S**pectra
SAS	**S**pecies **A**ssociated **S**pectra
SE	**S**timulated **E**mission
SNR	**S**ignal-to-**N**oise **R**atio
SVD	**S**ingular **V**alue **D**ecomposition
SW	**S**ulzer-**W**ieland
T$_n$	**T**riplet n state
TA	**T**ransient **A**bsorption
TCSPC	**T**ime **C**orrelated **S**ingle **P**hoton **C**ounting
TRIR	**T**ime **R**esolved **I**nfra**r**ed
UV	**U**ltra**v**iolet
UV-Vis	**U**ltra**v**iolet-**Vis**ible
VC	**V**ibrational **C**ooling
VER	**V**ibrational **E**nergy **R**elaxation
XPM	Cross-**P**hase **M**odulation