Analytica Chimica Acta 1273 (2023) 341545

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Kernelizing: A way to increase accuracy in trilinear decomposition analysis of multiexponential signals

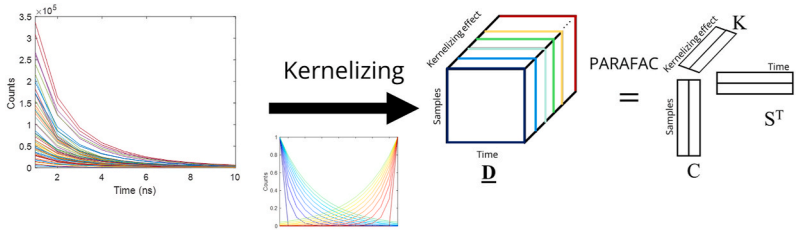
Adrián Gómez-Sánchez [a,b,*], Raffaele Vitale [b], Olivier Devos [b], Anna de Juan [a], Cyril Ruckebusch [b]

[a] Chemometrics Group, Universitat de Barcelona, Diagonal, 645, 08028, Barcelona, Spain
[b] Univ. Lille, CNRS, UMR 8516, LASIRe, Laboratoire Avancé de Spectroscopie pour Les Intéractions La Réactivité et L'Environnement, F-59000, Lille, France

HIGHLIGHTS

- A tensorization approach based on convolution of exponential decays is proposed.
- The new Kernelizing approach is able to generate trilinear from bilinear data.
- This new unmixing approach provides the pure monoexponential components.
- Kernelizing can be applied even if a low number of sampling points is only available.
- Kernelizing has been tested in simulations, experiments and FLIM images.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Prof. L. Buydens

Keywords:
Kernelizing
Multiexponential decays
Tensorization
Trilinear models
Parallel factor analysis-alternating least squares (PARAFAC-ALS)

ABSTRACT

The unmixing of multiexponential decay signals into monoexponential components using soft modelling approaches is a challenging task due to the strong correlation and complete window overlap of the profiles. To solve this problem, slicing methodologies, such as PowerSlicing, tensorize the original data matrix into a three-way data array that can be decomposed based on trilinear models providing unique solutions. Satisfactory results have been reported for different types of data, e.g., nuclear magnetic resonance or time-resolved fluorescence spectra. However, when decay signals are described by only a few sampling (time) points, a significant degradation of the results can be observed in terms of accuracy and precision of the recovered profiles.

In this work, we propose a methodology called Kernelizing that provides a more efficient way to tensorize data matrices of multiexponential decays. Kernelizing relies on the invariance of exponential decays, i.e., when convolving a monoexponential decaying function with any positive function of finite width (hereafter called "kernel"), the shape of the decay (determined by the characteristic decay constant) remains unchanged and only the preexponential factor varies. The way preexponential factors are affected across the sample and time modes is linear, and it only depends on the kernel used. Thus, using kernels of different shapes, a set of convolved curves can be obtained for every sample, and a three-way data array generated, for which the modes are sample, time and kernelizing effect. This three-way array can be afterwards analyzed by a trilinear decomposition method, such as PARAFAC-ALS, to resolve the underlying monoexponential profiles. To validate this new approach and assess its performance, we applied Kernelizing to simulated datasets, real time-resolved fluorescence spectra

* Corresponding author. Chemometrics Group, Universitat de Barcelona, Diagonal, 645, 08028, Barcelona, Spain.
E-mail address: agomezsa29@alumnes.ub.edu (A. Gómez-Sánchez).

https://doi.org/10.1016/j.aca.2023.341545
Received 14 February 2023; Received in revised form 9 June 2023; Accepted 18 June 2023
Available online 19 June 2023

collected on mixtures of fluorophores and fluorescence-lifetime imaging microscopy data. When the measured multiexponential decays feature few sampling points (down to fifteen), more accurate trilinear model estimates are obtained than when using slicing methodologies.

1. Introduction

The analysis of exponential decay signals is usually performed by multiexponential fitting approaches that allow extracting the characteristic decay constants and preexponential factors of the different monoexponential components [1]. However, multiexponential fitting remains difficult due to the high natural correlation among the monoexponential decays of these individual components and becomes even more complex for signals exhibiting a low signal-to-noise ratio [2]. Besides, the results obtained can be very user-dependent since selecting the correct number of monoexponential components and setting appropriate initial parameters for the fitting are tasks that require expertise and are often based on a trial-and-error approach.

In this context, factor analysis can constitute a good alternative to multiexponential fitting. Indeed, specific chemometric approaches are available to solve the unmixing (curve resolution) problem for exponential mixtures and have been successfully applied in e.g., Nuclear Magnetic Resonance (NMR) [3,4] and Time-Resolved Fluorescence Spectroscopy (TRFS) [5–7]. Among these approaches, PowerSlicing [8] is a method aimed at resolving mixtures of monoexponential decays, based on (i) the reorganization of the collected dataset into a three-way data array by a so-called slicing approach and (ii) the subsequent application of Parallel Factor Analysis-Alternating Least Squares (PARAFAC-ALS) [9]. From a broad perspective, data slicing can be considered a tensorization approach [10] which consists of splitting the exponential decays of a data matrix, say **D**, into several equally sized slabs or "slices" covering different signal time ranges separated by a certain lag. The slices obtained are afterwards rearranged into a three-way data array **D**, to which a trilinear decomposition method is applied (see Fig. 1A). Trilinearity offers the advantage of uniqueness, being trilinear models more robust to noise and less affected by the choice of the initial estimates as long as the datasets analyzed have full rank [9,11]. These properties enhance significantly the capacity for unmixing multiexponential signals even in conditions of complete window overlap and high correlation among profiles. However, methodologies like PowerSlicing usually require that the measured signals encompass many sampling points (hundreds), so that the slicing procedure can be efficiently performed returning accurate results. In real scenarios where the exponential signals would consist of a few tenths of sampling points, results could be much less accurate or even incorrect.

In this work, we propose an alternative approach, called Kernelizing, to tensorize multiexponential signals characterized by only few sampling points. Kernelizing exploits the following invariance property of exponential functions, i.e., when convolving a monoexponential decaying function with any positive function of finite width (hereafter called "kernel"), the shape of the monoexponential decay (determined by the characteristic decay constant) remains unchanged and only the preexponential factor varies. This approach provides a new way to build three-way data arrays from the measured two-way data matrices of decay curves. To do this, each measured exponential decay (each row of the data matrix) is convolved with a set of different kernels, yielding new signals for which the decay constants of the individual monoexponential components are unchanged and only the corresponding preexponential factors are modified, but preserving the relative proportion of the components across the samples analyzed. Such an operation yields the "slices" needed to build a three-way trilinear data array from the original bilinear data to which a trilinear data decomposition, such as PARAFAC-ALS, can be applied to extract decay constants and concentration profiles of individual components (see Fig. 1B).

Both PowerSlicing and Kernelizing generate slices that are arranged

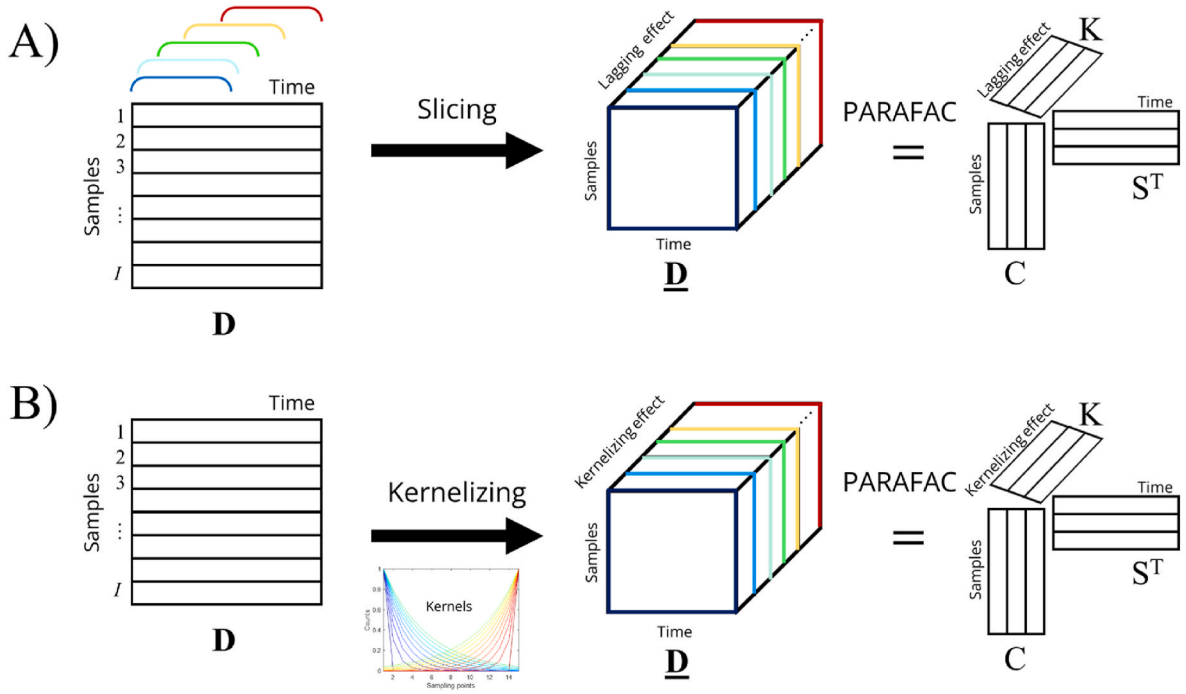


Fig. 1. Multilinear decomposition of exponential signals A) Slicing allows tensorizing a bilinear dataset into a trilinear dataset obtained by the selection of slices of the original decay separated by certain lags. B) Kernelizing allows tensorizing a bilinear dataset into a trilinear dataset convolving the original signals with a set of kernels.

Table 1
Composition of each analyzed mixture (expressed as volume ratios of standard solutions of pure dyes).

Mixture	ATTO 647 N (v/v)	ATTO 655 (v/v)	ATTO 665 (v/v)
1	1/3	1/3	1/3
2	4/6	1/6	1/6
3	1/6	4/6	1/6
4	1/6	1/6	4/6
5	5/12	5/12	2/12
6	2/12	5/12	5/12
7	5/12	2/12	5/12

in a three-way data array. However, the substantial difference between PowerSlicing (or other methodologies based on data lagging) and the Kernelizing approach is that the latter can provide a three-way array with an unlimited number of slices (as many slices as the number of kernels used), whereas PowerSlicing can only work with a limited number of slices, defined as a function of the total number of data points resulting from the sampling of the measured exponential signals. Such a difference becomes crucial when working with exponential signals with a low number of sampling points, a scenario in which Kernelizing improves significantly the accuracy and precision of the monoexponential profiles resolved.

To assess the performance of the proposed approach, several datasets simulated considering mixtures of monoexponential decays with different numbers of sampling points were studied. To complement these simulations, TRFS datasets were also investigated. TRFS is a well-established spectroscopic technique aiming at measuring the emission decay of a fluorophore in the picosecond to nanosecond timescale for the characterization of its lifetime (decay constant) [5]. Fluorescence lifetime is fluorophore-specific, dependent on the physicochemical environment probed, and provides valuable information about the sample. A real dataset was obtained gathering measurements performed on mixtures of known composition of ATTO fluorophores in solution using Time Correlated Single Photon Counting (TCSPC), a TRFS-based technique. Both PowerSlicing and Kernelizing were applied to the simulated and the TCSPC datasets (for which the ground truth solutions were known) to illustrate the benefits of the new approach proposed. Finally, the potential of Kernelizing was tested on a challenging real example of Fluorescence-Lifetime Imaging Microscopy (FLIM) [12], another TRFS-based technique widely used in the bioimaging field [13].

Although all the case-studies presented involve TRFS measurements, the results and conclusions inferred can be generalized to any analytical signal following a multi-exponential decaying behavior.

2. Datasets and software

The Kernelizing methodology was tested on simulated and experimental TRFS datasets. All datasets are representative of challenging scenarios for curve resolution approaches (high correlation among decay profiles, no selectivity, i.e., absence of subwindows with pure channels in either the concentration or the decay direction, and significant amount of Poisson-structured noise). The simulated dataset 1 consists of a set of multiexponential decay curve signals sampled at 500 time points (columns of the data matrix). The simulated dataset 2 consists of the same signals sampled at 50 points. The simulated dataset 3 consists of the same signals, with only 15 sampling points. The ATTO experimental dataset contains exponential decay curves of known mixtures of fluorescence dyes, originally sampled at 1500 time points. Finally, the ConvM dataset corresponds to a real FLIM image of a *Convallaria Majali*, for which pixel decay curves were sampled at 16 time points only. Further details about each dataset are given below. It should be noted that all datasets are full-rank.

2.1. Simulated dataset 1

Three pure monoexponential decays were simulated with decay constants of 0.8 ns, 1.4 ns and 2.4 ns, respectively, and with 500 equispaced sampling points. The three pure exponential profiles were organized in a matrix sized 3 × 500. These profiles show strong correlation with complete window overlap, i.e., all the pure monoexponential decays have a correlation coefficient ≥0.9 among them and cover the full time range concerned. A total of 200 mixtures with different composition were simulated to build a concentration matrix, sized 200 × 3. The two-way dataset 1 (200 × 500) of multiexponential decay curves was generated by matrix multiplication of the concentration and pure monoexponential decay matrices. All samples have a significant contribution of the three components, resulting in a non-trivial unmixing problem, since no pure sample for any component exists. Poisson noise was added representing approximately 10.0% of the global signal.

2.2. Simulated dataset 2

As for dataset 1, three pure monoexponential decays with constants of 0.8 ns, 1.4 ns and 2.4 ns were used. However, in this case, they were downsampled (the total time range covered was unchanged) and only 50 sampling points were considered , which resulted in a pure profile matrix sized 3 × 50. The concentration matrix, sized 200 × 3, was identical to the one exploited for dataset 1. The dataset 2, sized 200 × 50, was generated by matrix multiplication of the concentration and pure monoexponential decay matrices. Poisson noise was added representing approximately 5.0% of the total signal.

2.3. Simulated dataset 3

The same pure monoexponential decays as for dataset 1 and 2, with constants of 0.8 ns, 1.4 ns and 2.4 ns, were used; however, only 15 sampling points were now considered. The concentration matrix was left unchanged. The resulting dataset 3, sized 200 × 15, was obtained as described before and Poisson noise was added representing approximately the 4.5% of the total signal.

2.4. Time Correlated Single Photon Counting (TCSPC) data of ATTO fluorophores

Seven mixtures were prepared using solutions of three commercial dyes (ATTO 647, ATTO 655 and ATTO 665 from ATTO-TEC GmbH, a. r.) at the volume ratios listed in Table 1. Standard solutions of pure dyes were prepared in phosphate buffer solution (PBS) at pH 7.4 with a concentration of $5{\cdot}10^{-7}$ M.

All TCSPC measurements were performed using a PicoQuant TCSPC system with a FluoTime 200 spectrometer equipped with a picosecond laser diode emitting at 640 nm with a pulse width <90 ps full width at half-maximum (FWHM) and a repetition rate of 8 MHz.

A microchannel plate photomultiplier tube (MCP-PMT) connected to a TCSPC system (TimeHarp260, time precision 20 ps, dead time 25 ns) with a bin time of 25 ps was used for detection. The instrumental response function (IRF) of system (75 ps FWHM) was measured using a nonfluorescent scattering solution (LUDOX colloidal silica solution). Measurements stopped when the maximum reached 10 000 counts. The signals were recorded at 700 nm using a band pass filter with a 4 nm band pass, which resulted, after cropping the non-exponential signal portion originated from the IRF, in a TRFS data matrix composed of seven fluorescence decay curves in the range 0–20 ns with 1500 time points each. Thus, the experimental ATTO dataset is sized 7 × 1500 and does not contain any pure sample. In addition, a second dataset with fewer sampling points was built by sampling the ATTO dataset decay curves once every 100 points. The reduced dataset obtained is sized 7 × 15.

To assess the accuracy of the results obtained from the analysis of the

mixtures, the fluorescence decays of the pure solutions of ATTO 655, ATTO 665 and ATTO 647N were measured and fitted by a monoexponential decay to retrieve an estimate of the ground-truth lifetimes (1.9 ns, 2.9 ns and 3.5 ns, respectively).

2.5. ConvM dataset

The ConvM dataset was generated from a FLIM image of *Convallaria Majali* acquired by G. Williams et al. [12]. Instrumental details are available in the original reference. In FLIM, for a given spectral channel, a multiexponential fluorescence decay is provided for each pixel. The resulting dataset is, therefore, a hypercube of four dimensions, sized 256 × 256 pixels × 512 spectral channels × 16 time points. For every pixel, the spectral dimension of the FLIM image was integrated by summing the intensity values of all the spectral channels at each time point in order to analyze only the time dimension and increase the signal-to-noise ratio. This operation provided a FLIM image of 256 × 256 pixels and 16 time-sampling points, with a decay curve associated to every pixel. The first 30 columns of pixels were cropped because the related signal was saturated. In addition, time channels 1 to 3 and 14 to 16 were removed because they presented a non-exponential behavior. After cropping, the final size of the image was 256 × 227 pixels × 11 sampling points. The related unfolded convallaria dataset had size 58 112 × 11.

2.6. Software

All in-house routines, scripts and analyses generated for the Kernelizing approach were performed using MATLAB 2021 (The Mathworks, Inc., Natick, MA). PowerSlicing was adapted from Engelsen et al. [8]. The N-Way331 toolbox [14] was used for PARAFAC-ALS analysis.

For all analyses, the PARAFAC-ALS convergence criterium was set as 10^{-10}%, this small value being justified by the high correlation of the monoexponential profiles to be retrieved. In this scenario, when pure profiles are very correlated, a clear change in the decay constants of the resolved profiles may result in a very small change in the model residuals. Therefore, the algorithm needs more iterations to refine the resolved profiles. The analysis was repeated 1000 times adding the same amount of Poisson-structured noise in each run for each simulated dataset. The results obtained across these multiple runs help in the assessment of the accuracy and precision of the solutions obtained.

3. Data analysis

Let us consider the dataset **D** shown in Fig. 1, constituted by a set of multiexponential decay curves from different samples, sized $I \times J$ (samples and number of time-sampling points, respectively). Using curve resolution methods and under certain constraints, such as non-negativity, the bilinear decomposition of **D** provides component profiles (i.e., decay profiles and related sample concentration profiles) with direct chemical meaning, identifiable as those of the pure chemical compounds present in the samples [15,16]. However, due to the high correlation and time overlap among the monoexponential signals of the pure components, the bilinear decomposition of the data, even under constraints, seldom results in unique solutions, which hinders the interpretation of the results [11].

As shown in Fig. 1A and B, a two-way dataset of decay curves can be transformed into a three-way data array through different tensorization approaches. These data arrays or tensors (say, generically, $\underline{\mathbf{D}}$), follow a trilinear model as defined in Eq. (1), [9].

$$\underline{\mathbf{D}} = \mathbf{C}(\mathbf{S} \odot \mathbf{K})^{\mathrm{T}} + \underline{\mathbf{E}} \qquad \text{Eq. 1}$$

where $\underline{\mathbf{D}}$ is sized $I \times J \times K$ and the N-components model is built from the data matrices **C** ($I \times N$), **K** ($K \times N$), and **S** ($J \times N$), with $\odot$ denoting the Khatri-Rao product. The matrix **C** contains the pure concentration profiles, **S** the related pure monoexponential decays and the matrix **K** is related to the way the initial two-way dataset of full decay curves is transformed into $\underline{\mathbf{D}}$. $\underline{\mathbf{E}}$ is the data array of residual variation unexplained by the model, sized $I \times J \times K$.

One of the most useful properties of trilinear data factorization is the uniqueness of the solutions obtained [9,11]. This property is especially useful to resolve datasets formed by mixtures of exponential decay curves, with high correlation among them and showing no selectivity. To perform the trilinear decomposition of the data array $\underline{\mathbf{D}}$, Parallel Factor Analysis-Alternating Least Squares (PARAFAC-ALS) [9] will be used to obtain the trilinear model given by the matrices **C**, **K** and **S**, which contain chemically meaningful profiles for the components in the system studied.

A parameter used to assess the fit quality of the PARAFAC-ALS model is the lack of fit (LOF) expressed as in Eq. (2).

$$\mathrm{LOF}\ (\%) = 100 \times \sqrt{\frac{\Sigma_{i,j,k} e_{i,j,k}^2}{\Sigma_{i,j,k} d_{i,j,k}^2}} \qquad \text{Eq. 2}$$

where $d_{i,j,k}$ is the ijk $^{\mathrm{th}}$ element of $\underline{\mathbf{D}}$ and $e_{i,j,k}$ is the residual associated with the reproduction of $d_{i,j,k}$ by the trilinear model.

To assess the quality of the final trilinear model, the core consistency diagnostic (CORCONDIA) [17] can be used as an indicator parameter. CONCORDIA takes values from 100 (perfect fit by a trilinear model) to 0 or even negative (when the fitted trilinear model is less appropriate).

Despite the fact that PARAFAC-ALS has been the chosen algorithm in this work, it is important to note that the Kernelizing approach would increase the accuracy of the solutions provided by any other trilinear decomposition method.

3.1. Tensorization methods: slicing

Trilinear data arrays can be built from bilinear matrices by adequately reorganizing the original data, an example of so-called Data Tensorizing [10]. In this context, Pedersen et al. [4] proposed to use a slicing methodology to generate trilinear data from bilinear data based on the mathematical properties of exponential decays. Later, PowerSlicing [8], which consists of an optimal way to perform such slicing, was introduced. PowerSlicing and other slicing techniques exploit the invariance of exponential functions taken at different lags (see Eq. (3)):

$$Ae^{-\frac{t+\Delta t}{\tau}} = e^{-\frac{\Delta t}{\tau}} Ae^{-\frac{t}{\tau}} \qquad \text{Eq. 3}$$

When the independent variable t is lagged ($t + \Delta t$), the characteristic decay time τ of the exponential function (i.e., its shape, $e^{-\frac{t}{\tau}}$) remains unchanged, and only the value of the preexponential factor (A) gets modified.

If several slices, lagged Δt from one another, are extracted from the same decay (Fig. 1A), they will share the same τ but their respective preexponential factors will be different. Thus, linking Eq. (3) with Eq. (1), the preexponential factor A is related to the concentration mode (**C**), the $e^{-\frac{t}{\tau}}$ term is related to the shape of the monoexponential decay (**S**) and the new preexponential term $e^{-\frac{\Delta t}{\tau}}$ is related to the lag, which is in turn linearly related to the signal and becomes the new dimension **K**. The same occurs if the decay curve results from a mixture of exponential decays: individual τ are unchanged, but the preexponential factors for all the components get modified as previously described. Thanks to this mathematical property, slices can be arranged as a three-way data array that can be readily decomposed (Eq. (1)). For slicing methodologies, the profiles in **K** describe how the preexponential factors vary over the K slices, i.e., they define the lagging effect per component.

When applying PowerSlicing, we define the number of slices K to consider for the tensorization step according to the relation $\frac{J}{2} \geq 2^{K-1}$, being J the number of sampling points of the decays. The size of every slice is then $J - 2^{K-1} + 1$. For instance, for a matrix built from expo-

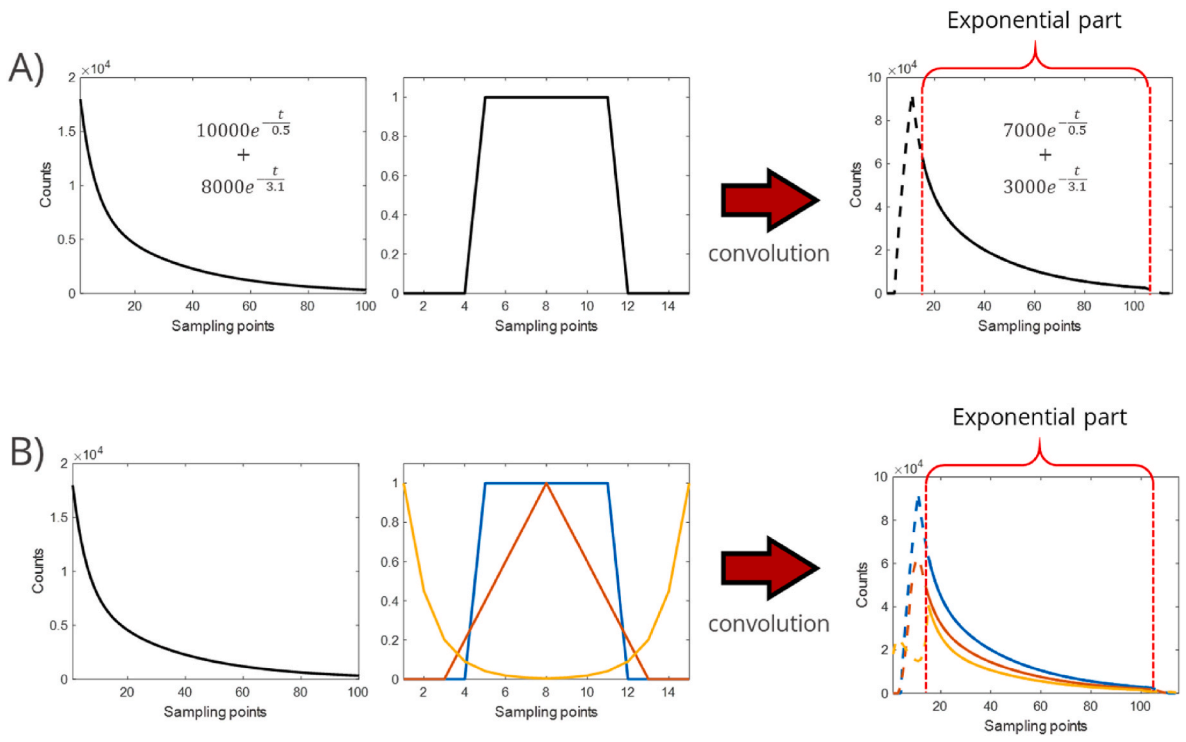


Fig. 2. Representation of a multiexponential decay resulting from the contributions of two monoexponential components (preexponential factors of 10 000 and 8000 and decay constant of 0.5 and 3.1, respectively) convolved with A) a single kernel (the preexponential factors change but the decay constants remain unchanged) and B) three kernels, yielding three new signal profiles, each of them characterized by different preexponential factors but the same decay constants. The dotted lines indicate the extremes of the convolved signals exhibiting non-exponential behaviors. Only the exponential part of the convolved signals is used for further trilinear decomposition analysis.

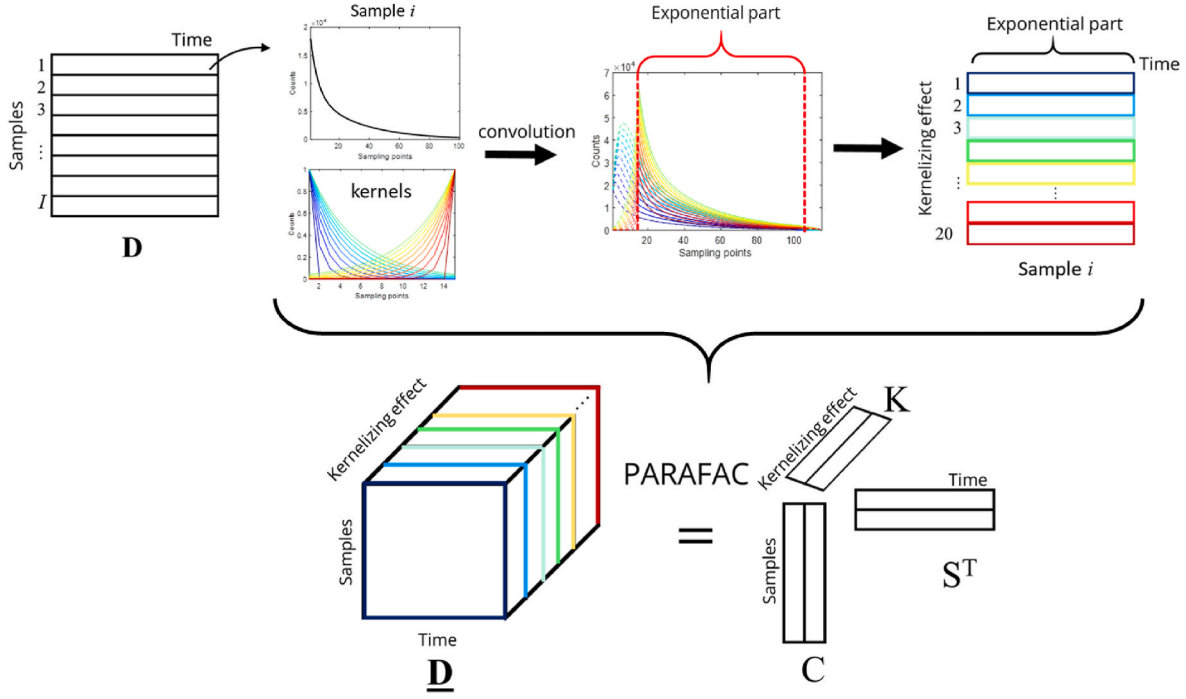


Fig. 3. Schematic representation of the Kernelizing approach. Each decay of the matrix **D** is convolved individually with a set of different kernels. Then, only the exponential parts of the resulting curves are kept and gathered into a matrix. After convolving every sample with the aforementioned kernels, a trilinear data array **D** can be built and subsequently analyzed by PARAFAC-ALS.

Table 2
Median of the correlation coefficients between simulated and recovered concentration profiles and of the decay constants recovered by PowerSlicing and Kernelizing in the 1000 analyzed runs. Values between brackets show the interval between the 2.5th and 97.5th percentiles of the corresponding metric. Median of CONCORDIA parameter across all runs is shown.

	True decay constant	Simulated dataset 1			
		PowerSlicing		Kernelizing	
		Concentration profile	Decay constant	Concentration profile	Decay constant
Component 1	0.80	0.98 [0.98, 0.99]	0.78 [0.76, 0.80]	0.99 [0.99, 0.99]	0.79 [0.78, 0.80]
Component 2	1.40	0.95 [0.91, 0.97]	1.39 [1.35, 1.43]	0.98 [0.97, 0.98]	1.40 [1.37, 1.43]
Component 3	2.4	0.96 [0.95, 0.97]	2.26 [2.13, 2.41]	0.98 [0.98, 0.99]	2.34 [2.26, 2.44]
CONCORDIA (%)		55		97	
	True decay constant	Simulated dataset 2			
		PowerSlicing		Kernelizing	
		Concentration profile	Decay constant	Concentration profile	Decay constant
Component 1	0.80	0.95 [0.94, 0.97]	0.76 [0.73, 0.78]	0.97 [0.96, 0.98]	0.77 [0.75, 0.79]
Component 2	1.40	0.86 [0.74, 0.92]	1.38 [1.30, 1.42]	0.94 [0.91, 0.96]	1.42 [1.37, 1.46]
Component 3	2.4	0.91 [0.87, 0.93]	2.16 [2.01, 2.35]	0.95 [0.93, 0.96]	2.27 [2.16, 2.41]
CONCORDIA (%)		24		77	
	True decay constant	Simulated dataset 3			
		PowerSlicing		Kernelizing	
		Concentration profile	Decay constant	Concentration profile	Decay constant
Component 1	0.80	0.92 [0.86, 0.94]	0.71 [0.60, 0.75]	0.94 [0.93, 0.95]	0.73 [0.69, 0.76]
Component 2	1.40	0.64 [0.56, 0.81]	1.29 [1.12, 1.40]	0.83 [0.72, 0.89]	1.41 [1.33, 1.47]
Component 3	2.4	0.83 [0.76, 0.88]	2.00 [1.86, 2.18]	0.89 [0.85, 0.91]	2.15 [2.04, 2.29]
CONCORDIA (%)		−2		24	

nential signals consisting of 1000 sampling points, nine slices can be generated, with 745 points each. If now only 15 sampling points are available, only three slices can be obtained, each of 12 sampling points. The effect of reducing the number of slices and the number of sampling points per slice on the stability of the trilinear model can be dramatic when few sampling points are available.

3.2. Tensorization methods: kernelizing

Kernelizing provides an alternative approach to generate trilinear data arrays from mixtures of monoexponential decay signals. It exploits the invariance of exponential functions resulting from the fact that by convolving a monoexponential decay function with any positive function (kernel), the preexponential factor gets modified but the characteristic decay time remains unchanged.

An example is shown in Fig. 2A, which illustrates this property. From left to right, an exponential signal of 100 sampling points generated from the sum of two monoexponential decays is convolved by a given kernel, with a width of 15 points. Once cut at the beginning and at the end (removing as many points as the size of the kernel window used), the resulting signal corresponds to a decay signal formed by the two same monoexponential components, but with different preexponential factors (for a mathematical proof and additional explanations see Eq. S1-7 in the Supplementary Material).

If a single decay curve is convolved with a set of K kernels (see Fig. 2B, $K = 3$), K new decays are obtained, characterized by different preexponential factors but the same decay constants. Then, if each decay curve (rows of **D)** is convolved with the same set of K kernels, the K convolved decays obtained can be arranged into a three-way data array (**D**) and analyzed by PARAFAC-ALS to obtain the underlying monoexponential contributions associated with the individual components.

A schematic representation of the Kernelizing approach is shown in Fig. 3. I samples, characterized by different mixtures of two monoexponential decays, and a set of 20 kernels of different shapes, providing 20 new convolved decays per sample, are here considered. Since the preexponential factors of the components change along the kernel direction, a third linearly independent dimension (mode) is obtained, and, thus, **D** results to be a trilinear array. Differently from slicing, this procedure remains applicable for signals characterized by very few sampling points.

It should be noted that kernel normalization is recommended to obtain convolved sets of decay curves with similar signal intensity for the subsequent PARAFAC-ALS analysis. Fig. 3 shows an example where kernels were normalized to a maximum value of 1.

Several features can be expected for the Kernelizing approach. First, the number of kernels used to build **D** does not depend on the number of sampling points of the decay curves and, hence, a rich trilinear dataset, with no limited number of slabs can always be generated. The second feature is the denoising action associated with the Kernelizing approach, since signal convolution can be interpreted as a weighted moving average methodology, where every point in the convolved signal results from the sum of the signal points covered by the kernel window, weighted by the related kernel coefficients. The third feature is linked to the wide diversity of kernel shapes that can be chosen to emphasize different parts of the original decay signal. For example, by choosing a kernel corresponding to a monotonically decreasing function, the emphasis would be on the components with longer decay times, since the decay points at longer times will be weighted with a larger coefficient in the convolved signal, (see the darkest blue kernel and the related

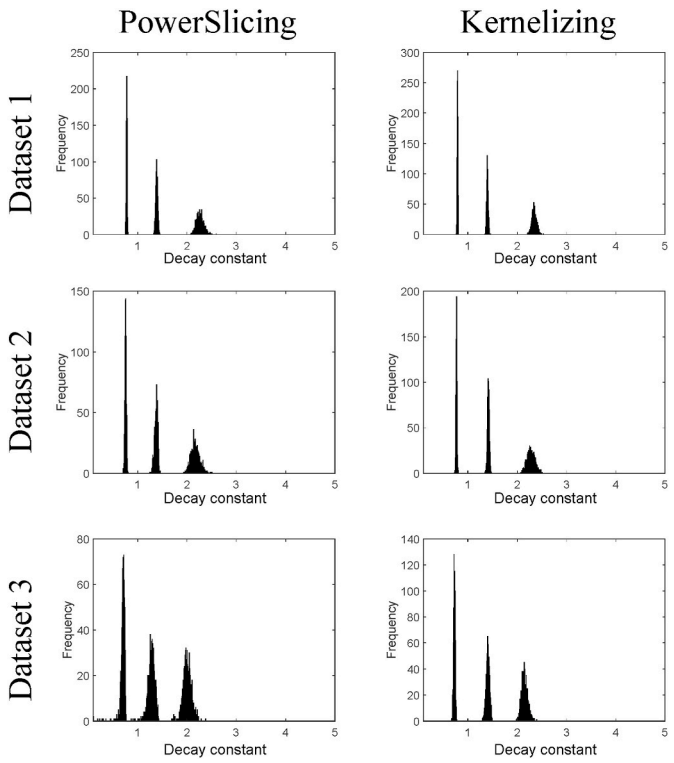


Fig. 4. Decay constants returned after the application of PowerSlicing and Kernelizing to the simulated datasets 1, 2 and 3 over the 1000 calculation runs. A significant reduction of the estimate scatter is observed for Kernelizing with respect to PowerSlicing when the number of sampling points is decreased.

convolved blue signal in Fig. 3). In our case, we have selected a set of kernels based on symmetric exponential functions to weight equally the short and long decay time components, as it is shown in Fig. 3. In general, it would be advisable to use a set of kernels that allows weighting in a similar way the entire range of points of the decay curves to be convolved.

4. Results and discussion

Kernelizing has been tested on simulated and time-resolved fluorescence experimental data acquired under controlled conditions. It has also been applied to a FLIM image for exploratory analysis. The results are discussed in the following sections.

4.1. Simulated examples

This subsection is divided in two parts. The first part aims at showing and discussing the performance of Kernelizing on simulated datasets with different numbers of sampling points (datasets 1, 2 and 3). The results are compared with those yielded by PowerSlicing. The second part aims at assessing the effect of key parameters of the Kernelizing approach on the results obtained.

4.2. Performance of the kernelizing approach

Datasets 1, 2 and 3 were simulated as explained in Section 2 and tensorized by both the Kernelizing and PowerSlicing approaches. The data arrays provided by Kernelizing were built by tensorizing the datasets 1, 2 and 3 using a set of 20 kernels, sized 20 $\times$ 250, 20$\times$ 25 and 20 $\times$ 8, respectively (in all cases, the kernel size is half the number of sampling points of the corresponding dataset). The typology and diversity of kernel functions were chosen according to the guidelines provided in Section 3 (see Fig. 3 and S3). The dimensions of the corresponding three-way data arrays are 200 $\times$ 20 $\times$ 250, 200 $\times$ 20 $\times$ 25 and 200 $\times$ 20$\times$ 7 for datasets 1, 2 and 3, respectively. In the case of PowerSlicing, considering the rules to define the number of slices and the sampling points therein, the dimensions of the three-way arrays generated are 200 $\times$ 8 $\times$ 373, 200 $\times$ 5 $\times$ 43 and 200 $\times$ 3 $\times$ 12, respectively.

PARAFAC-ALS was used to analyze the three-way data arrays resulting from both tensorization procedures. Initial estimates were set as the profiles yielded by the best-fitting models obtained after several PARAFAC-ALS runs started with a variety of initial estimates and fitted using only a few iterations [14].

To compare the results obtained by both the Kernelizing and PowerSlicing approaches, the individual pure component decay profiles in **S** were fitted by a monoexponential model to extract their respective characteristic decay constants and the correlation coefficients between the recovered concentration profiles in **C** and their corresponding ground truth profiles were calculated (see Table 2). The analysis was repeated 1000 times adding the same amount of Poisson-structured noise in each run for each dataset. To assess the accuracy and spread of the final solutions, the median and the 2.5th-97.5th percentile interval of these metrics were considered for every component. For each run, the CONCORDIA was calculated. The lack of fit of all models was found in agreement with the quantity of noise added to the simulated dataset (data not shown).

For the tensorized dataset 1 (500 sampling points), the values of the three decay constants are well recovered by both Powerslicing and Kernelizing. The full distribution of the decay constants is shown in Fig. 4. The exponential profiles and related decay constants are very well recovered for components 1 and 2, whereas a higher scatter, slightly more pronounced for Powerslicing, can be observed for component 3.

The correlation coefficients obtained for the concentration profiles are satisfactory for both approaches (equal or higher than 0.9 for all components), despite a slightly lower value for the concentration profile of component 2 returned by Powerslicing. Finally, it should be noted that a clear difference exists in the CONCORDIA values, 55% for PowerSlicing *vs* 97% for Kernelizing, meaning that the PARAFAC-ALS model yielded by Kernelizing is closer to an ideal trilinear model than the one yielded by PowerSlicing.

Overall, although Kernelizing provides slightly better results, both approaches guarantee a satisfactory performance when dealing with exponential decay signals for which a sufficiently number of sampling points is available.

When inspecting the results for datasets 2 and 3 (featuring 50 and 15 sampling points, respectively) the differences between PowerSlicing and Kernelizing become more pronounced (see Table 2). For the concentration profiles, Kernelizing provides correlation coefficients very close or higher than 0.9 in all components and datasets whereas the accuracy of Powerslicing decreases, as can be seen from the outcomes obtained for component 2 in dataset 2 and components 2 and 3 in dataset 3. Looking at the decay constants, correct decay constant values are recovered by Kernelizing for datasets 2 and 3 (with only a small bias for component 3 in the latter case). An additional relevant fact is that all Kernelizing models are very stable, the spread of the calculated decay constants does not change across datasets (see Fig. 4 and Table 2). Conversely, biased results are obtained for the decay constants of component 3 in dataset 2 and components 2 and 3 in dataset 3 when applying PowerSlicing. Additionally, the spread of the solutions is significantly larger when the number of sampling points decreases, which derives from the instability PARAFAC-ALS exhibits when taking into account reduced numbers of slices. Another indicator of the quality of the modelling approach is the CONCORDIA value. It can be observed that this parameter decreases for both approaches from dataset 1 to dataset 3, but the effect is more pronounced for PowerSlicing. All the differences mentioned above stem from the fact that reducing the number of sampling points heavily affects the amount of information that can be encoded in the three-way data array generated by Powerslicing since less slices can be built (five and

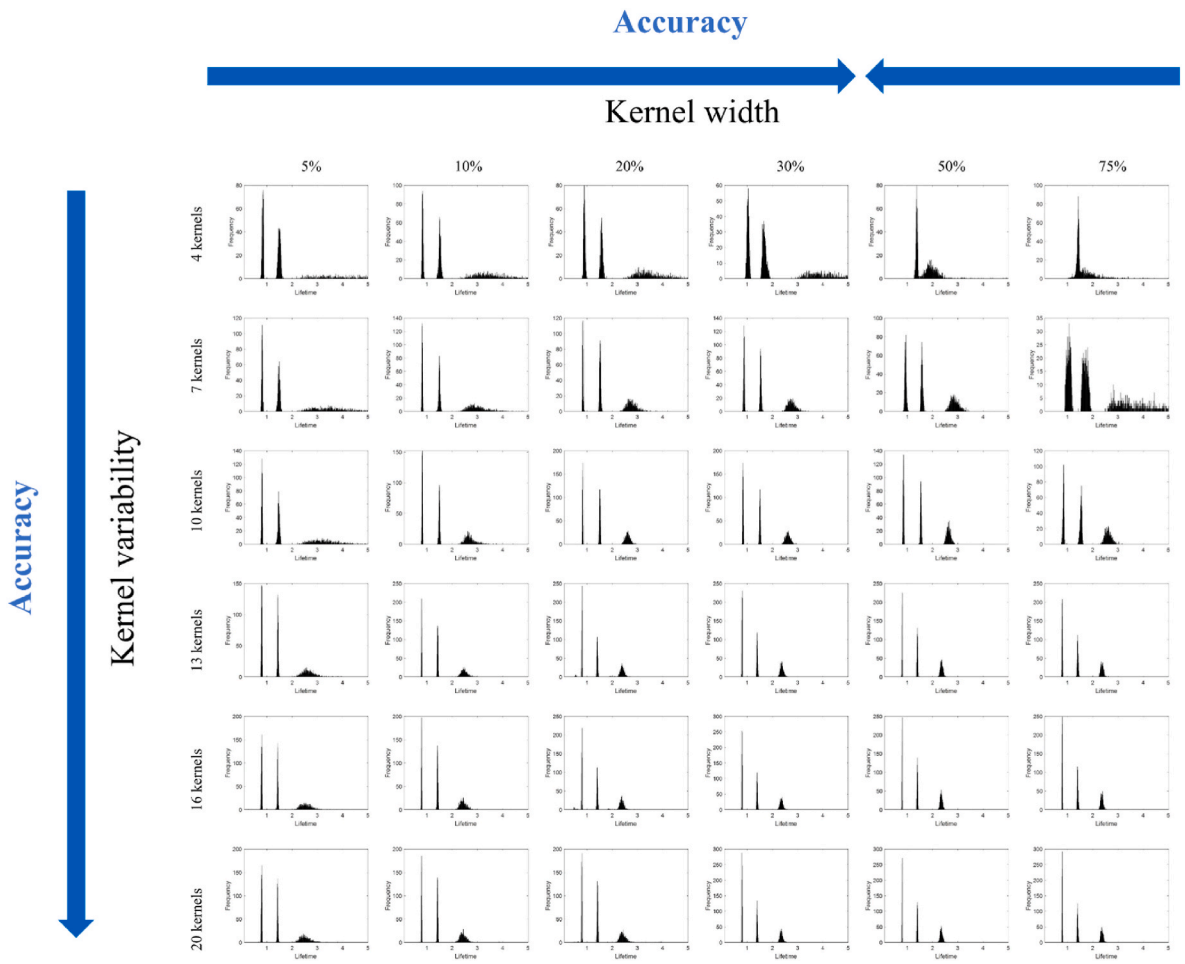


Fig. 5. Dispersion of the values of the decay constants recovered using Kernelizing on dataset 1. From top to bottom, the number and diversity of the kernel functions increase. From left to right, the width of the kernel functions (expressed as percentage of points over the total number of sampling points of the original decay curves) increases.

Table 3
Summary of results yielded by PowerSlicing and Kernelizing for the ATTO dataset in the two tested scenarios.

	True lifetime (ns)	ATTO data (1500 sampling points)				ATTO data (15 sampling points)			
		PowerSlicing		Kernelizing		PowerSlicing		Kernelizing	
		Concentration profile*	Lifetimes (ns) [+]	Concentration profile*	Lifetimes (ns)[+]	Concentration profile*	Lifetimes (ns) [+]	Concentration profile*	Lifetimes (ns) [+]
ATTO 655	1.9	0.75	1.786 [1.783–1.789]	0.76	1.887 [1.886–1.887]	0.70	1.86 [1.85–1.88]	0.72	1.93 [1.92–1.94]
ATTO 665	2.9	0.86	2.435 [2.432–2.438]	0.87	2.486 [2.486–2.486]	0.83	2.64 [2.62–2.66]	0.88	2.8 [2.79–2.83]
ATTO 647 N	3.5	0.90	3.419 [3.416–3.421]	0.99	3.494 [3.494–3.495]	0.77	3.65 [3.61–3.68]	0.86	3.74 [3.72–3.77]

[+] Monoexponentially fitted lifetime and 95% confidence interval associated with the fitting error.

three for datasets 2 and 3, respectively). This, together with the decreasing number of sampling points per slice, hinders the correct recovery of monoexponential decays. For the Kernelizing approach, the number of sampling points per slice is also reduced, but the number of slabs remains unchanged (20 for all datasets). The shape diversity of the kernels is also preserved and, as a consequence of both facts, a three-way data array with large variability of information can still be generated in the worst-case scenario and satisfactory results can be obtained.

4.3. Effect of the kernel size and shape variation on the final solutions

In this section, we investigate several key features of the Kernelizing approach, such as the number of kernels used, their widths and shapes.

Six different sets of kernels were generated, enclosing from four to 20 exponential functions with different shapes, chosen so that the set of 20 kernels weights similarly short and long decay time components. The kernel sets used are shown in Fig. S3. In addition, for each set, six different kernel widths were considered (covering 5, 10, 20, 30, 50 and 75% of the total number of sampling points of the data). This results in a total of 36 different sets of kernels that were used to build the three-way data arrays corresponding to datasets 1, 2 and 3. For each combination of kernel shape and kernel width, PARAFAC-ALS analysis was repeated 1000 times adding to the data the same amount of Poisson-structured noise in each run.

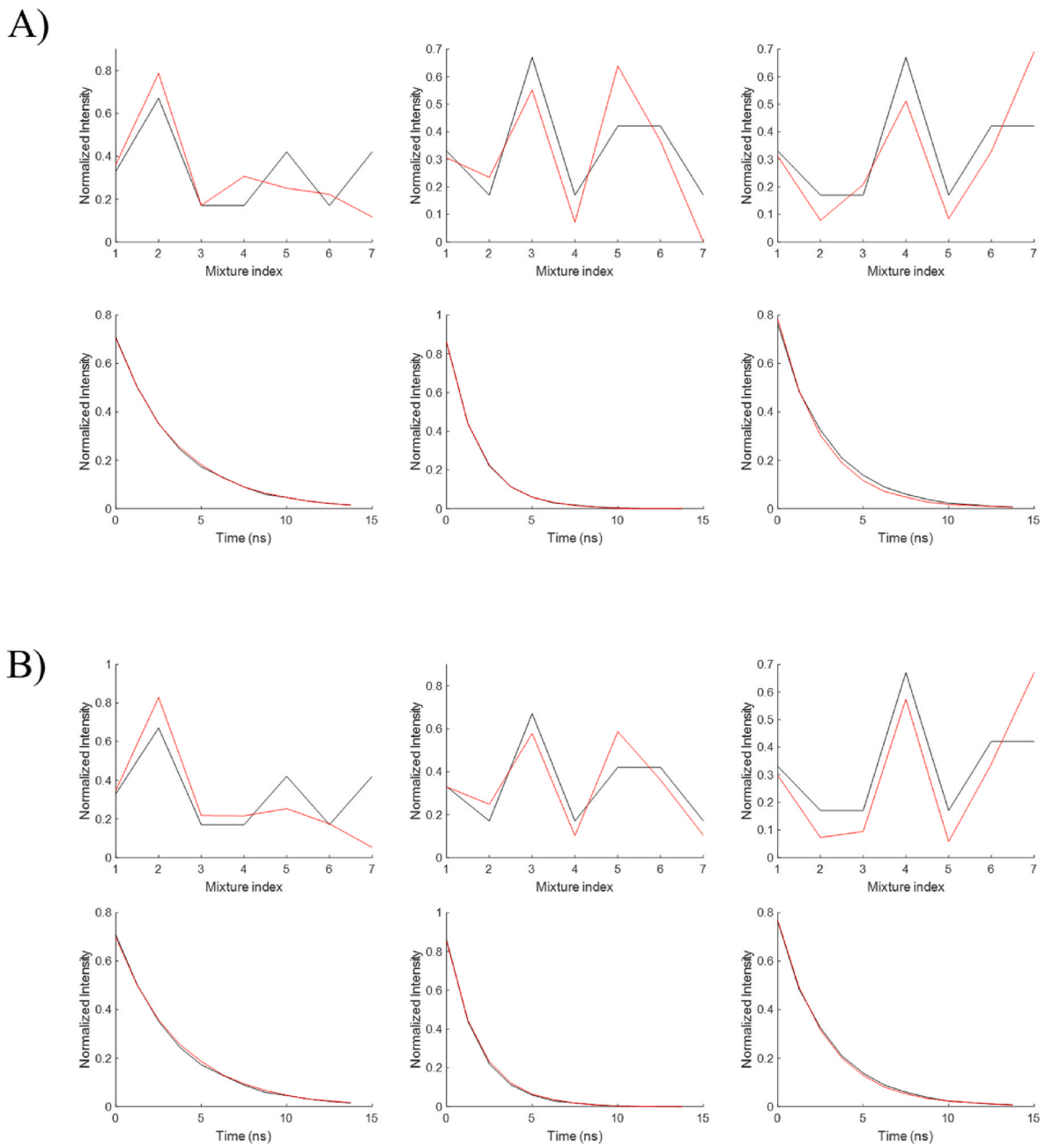


Fig. 6. ATTO dataset (15 sampling points). A) Top panel, concentration profiles recovered by PARAFAC-ALS for PowerSlicing (red) and expected concentration profiles (black). Bottom panel, pure fluorescence decays recovered by PARAFAC-ALS for PowerSlicing (red) and expected pure fluorescence decays (black). B) Top panel, concentration profiles recovered by PARAFAC-ALS for Kernelizing (red) and expected concentration profiles (black). Bottom panel, pure fluorescence decays recovered by PARAFAC-ALS for Kernelizing (red) and expected pure fluorescence decays (black). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 5 shows the decay constants obtained from the analysis of dataset 1. Looking at the results from top to bottom, it can be clearly observed that the higher the number of kernels and the wider the diversity of their shape, the closer to the ground truth the solutions are. For this reason, a key point is to always use a high number of kernels with very different shapes, since more variance is induced in the preexponential factors of all components, irrespective of their characteristic decay time (short or long).

Looking at the results in Fig. 5 from left to right, the effect of the kernel width can be assessed. In general, the larger the kernel, the more signal needs to be removed (see Fig. 3) and, thus, the lower the amount of available information exploitable for the resolution of strongly overlapping monoexponential components. On the other hand, when the kernel is too narrow, the variability induced in the preexponential factors is too small, and so is the denoising action, both effects resulting in a larger dispersion of the results. In practice, a compromise should be found and for the cases explored here, choosing a kernel width in the range between 20% and 50% of the total number of sampling points provided good results. However, different approaches may be advisable when coping with real-world datasets for which the ground truth is unknown, e.g., generating replicates by the noise addition method [18] and looking at the spread of the final results.

As a final conclusion, it is also important to understand the interaction between the effects of the key factors described above. When the number of kernels increases, the kernel width can be significantly reduced and, hence, the part of the convolved signals with non-

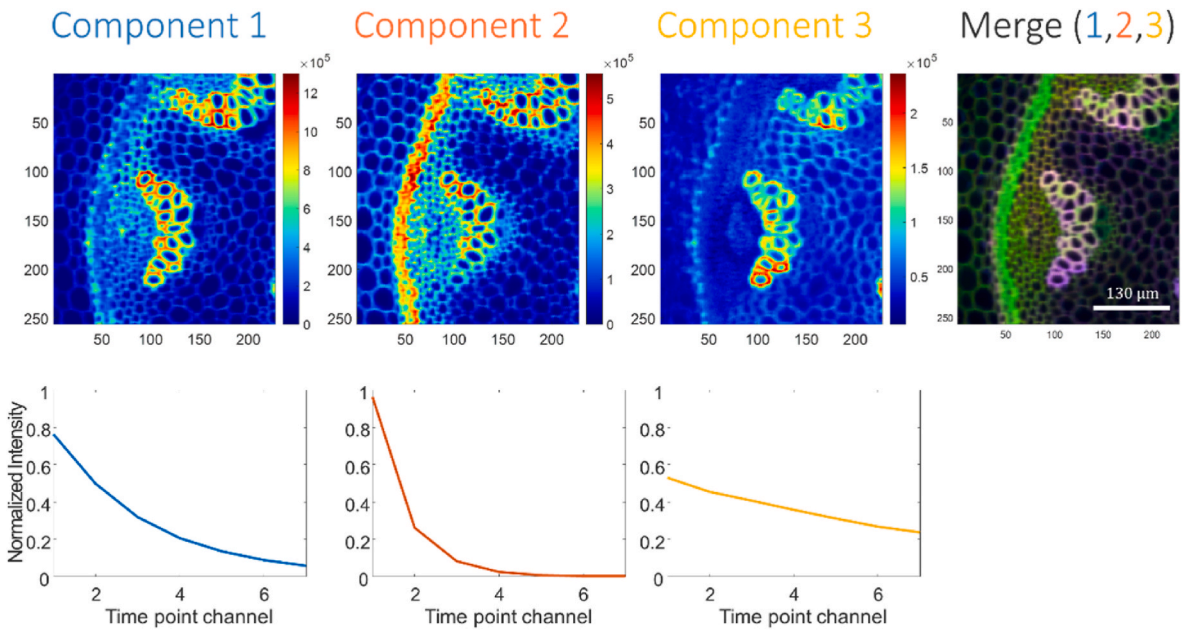


Fig. 7. FLIM dataset. Pure distribution maps (top) and pure fluorescence decays (bottom) recovered by Kernelizing-PARAFAC-ALS).

exponential behavior minimized.

The results obtained for datasets 2 and 3 provide the same general conclusions as for dataset 1 and are presented in the Supplementary Material (Figs. S4 and 5).

4.4. ATTO fluorescence data

The ATTO dataset, composed of fluorescence decays of seven ternary mixtures of three ATTO dyes sampled at 1500 points, was used to compare the results obtained by both Kernelizing and Powerslicing on real data. Kernelizing was applied using a set of 20 kernels of 300 point width (20%), resulting in a data array sized $7 \times 20 \times 1200$. PowerSlicing was applied using 10 slices and the corresponding data size was $7 \times 10 \times 989$. The tensorized data arrays were then decomposed by means of PARAFAC-ALS. After this, the resolved time profiles were fitted and the respective lifetimes extracted with their corresponding fitting error (95% confidence). The CONCORDIA parameter was also calculated.

Table 3 shows a summary of the results obtained. For full length signals, the concentration profiles and decay constants are, generally, well recovered by both approaches when they are compared with the ground truth (Fig. S6), despite the high correlation among the undelying monoexponential profiles and the very low number of samples handled. The decay constants of the components found for PowerSlicing and Kernelizing match well the expected ones. However, a small deviation for the component ATTO 665 can be observed for both approaches. The concentration profiles are also in good agreement with the true ones. Thus, in this scenario, both approaches are found to perform equivalently from a practical point of view, which is in line with the results obtained on simulated data for a large number of sampling points (dataset 1).

In a second step, the sampling points in the ATTO dataset were reduced to 15 (see Section 2). Kernelizing (20 kernels, 3 points width) was then applied to the resulting signals and a $7 \times 20 \times 12$ data array was obtained. PowerSlicing was applied to the same data using three slices which yielded a $7 \times 3 \times 12$ data array. PARAFAC-ALS was used to analyze both datasets. Results are shown in Fig. 6. It can be observed that the concentration profiles and the decay constants retrieved by PARAFAC-ALS are more consistent with the ground truth when employing the Kernelizing approach rather than PowerSlicing, especially for components 2 and 3. This can be explained by the fact that only three slices are here available for Powerslicing as opposed to the 20 kernelized versions of the initial dataset, as previously observed for dataset 3. It is also important to notice that the error associated with the fitted lifetimes is higher for Powerslicing, indicating that the pure decay curves extracted may be a bit further from the pure monoexponential shapes expected.

4.5. Analysis of FLIM data

A FLIM image of *Convallaria majali* has also been investigated for illustrative purposes. The FLIM data was made available by Williams et al. [12]. The FLIM image has size $256 \times 227 \times 11$ after preprocessing (see Section 2), where the first two dimensions are the *x*- and *y*-spatial directions and the third represents the 11 sampling points of the decay curve of every pixel, respectively. A set of 20 kernels of 3 time points each was used to tensorize the data, resulting in a dataset sized $256 \times 227 \times 20 \times 8$. The dataset was decomposed (after unfolding along the pixel direction) by a three-component PARAFAC-ALS model. The concentration profiles obtained were refolded to recover the 2D concentration maps of every component. Fig. 7 shows the pure component concentration maps and the related pure monoexponential decays obtained.

As can be observed, despite no ground truth is available, the three resolved concentration maps highlight quite specific biological zones of the vegetal tissue. A tentative assignment would be the following: component 1 relates to the xylem and could correspond to the safranin dye linked to lignin. Besides, the endodermis is highlighted. The safranin dye stains lignin, which is generally located on the xylem and the endodermis [19]. Component 2 (fast decay) generally appears in the mesophyll cells as well as in some lignified cells and might be related to the fast-green stain. It has been reported that fast-green stains well the phloem and cellulosic cell walls in the pith [19]. On the other hand, component 3 (slow decay) appears specifically on the xylem, differentiating multiple environments for lignin. These results are in agreement with Kaminski et al. [19], who characterized a similar sample by means of excitation-emission spectroscopy.

As all the investigated components have biological sense and were satisfactorily recovered in spite of the high complexity of the case-study dealt with, this last example shows the potential of the Kernelizing

approach for the analysis of FLIM images of vegetal tissues featuring a low number of sampling points.

5. Conclusions

In this work, Kernelizing is proposed as a very efficient way to obtain trilinear data arrays with a high degree of variability from bilinear data matrices of multiexponential decays. The richness of information encoded in these kernelized three-way arrays allows PARAFAC-ALS resolving chemical mixtures into individual components characterized by monoexponential decays with an equal or higher accuracy than well-established slicing approaches (such as PowerSlicing).

Thus, although PowerSlicing is a fast and robust method that can serve the same purpose as Kernelizing in most practical situations, we have identified specific scenarios for which the robustness of the PowerSlicing solutions can be questioned, i.e., situations for which very few slices can be obtained because the number of sampling points in the original multiexponential curves is low. Such a problem does not affect the proposed Kernelizing approach since the number of convolved decays that can be generated for each sample signal is not limited by the number of sampling points. As has been proven, the possibility to choose the number and shapes of the kernels used is an excellent asset to increase the variability in the three-way arrays to be analyzed and, consequently, the accuracy and precision of the solutions obtained.

Kernelizing has been found very useful to handle multiexponential measurements for which binning is required to increase the signal-to-noise ratio or whose number of sampling points is low due to instrumental limitations. Fluorophores characterized by very similar decaying behaviors, for example, could be unmixed in a FLIM imaging case-study, where the number of sampling points was limited due to specific features of the instrument resorted to.

At this point, the main aspect that needs further exploration is the choice of the kernel width. In this study, a trial-and-error approach was utilized, but we acknowledge that a more systematic strategy would be useful, e.g., defining the width of the kernels based on the characteristics of the dataset (number of sampling points, noise, etc.) would further simplify the generalized use of the Kernelizing approach.

Finally, it is worth pointing out that the results and conclusions drawn in this article mainly relate to the PARAFAC analysis of TRFS datasets, but can be generalized to any measurement that can be expressed by multiexponential decay curves and to any algorithm devoted to perform trilinear decomposition analysis.

CRediT authorship contribution statement

Adrián Gómez-Sánchez: Conceptualization, Methodology, Software, Formal analysis, Investigation, Discussion, Data curation, Writing – original draft, Visualization. **Raffaele Vitale:** Conceptualization, Discussion, Writing – review & editing. **Olivier Devos:** Data acquisition, Discussion, Writing – review & editing. **Anna de Juan:** Resources, Writing – original draft, Discussion, Supervision, Funding acquisition. **Cyril Ruckebusch:** Conceptualization, Methodology, Formal analysis, Resources, Writing – original draft, Discussion, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare no competing interests.

Data availability

Data will be made available on request.

Acknowledgements

A. G.-S. and A.J. acknowledge financial support from the Spanish government Project PID 2019-1071586 B-IOO and the Catalan government (2021 SGR 00449). A. G.-S. acknowledges scholarships from the MOBLILEX U Lille program and from the Santander bank.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2023.341545.

References

[1] H. Lemmetyinen, N.V. Tkachenko, B. Valeur, J.I. Hotta, M. Ameloot, N.P. Ernsting, T. Gustavsson, N. Boens, Time-resolved fluorescence methods (IUPAC technical report), Pure Appl. Chem. 86 (12) (2014) 1969–1998.
[2] A.A. Istratov, O.F. Vyvenko, Exponential analysis in physical phenomena, Rev. Sci. Instrum. 70 (2) (1999) 1233–1257.
[3] W. Windig, B. Antalek, Direct exponential curve resolution algorithm (DECRA): a novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles, Chemometr. Intell. Lab. Syst. 37 (2) (1997) 241–254.
[4] H.T. Pedersen, R. Bro, S.B. Engelsen, Towards rapid and unique curve resolution of low-field NMR relaxation data: trilinear SLICING versus two-dimensional curve fitting, J. Magn. Reson. 157 (1) (2002) 141–155.
[5] T. Ohno, Z. Wang, R. Bro, PowerSlicing to determine fluorescence lifetimes of water-soluble organic matter derived from soils, plant biomass, and animal manures, Anal. Bioanal. Chem. 390 (8) (2008) 2189–2194.
[6] O. Devos, M. Ghaffari, R. Vitale, A. de Juan, M. Sliwa, Ruckebusch, Multivariate curve resolution slicing of multiexponential time-resolved spectroscopy fluorescence data, Anal. Chem. 93 (37) (2021) 12504–12513.
[7] D. Cevoli, S. Hugelier, R. Van den Eynde, O. Devos, P. Dedecker, C. Ruckebusch, Multilinear Slicing for curve resolution of fluorescence imaging with sequential illumination, Talanta 241 (2022), 123231.
[8] S.B. Engelsen, R. Bro, PowerSlicing, J. Magn. Reson. 163 (1) (2003) 192–197.
[9] R. Bro, PARAFAC. Tutorial and applications, Chemometr. Intell. Lab. Syst. 38 (2) (1997) 149–171.
[10] O. Debals, L.D. Lathauwer, August. Stochastic and deterministic tensorization for blind signal separation, in: International Conference on Latent Variable Analysis and Signal Separation, Springer, Cham, 2015, pp. 3–13.
[11] R. Tauler, A. Smilde, B. Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, J. Chemometr. 9 (1) (1995) 31–58.
[12] G.O. Williams, E. Williams, N. Finlayson, A.T. Erdogan, Q. Wang, S. Fernandes, A. R. Akram, K. Dhaliwal, R.K. Henderson, J.M. Girkin, M. Bradley, Full spectrum fluorescence lifetime imaging with 0.5 nm spectral and 50 ps temporal resolution, Nat. Commun. 12 (1) (2021) 1–9.
[13] W. Becker, Fluorescence lifetime imaging–techniques and applications, J. Microsc. 247 (2) (2012) 119–136.
[14] The N-Way Toolbox for MATLAB Version 3.31, 2022, 18/11, http://www.models.life.ku.dk/nwaytoolbox/download. Last access: 9/6/2023.
[15] A. de Juan, R. Tauler, Multivariate curve resolution-alternating least squares for spectroscopic data, Data Handling Sci. Technol. 30 (2016) 5–51 (Elsevier).
[16] S. Benabou, C. Ruckebusch, M. Sliwa, A. Avino, R. Eritja, R. Gargallo, A. de Juan, Study of conformational transitions of i-motif DNA using time-resolved fluorescence and multivariate analysis methods, Nucleic Acids Res. 47 (13) (2019) 6590–6605.
[17] R. Bro, H.A. Kiers, A new efficient method for determining the number of components in PARAFAC models, J. Chemometr.: A Journal of the Chemometrics Society 17 (5) (2003) 274–286.
[18] J. Jaumot, R. Gargallo, R. Tauler, Noise propagation and error estimations in multivariate curve resolution alternating least squares using resampling methods, J. Chemometr.: A Journal of the Chemometrics Society 18 (7-8) (2004) 327–340.
[19] C.F. Kaminski, R.S. Watt, A.D. Elder, J.H. Frank, J. Hult, Supercontinuum radiation for applications in chemical sensing and microscopy, Appl. Phys. B 92 (3) (2008) 367–378.