

Kernel methods in fluorescence lifetime imaging microscopy

MATHEMATICS FOR INDUSTRY

⁵ *Kika Dimitric, Jonathan Grube, Matthijs Holstege, Roxy van de Kuilen, Ben Mason, Marcel Masque Salgado, Jesper van Soest, Splinter Suidman, Teun van Wezel*

MSc Mathematical Sciences

Supervisor:
Dr. Palina Salanevich

Contents

	1 Management Summary	1
	2 Introduction	1
	3 Problem description	2
15	4 Background	3
	5 Data	3
	5.1 Experimental data	3
	5.1.1 Data structure	3
	5.1.2 Data pre-processing	4
20	5.1.3 Truncation	4
	6 Method	4
	6.1 Experimental pipeline	5
	6.2 Fitting a single-dye model to detect single-dye pixels	6
	6.2.1 Single-dye models for pixels with a single dye present	6
25	6.2.2 Identifying pixels in the image with only a single dye present	8
	6.3 Gaussian mixture models to predict lifetime from single-dye analysis	9
	6.4 Double-dye analysis	9
	6.4.1 Fixed lifetimes	9
	6.4.2 Allowing lifetimes to vary per pixel	11
30	6.4.2.1 Laplace Method	11
	6.4.2.2 Performance of methods for varying lifetimes	13
	6.4.2.3 Overview of performance of methods for varying lifetimes	15
	7 Analysis on real data	16
	8 Conclusion	22
35	8.1 Recommendations and further work	23
	9 Acknowledgements	23
	References	24
	A Analysis of methodology	25
	A.1 Truncation and Laplace method	25
40	A.2 Generation of synthetic data	25
	A.3 Accuracy of identifying single-dye pixels	26
	A.4 Choosing a truncation parameter	26
	A.5 Parameter degeneracy	28
	B Analytic solution	28

1 Management Summary

In fluorescence lifetime imaging microscopy, biologists inject a biological sample with fluorescent dyes, and create images by analyzing the fluorescence process of these dyes. Specifically, the fluorescence lifetime is used to distinguish between dyes and determine their relative concentrations in the sample. Images created through this process help biologists understand the biological and chemical properties of this sample.

In this report, we present different methods of extracting the lifetime and concentration of each dye, based on different assumptions about the data. We consider the case where there are two dyes present in the sample. We distinguish between the cases where the lifetimes are assumed to be constant over the whole image and where the lifetimes may vary slightly throughout the image. In both cases, we use the computed lifetimes to determine the relative concentration of the two dyes at each pixel. These relative concentrations can be used to create images of the sample. We present a method where we first estimate the two dye lifetimes across the whole image by assuming that there is only one dye present. We use these estimates to obtain bounds on how much we expect the lifetimes to vary within the image, and estimate the concentration and lifetime of each dye in each pixel under these restrictions. In the case where we assume the lifetimes to be constant, we use the estimates for the lifetimes to determine the concentration of each dye in each pixel. In both the constant-lifetime and varying-lifetime case, we use methods based on the Fourier and Laplace transforms to extract the estimates for the lifetimes and concentrations in each step, since these are computationally efficient. We analyse and compare each of these methods on both experimental and synthetic data. We evaluate the methods on synthetic data, by investigating their sensitivity to noise, the number of samples used and the underlying lifetimes and concentrations of the dyes. For several parameter configurations we analyse how to choose parameters for the aforementioned Fourier and Laplace methods and under what conditions these methods fail to yield good results.

If the underlying lifetimes are constant over the image, the constant-lifetime method is shown to be more accurate. The analysis shows that if the lifetimes vary over the image, the varying-lifetime method, although computationally more expensive, yields more accurate predictions.

2 Introduction

Fluorescence microscopy is a non-invasive method to study biological cells. Using an optical microscope, also known as a light microscope, biologists can study the molecular structure of these cells by observing what happens after a laser is shone on a sample. The data resulting from this process is analysed and enhanced by companies like Scientific Volume Imaging (SVI), a company that offers software to assist in working with images created through this process. SVI has developed varied software applicable in this branch of biology, such as tools to improve image resolution, correct data noise or create sophisticated animations.

In fluorescence microscopy, biologists use fluorescent chemical compounds called *dyes* that attach to certain molecular structures in the tissue they are interested in. An example of such dyes is the commonly used dye fluorescein isothiocyanate. It attaches to amino, sulfhydryl, imidazolyl, tyrosyl or carbonyl groups on proteins. With fluorescence microscopy, we can visualize how these dyes attach to different parts of the tissue in a non-invasive way. After shining a laser of a specific wavelength at this tissue, the dye fluorescein isothiocyanate absorbs some of the energy from that laser: some of the electrons present in the molecule become ‘excited’. After a short period of time, the electrons ‘de-excite’, falling back into a (more stable) lower energy state, emitting a photon in the process. This process is called fluorescence. The fluorescent light that is emitted by the dye in the tissue is then filtered from the brighter excitation light by specific filters or mirrors in the light microscope.

This process is completed within a few nanoseconds. We call the average time that a dye is in an excited stage the *fluorescent lifetime* [1].

Biologists often use multiple dyes to get a better understanding of a tissue, relying on the different properties of the fluorescent chemical compounds. For example, they can use one dye that attaches to proteins present in cell membranes to show the cell outline and another dye that attaches to nucleic acids to label DNA present. It may happen that two or more dyes attach to the same or nearby matter in the tissue with an unknown concentration. To thoroughly understand the structure of the organic tissue, biologists need to know which dye is present where in the tissue. In the case where the two dyes emit different wavelengths, the images obtained with the microscope can already give clarity about the structure of the tissue via the colors. But sometimes the two dyes attaching to the different molecular structures emit the same wavelength. In that case, the dyes can only be distinguished through their lifetimes. This is why we want to be able to determine the fluorescence lifetimes of the dyes used in our data, which helps find the concentration of each dye for each pixel in the image.

In § 3, we state the problem we are focusing on, followed by examples of possible applications of fluorescence microscopy in § 4. Then, in § 5, we explore the experimental data we used in this project. In § 6, we discuss the different methods used in this project to find dye lifetimes. This includes the least squares methods and methods based on the Fourier and Laplace transform. We analyse these methods on synthetic data in the case where we assume the lifetimes to be constant over the whole image, as well as the case where the lifetimes may vary slightly throughout the image. In § 7, we analyse our methods on real data obtained that we received from SVI. We end the report by stating our conclusions, including recommendations for future work, in § 8.

3 Problem description

In this project, we will focus on reconstructing lifetimes of the dyes used in a fluorescence microscopy experiment. Fluorescent dyes, also called *fluorophores*, decay exponentially. That is, given a total number of molecules N_0 of a certain type excited by a laser, the number of molecules in an excited state at time t is modelled by

$$N(t) = N_0 e^{-t/\tau_0},$$

where τ_0 is the *mean lifetime* of the fluorophore. When multiple fluorophore types are present in the image, the observed decay is a sum of exponentials

$$N(t) = \sum_i N_i e^{-t/\tau_i}.$$

Observations of the data $N(t)$ can be made with precision at each pixel in an image of biological matter. In this report, we detail methods to predict the lifetimes of the fluorophores globally (once for an image) and use them to recover the underlying concentrations $N_k/\sum_i N_i$ of each fluorophore at each pixel. Due to interactions of the fluorophores with the sample, the lifetimes τ_i may also vary slightly per pixel, so we also consider methods that predict the lifetimes per pixel.

There are multiple approaches to determine the fluorescence lifetime of fluorophores. The golden standard is Fluorescence Lifetime Imaging Microscopy (FLIM) of which are two types: Time-Domain FLIM and Frequency-Domain FLIM. In the Time-Domain FLIM the time delay is measured between excitation and emission using a laser pulse. In the Frequency-Domain FLIM the phase shift in emission is measured using modulated excitation light. Generally, the Time-Domain approach is more accurate but slower than the Frequency-Domain approach [2]. In our case, the former is used.

Fluorescence lifetime estimation is typically done using three main methods: curve-fitting, phasor, and deconvolution techniques, each utilizing different mathematical approaches to extract lifetime information from time-resolved fluorescence data. For our project, SVI asked us to explore fluorescence lifetime estimation using kernel methods. While ‘kernel methods’ often refer to machine learning techniques, the term also applies to mathematical operations like integral transforms and deconvolution. Fourier and Laplace transforms, which involve kernel functions to analyse signals, fit within this broader definition of kernel-based methods. With this in mind, we applied Fourier and Laplace transforms as deconvolution methods and compared their results with those from a least squares curve-fitting method.

4 Background

This section gives an outline of some applications of fluorescence microscopy.

Fluorescence lifetime imaging microscopy is increasingly being integrated into biological and also medical research and clinical practice, offering a powerful tool for non-invasive diagnostics with its ability to distinguish fluorophores based on only their lifetimes. This makes it particularly useful for applications where molecular environment sensing and multi-dye imaging are essential. Besides being significantly less expensive than for example MRI, it also provides a safer, more convenient alternative to conventional dye-based imaging, where only intensity is measured.

FLIM has been successfully implemented in brain surgery to enhance tumour detection, transitioning from a prototype in early 2021 to clinical use by the end of 2022 [3, 4]. Beyond neurosurgery, it has shown potential in cardiovascular diagnostics, where intravascular FLIM has been used to assess plaque composition and identify damaged areas in the walls of arteries, aiding in early disease detection [5]. Additionally, FLIM has been applied to differentiate tissue types, demonstrating its ability to reveal biochemical and structural variations, making it a valuable tool for non-invasive medical imaging and diagnostics [6].

In biological research, FLIM is commonly used for studying protein-protein interactions, cell signalling and other molecular events [7]. These events are measured by determining the Förster Resonance Energy Transfer (FRET), which is a physical phenomenon where energy is transferred non-radiatively from a light-sensitive molecule to another through dipole-dipole interactions. Measuring FRET is highly specific since it occurs at a nanometre scale. High activity between proteins can suggest a cancer-related process is active.

Görlitz et al. [8] introduce a new microscope that combines super-resolution imaging (SIM) with fluorescence lifetime imaging to map molecular functions and their connection to biological structures at a finer scale than conventional microscopy. The method was used to study protein interactions and DNA changes in living cells, providing detailed insights into molecular behaviour within biological samples.

5 Data

5.1 Experimental data

5.1.1 Data structure

In this report, an *image* is a three-dimensional array of natural numbers. The first two dimensions in the array correspond to two spatial coordinates of the sample under consideration. The third dimension is temporal, binned at a resolution of 0.1 nanoseconds. We call each spatial coordinate pair a *pixel*. At pixel (x, y) and time bin t (in tenths of nanoseconds), the value d_t at (x, y, t)

corresponds to the number of photons detected by the microscope at pixel (x, y) in the time frame $(t, t + 1]$ over all the repetitions of the experiment.

In the dataset we consider in this report, we assume the presence of up to two distinct types of fluorophores at each pixel, with mean lifetimes τ_0 and τ_1 respectively. Thus, we hope to model the data at each pixel by the equation

$$N(t) = N_0 e^{-t/\tau_0} + N_1 e^{-t/\tau_1}.$$

5.1.2 Data pre-processing

We first pre-process the data at each pixel in the image into data we can model via exponential decay. Fixing now a pixel (x, y) , we consider the sequence $(d_0, d_1, \dots, d_t, \dots, d_l)$ of length $l + 1$ and define a sequence $(D_t)_t$ by

$$D_t = \sum_{i=t}^l d_i.$$

The value D_t represents the number of photons observed after time t , which we take as a proxy for the number of molecules in the excited state at time t and model by a sum of exponential decays.

5.1.3 Truncation

In the single-dye experiment, particles decay and emit photons exponentially. However, because observations are only done over a limited time frame, the obtained data might be incomplete: after a given time stamp, no data is collected. We thus introduce some error by fitting the equation $N(t) = N_0 e^{-t/\tau}$ directly to the observed data; particles might still decay and emit photons that would have been detected by the microscope if it collected data over a longer time frame. In this section, we analyse the influence of the truncated character of the data on the accuracy of estimating the lifetime. For the phasor method using the Laplace transformation, a model incorporating truncation is discussed in [9, *Cut-Off Correction*]; we discuss the problem in a larger generality here and discuss the Laplace transformation specifically in Appendix A.1.

As before, we assume an exponential decay model $N(t) = N_0 e^{-t/\tau}$ where N_0 is the total number of photons emitted at a particular pixel and τ is the lifetime. (Note that N_0 is the number of *emitted* photons, which might be larger than the number of *observed* photons.) Moreover, we let T denote the length of the time frame over which observations take place. Then we expect the number of photons that are emitted after time T , and thus not observed, to be

$$N(T) = N_0 e^{-T/\tau}.$$

In other words, we should theoretically model the observed data as $\tilde{N}(t) = N(t) - N(T)$. In our experimental dataset, we have $N_0 \sim 10^2$, $\tau \sim 10^1$ and $T \sim 10^2$, so the order of magnitude of this quantity is $N(T) \sim 10^{-3}$. Hence, the error introduced by using a non-truncated model is negligible for our purposes: almost all the observable emissions occur within the observation time frame. For datasets where the quantities N_0 and τ vary, for example if τ is much larger, care should be taken to observe over a time frame that is sufficiently long as to minimize this error.

6 Method

Given an image as described in § 5.1.1, we wish to predict, at each pixel, the relative concentration of each dye. In this section, we explain the methods we developed for making this prediction. We investigate assumptions that our methods are based on. We use computationally generated data, described in Appendix A.2, to evaluate different methods.

6.1 Experimental pipeline

We now describe the experimental pipeline we apply to the image to obtain the prediction of the relative concentration of each dye at each pixel. At each step in our pipeline, various approaches are considered.

Recall that the pre-processed data at pixel p is a sequence $(D_{p,0}, \dots, D_{p,t}, \dots, D_{p,l})$, where $D_{p,t}$ denotes the number of photons that remain to be observed at time t at pixel p . The first assumption in our model is the presence of pixels in the image at which the concentration of one of the dyes is low, and thus for which we may assume only one dye is present. This is also a key assumption in the phasor method. We assess the extent to which this is true in the experimental data we consider in § 7.

We begin by fitting a single exponential decay model to the pre-processed data at each pixel p . Recall that such a model is of the form

$$N_p(t) = N_{p,0}e^{-t/\tau_p},$$

where τ_p is the model’s predicted mean lifetime at pixel p . The prediction is evaluated against the experimental data by computing the mean squared error. Explicitly, we compute

$$\text{MSE}(D_p, N_p) = \frac{1}{l+1} \sum_t (D_{p,t} - N_{p,t})^2.$$

In this report we compare three approaches to fitting a single exponential decay model to fluorescence imaging data. We use nonlinear least squares as a baseline method. We then compare the performance of the Fourier transform method and the Laplace transform method to this baseline. This comparison is carried out in § 6.2.1.

The next assumption in our pipeline is that the pixels for which our single dye model yields a low mean square error are pixels where only one dye is present. This assumption is evaluated on synthetic data in Appendix A.3. Under this assumption, the τ values predicted by the single dye model at pixels where the MSE is low can be used to predict the lifetimes of the two present dyes (remember that we assume only two distinct dyes are present). Further assuming that these predicted τ values will follow a mixture of two normal distributions with different means, we fit a Gaussian Mixture Model (GMM) to a histogram of the lifetime values to obtain predictions for the underlying τ_0 and τ_1 values of dyes 0 and 1; see § 6.3.

Using the predicted mean lifetimes of the pair of fluorophores in the image, the next step in our pipeline is to fit a model of the form

$$N_p(t) = N_{p,0}e^{-t/\tau_0} + N_{p,1}e^{-t/\tau_1}$$

to the pixels for which the single dye fit is poor. The assumption is that at these pixels the concentration of both dyes is nonzero, and thus that a multi-exponential model better describes the system. From this fit we recover $N_{p,0}$ and $N_{p,1}$, the predicted number of excited molecules of each fluorophore type. When the τ_0, τ_1 values are assumed to be constant over the image, this is a model with one unknown parameter, since we have that $D_0 = N_p(0) = N_{p,0} + N_{p,1}$. We again assess the suitability of the Fourier transform and Laplace transform models in predicting the $N_{p,i}$ against a nonlinear least squares baseline model.

As an alternative to the assumption that the mean lifetimes are fixed throughout the image, we may allow the lifetimes to vary per pixel. Such variation could arise, for example, due to interaction of the fluorophores with matter local to regions of the image. In this case, the multi-exponential model has *three* unknowns, $\tau_{p,0}, \tau_{p,1}$ and $N_{p,0}$. In both the Fourier transform and Laplace transform approaches, predicting these parameters requires solving a system of quadratic equations. We describe this procedure and evaluate it on synthetic data in § 6.4.2.

6.2 Fitting a single-dye model to detect single-dye pixels

Given a pixel p , our assumption is that there are unknown concentrations N_0, N_1 of the two dyes in the image. However, since each dye binds to specific molecules, we further assume that there are at least some pixels in the image for which the value N_0 or N_1 is small or zero. If these pixels can be identified, their data can be used to infer the mean lifetimes of the dyes present in the sample by fitting a single exponential decay model $N(t) = Ne^{-t/\tau}$. These predicted mean lifetimes can be used to predict the relative concentrations of the fluorophores for the pixels where both are present. In this section, we detail and analyse three methods to predict the mean lifetime of a fluorophore given single exponential decay data.

Note that under the assumption that there is low to little variation of the lifetimes across the image, the number of pixels for which one of the N_i is low or zero need not be high. Indeed, theoretically, one such pixel for each N_i would suffice to obtain a prediction of τ_i . However, the assumption that there exist pixels with only one fluorophore present is stronger than the assumption made by the standard *phasor* method, which only requires the sample to be heterogeneous in concentration of the N_i and uses this variation in the data across the image to infer τ_0 and τ_1 . Overall, however, our analysis of the experimental data supports the hypothesis that pixels with only one fluorophore are present in the image, see § 7.

6.2.1 Single-dye models for pixels with a single dye present

We evaluate three methods of fitting an exponential decay model $N(t) = Ne^{-t/\tau}$ to synthetic data that models a single type of fluorophore: Nonlinear least squares, the Fourier transform method, and the Laplace transform method.

Nonlinear least squares We use the nonlinear least squares method as a baseline. The `curve_fit` function from `scipy.optimize` is used as an implementation.

Fourier and Laplace transform The Laplace transform of the exponential decay model $N(t)$ is given by

$$\mathcal{L}[N](s) = \frac{N_0}{s + 1/\tau}$$

for $s \in \mathbb{C}$. We view the Fourier transform as a special case of the Laplace transform where the frequency domain is purely imaginary, i.e. we have $\mathcal{F}[N](\omega) = \mathcal{L}[N](i\omega)$ for $\omega \in \mathbb{R}$. Rearranging the equation for the Laplace transform, we obtain an equation for the lifetime:

$$\tau = \frac{\mathcal{L}[N](\omega)}{N_0 - s \cdot \mathcal{L}[N](s)}.$$

We estimate the term $\mathcal{L}[N](s)$ from simulated data by numerical integration using Simpson's rule, and then use the equation above to obtain an estimate for τ .

To evaluate the suitability of these methods, we assess their predictions on simulated data generated across a range of parameters. In Figure 1, the prediction of τ is displayed for each method as the number of samples, the underlying τ , and the cutoff value T (see § 5.1.3) vary. While the Fourier and Laplace transforms show greater variation across trials than the baseline method, they both converge to the true τ values whenever $\tau \lesssim 50$ as long as the cutoff is larger than about $T = 100$. We also observe that the baseline method fails to converge to the correct τ value for $\tau > 50$. This is explained by the truncation error increasing for large τ values. For example, for $\tau = 100$, $N_0 = 500$, $T = 300$, the error introduced by truncation is $N(T) \approx 25$.

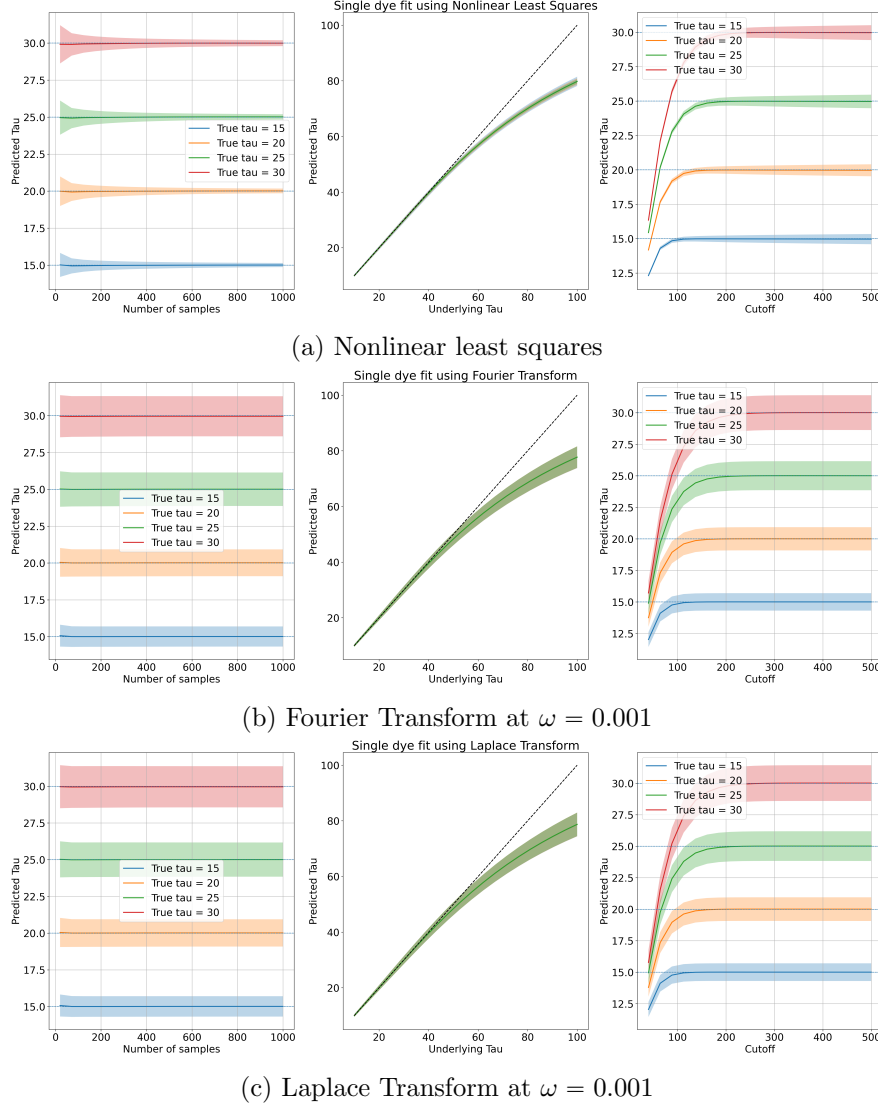


Figure 1: **Fitting a single-dye model to synthetic single-dye data.** Predicted τ values by each model as parameters used to generate synthetic data are varied. Using truncated noisy exponential decay data with $N_0 = 500$, truncation cutoff $T = 250$ (except for right-most plots, where it is varied, and for which the number of samples is 300), noise coefficient $N_{\text{scale}} = 1$. Solid lines are mean values across 200 trials. Shaded areas show standard deviation across trials.

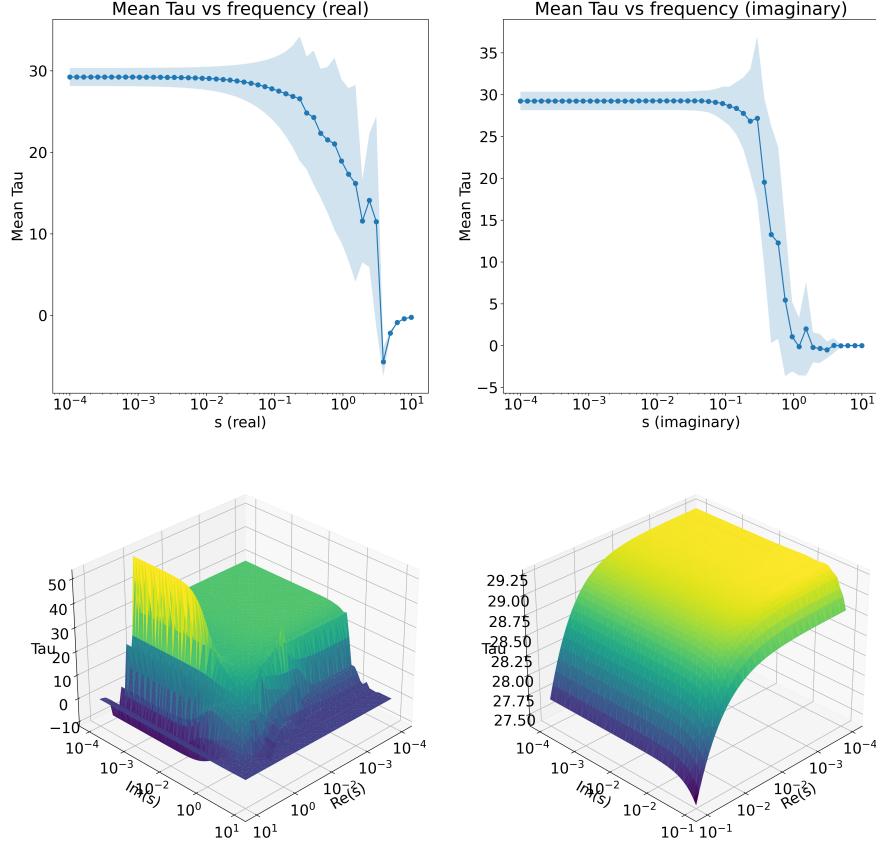


Figure 2: **Predicted τ by Laplace transform method as frequency is varied.** True lifetime is $\tau = 30$. Using truncated noisy exponential decay data with $N_0 = 500$, truncation cutoff $T = 250$, noise coefficient $N_{\text{scale}} = 1$. In the top row of plots, purely real and purely imaginary frequencies are used to compute the Laplace transform, from which a predicted τ is obtained. Solid lines are mean predicted τ values across 200 trials, shaded areas show standard deviation across trials. In the bottom row, the frequency is allowed to vary across subsets of $\mathbb{C}$. The left subplot varies frequencies in $[10^{-4}, 10^0]^2$; the right subplot varies frequencies in $[10^{-4}, 10^{-1}]^2$. Mean predicted τ values across 200 trials are displayed.

The Laplace and Fourier transform methods rely on a choice of frequency parameter. In Figure 2 the frequency parameter is varied across subsets of the complex plane. We obtain excellent prediction of τ for $|s| < 10^{-2}$, and consistently incorrect predictions for larger values. We hypothesize that when the Laplace transform values are lower than the tolerance of the numerical integrator (which occurs when the frequency increases), the predicted τ becomes meaningless.

6.2.2 Identifying pixels in the image with only a single dye present

After fitting each pixel with an exponential decay model that assumes the pixel contains only a single fluorophore, we compute the mean square error (MSE) of the predicted decay curve and the observed decay at each pixel. In our analysis of the experimental data we discard pixels where the MSE is above 2.0. We choose the cutoff manually after inspection of the number of pixels where the single-dye fit is below various cutoff candidates. See § 7 for details on experimental results.

6.3 Gaussian mixture models to predict lifetime from single-dye analysis

After identifying the pixels where (approximately) only a single dye is present, we can estimate the lifetimes of the two dyes over the entire image. Combining the obtained per-pixel lifetime estimations, we obtain a histogram of lifetimes. If the lifetimes of the two dyes are sufficiently different, we assume the obtained lifetime values follow a bimodal normal distribution; this assumption is verified by the provided experimental data (see § 7). The two means of the mixed normal distributions should be the values of the lifetimes of the two dyes. We use a Gaussian Mixture Model from the Python library `scikit-learn`¹ to estimate these two means, providing estimates for the lifetimes of the two dyes. Moreover, we use the obtained estimated standard deviations as bounds for estimating per-pixel lifetimes when we do not assume the lifetimes are constant over the entire image.

6.4 Double-dye analysis

After predicting the lifetimes τ_0 and τ_1 from the single-dye analysis, we fit the multi-exponential model

$$N(t) = N_0 e^{-t/\tau_0} + N_1 e^{-t/\tau_1}$$

to estimate the concentrations of the two dyes. In this section, we investigate the Laplace method for estimating N_0 and N_1 .

6.4.1 Fixed lifetimes

In this subsection, we investigate how the assumption that the lifetimes τ_0 and τ_1 are constant influences the accuracy of our model. That is, we assume that the lifetimes of the two dyes do not vary across the image.

We will investigate four different parameters for our model, and how they influence model accuracy (under the constant-dye-lifetime assumption). The parameters are the s_0 value, *noise level*, the *cutoff value*, and the *number of samples*. We will consider these one by one.

Across all these graphs, if a parameter is not varying in a graph, it is consistently chosen as follows: $s_0 = 0.01$, $N_{scale} = 0.5$, cutoff is 300, number of samples is 500.

For each parameter, we show plots demonstrating its impact on model fit. We plot these for three different pairs of values for τ_0 and τ_1 (namely $(\tau_0, \tau_1) \in \{(14, 32), (25, 30), (2, 45)\}$). The idea behind picking these pairs is that it should give some insight into how resilient a method is to two dyes with very similar lifetimes. The data for these plots are generated for 200 trials, where for each trial, we randomly generate noisy data and fit a model to it, with fixed N_0, N_1, τ_0, τ_1 , with the chosen parameter varying over the horizontal axis. We draw a line for the mean fitted N_0 and N_1 across the different trials, and the surrounding shading shows the standard deviation for the fit across the trials.

We also show the impact of varying each of these parameters on the nonlinear least squares method (implemented using SciPy's `curve_fit`) as a baseline. The parameter s_0 is specific to the Laplace model and thus not varied for the baseline model.

The influence of the s_0 parameter is shown in Figure 3. Similar to the single dye case, we observe that small frequencies yield better predictions, with predictions becoming consistently incorrect across trials for $s_0 \gg 10^{-2}$. This effect is less pronounced as the difference between lifetimes increases.

¹<https://scikit-learn.org/>

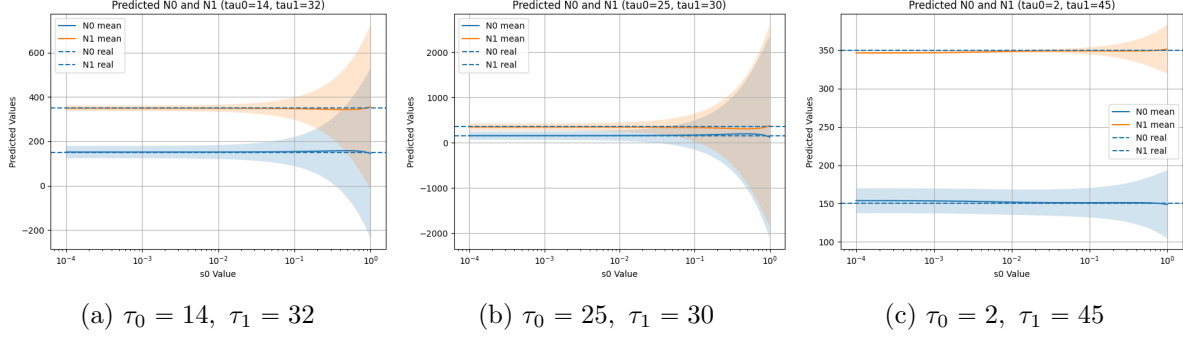


Figure 3: Effect of choosing s_0 on predicted N_0 and N_1 of Laplace method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

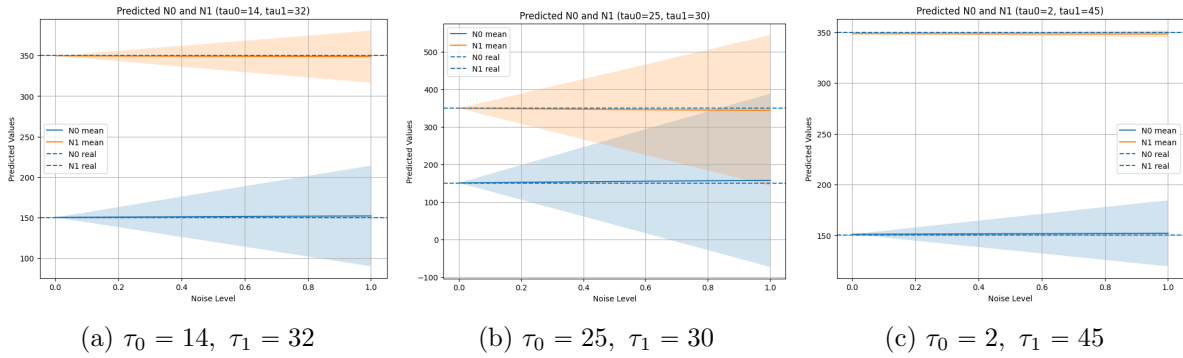


Figure 4: Effect of choosing noise level on predicted N_0 and N_1 of Laplace method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

In Figure 4, we investigate the effect of noise on the accuracy of the Laplace method. We observe increased standard deviation across trials as the noise parameter is increased, with accuracy of the mean prediction being consistently high across noise levels.

In Figure 5 we compare the influence of noise on the Laplace model's performance to its influence on the `curve_fit` model's performance. We see that the Laplace method is approximately an order of magnitude more sensitive to noise than the `curve_fit` method.

As discussed in § 5.1.3, the finite recording length of measured decay introduces a bias term in our model which we do not correct for. When recording duration (truncation) is small, this term is large, and we thus expect incorrect prediction in the double dye case, similar to the failure to correctly predict lifetimes in the single dye case. Figure 6 and Figure 7 show the effect of the length of recording for the Laplace and `curve_fit` methods, respectively. Both models converge relatively quickly to the correct N_0 and N_1 value as the cutoff value increases, in behavior similar to the single dye models.

Finally, we investigate the impact of the number of samples. The graphs are shown in Figure 8 and Figure 9. Most notably, these plots show that even a hundred samples can be enough to give an accurate prediction for N_0 and N_1 , both for the Laplace model and the least squares model. Note also that the variance in the Laplace model's prediction is constant over the number of samples after convergence, while the least squares model shows decreasing variance as the number of samples increases.

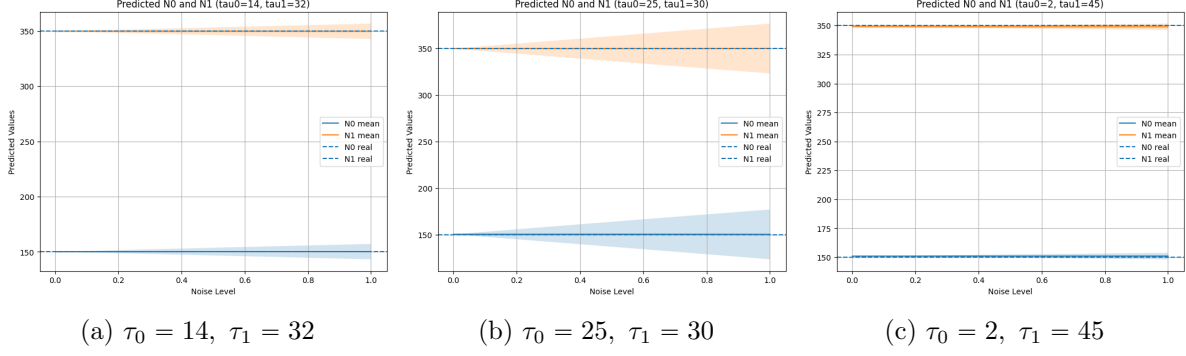


Figure 5: Effect of choosing noise level on predicted N_0 and N_1 of `curve_fit` method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

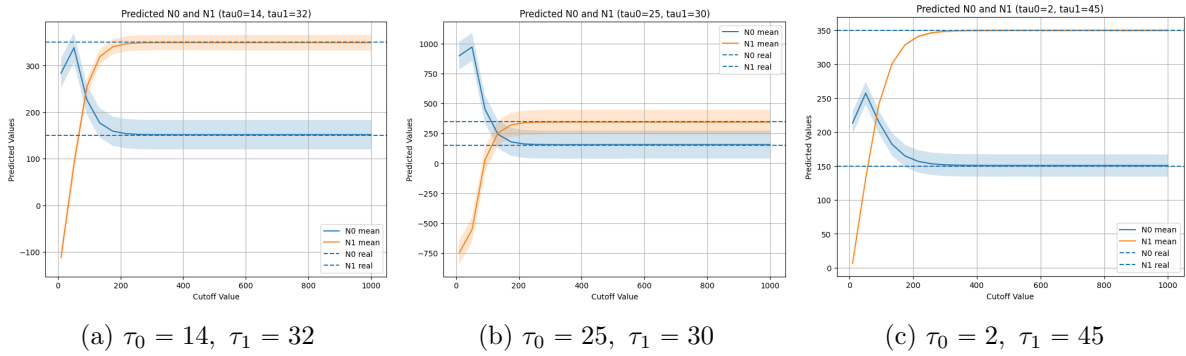


Figure 6: Effect of choosing cutoff level on predicted N_0 and N_1 of Laplace method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

6.4.2 Allowing lifetimes to vary per pixel

The methods from the previous subsection assume the lifetimes of the dyes do not vary per pixel. In practice, however, the lifetimes across the image may vary due, for example, to interactions with the chemical environment in the sample. Such variation would make the methods from the previous subsection less accurate on real world data.

6.4.2.1 Laplace Method

Recall that we want to fit a model of the form

$$N(t) = N_0 e^{-t/\tau_0} + N_1 e^{-t/\tau_1} \quad (6.1)$$

to our data. The Laplace transform of $N(t)$ at a frequency $s \in \mathbb{C}$ is given by

$$\begin{aligned} & \int_0^\infty \left(N_0 e^{-t/\tau_0} + N_1 e^{-t/\tau_1} \right) e^{-st} dt \\ &= N_0 \int_0^\infty e^{-t/\tau_0 - st} dt + N_1 \int_0^\infty e^{-t/\tau_1 - st} dt \\ &= N_0 \left[\frac{-1}{s + 1/\tau_0} e^{-t/\tau_0 - st} \right]_{t=0}^{t=\infty} + N_1 \left[\frac{-1}{s + 1/\tau_1} e^{-t/\tau_1 - st} \right]_{t=0}^{t=\infty} \\ &= \frac{N_0}{s + 1/\tau_0} + \frac{N_1}{s + 1/\tau_1}. \end{aligned}$$

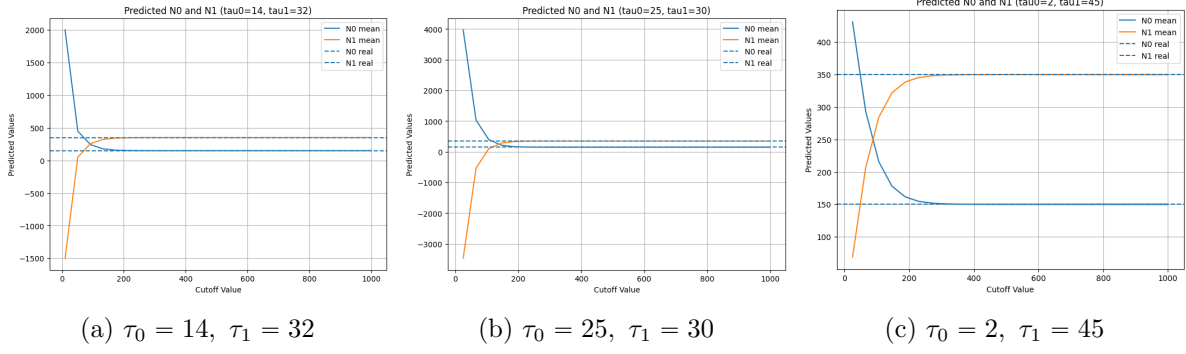


Figure 7: Effect of choosing cutoff level on predicted N_0 and N_1 of `curve_fit` method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

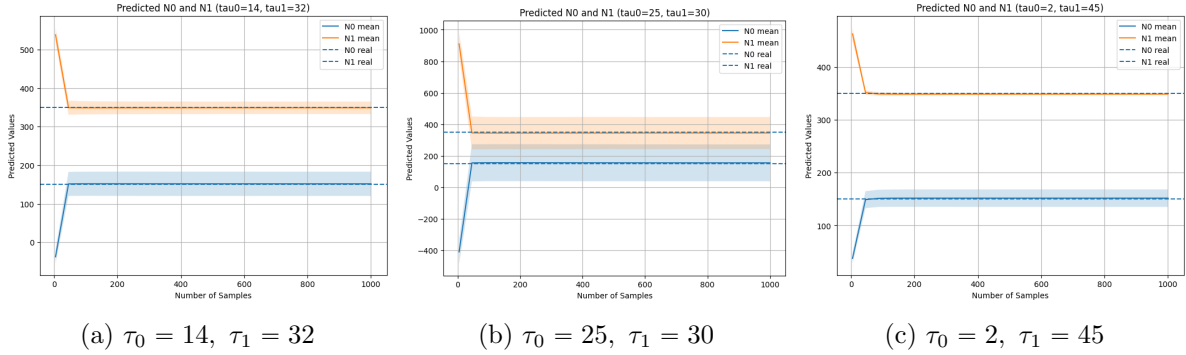


Figure 8: Effect of number of samples on predicted N_0 and N_1 of Laplace method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

If τ_0 and τ_1 are unknown (in addition to N_0 and N_1), applying the Laplace transformation for some s_0 is not enough, since it only gives two equations for four unknowns. However, we can repeatedly apply the Laplace transformation for different frequencies to get additional equations relating our unknowns. We will now describe this in detail. We consider three frequencies s_0, s_1, s_2 and the values from the Laplace transform at these frequencies: z_0, z_1 and z_2 respectively. The values of z_0, z_1 and z_2 are again estimated by performing numerical integration using Simpson's rule. We then solve the following system of equations numerically:

$$\begin{aligned}
 z_0 &= \frac{N_0}{s_0 + 1/\tau_0} + \frac{N_1}{s_0 + 1/\tau_1} \\
 z_1 &= \frac{N_0}{s_1 + 1/\tau_0} + \frac{N_1}{s_1 + 1/\tau_1} \\
 z_2 &= \frac{N_0}{s_2 + 1/\tau_0} + \frac{N_1}{s_2 + 1/\tau_1} \\
 N &= N_0 + N_1.
 \end{aligned} \tag{6.2}$$

Here N denotes the total number of measured photons at a given pixel, which can be determined directly from the data. We choose three frequencies in order to obtain a system of four equations, since there are four variables (τ_0, τ_1, N_0, N_1) that need to be computed.

To simplify our implementation, we substitute $N_1 = N - N_0$ and solve a system of three equations, and then compute N_1 from our prediction of N_0 . The `least_squares` method requires

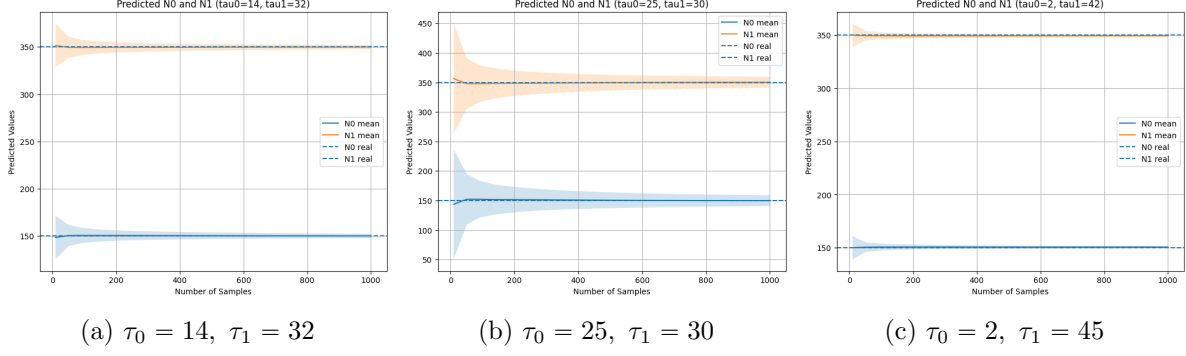


Figure 9: Effect of number of samples on predicted N_0 and N_1 of `curve_fit` method, constant lifetimes τ_1, τ_2 . The real concentration values are $N_0 = 150$ and $N_1 = 350$.

an initial guess and bounds for the variables that are to be determined. Let τ'_0, τ'_1 be the correct values. For τ_0 we use 0 and halfway between τ'_0 and τ'_1 as bounds. For τ_1 we use halfway between τ'_0 and τ'_1 and 100 as bounds. The initial guesses for τ_0 and τ_1 respectively, are halfway between their bounds. The bounds we are using in this section are slightly wider than bounds obtained from real data. This way, we are not inflating the performance of our methods by using unrealistically accurate bounds. Our initial guess for N_0 is N , the measured number of emitted photons. We do not set bounds on N_0 and N_1 . When applying this method in production, we suggest basing the bounds of τ_0 and τ_1 on the standard deviations on globally predicted lifetimes obtained in in § 6.3.

We use Sage to determine the number of solutions of the system (6.2) for various values of $z_0, z_1, z_2, s_0, s_1, s_2$ and N . For the pairs $(\tau_0 = 14, \tau_1 = 32)$, $(\tau_0 = 25, \tau_1 = 30)$ and $(\tau_0 = 2, \tau_1 = 45)$, we show that the system (6.2) has two rational solutions for all options, when s_0 ranges over $\{0, 0.01, \dots, 0.19\}$, s_1 ranges over $\{2, 2.1, \dots, 3\}$, s_2 ranges over $\{30, 31, \dots, 39\}$ and N_0 ranges over $\{100, 200, 300, 400\}$. Here we take $N_1 = 500 - N_0$. Since the system (6.2) is symmetric, this actually gives us a unique rational solution. We expect that for reasonable values (for example s_0, s_1 and s_2 are pairwise distinct), the system (6.2) has a unique rational solution. This means that, in theory, our method will solve the problem. In Appendix B we elaborate further on this method and describe an analytic solution to the system of equations in the case that $s_1 = 2s_0$ and $s_2 = 3s_0$. Here we also analyse the stability of this solution when noise is added. We find that even with small amounts of noise, this method is unreliable and can predict nonsensical solutions.

6.4.2.2 Performance of methods for varying lifetimes

Figures 10 to 16 show the performance of the methods when varying different parameters. The line gives the mean value after running our methods 200 times at each data point. The shaded area gives the standard deviation. A large shaded area shows that our methods don't work very well. During real world applications, we do not run our methods 200 times. When a parameter is not the one that is being varied, we use the following standard values. The graphs support the idea that these are reasonable choices. As explained in Appendix A.4, we choose 300 as our truncation value. We use $s_0 = 0.01, s_1 = 0.05, s_2 = 0.02$. The standard noise is 0.5 and the standard number of points is 500. We use $N_0 = 150$ and $N_1 = 350$. We investigate three cases $(\tau_0 = 14, \tau_1 = 32)$, $(\tau_0 = 25, \tau_1 = 30)$ and $(\tau_0 = 2, \tau_1 = 45)$. As a benchmark to compare our method against, we again use a nonlinear least squares method (implemented using `curve_fit` from `scipy`) to directly fit a solution to (6.1). Since we predict the N_1 value by computing $N - N_0$, one would expect the plots to contain a certain symmetry. Note that the standard deviations are not symmetric in all plots.

The reason for this is that we also add noise to the first data point, so the N varies per run.

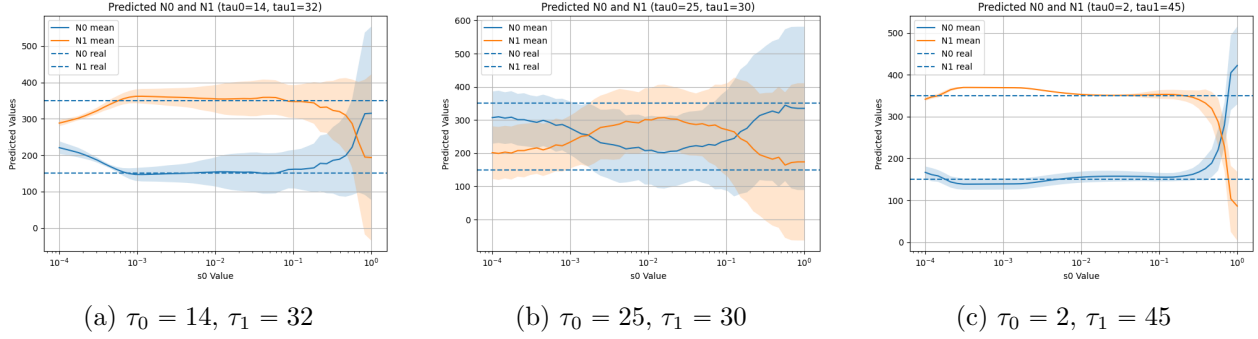


Figure 10: Effect of choice of s_0 on predicted N_0 and N_1 of Laplace method, varying lifetimes τ_1, τ_2 .

Parameter: s_0 When constructing Figure 10, we used $s_1 = 5s_0$ and $s_2 = 2s_0$. Figure 10 shows that the choice of s_0, s_1 and s_2 has a large influence on the accuracy of our method. In these cases our method seems to work best for $s_0 = 0.01$. We again hypothesize that for large values of s_0, s_1, s_2 , the approximated Laplace transforms are smaller than the numerical integrator's tolerance, yielding poor predictions. For small values of s_0 , our choice of s_1 and s_2 as multiples of s_0 means all three values are very similar, so the resulting equations are also very similar. In this case we hypothesize that the problem becomes under-determined. We thus hypothesize that choosing $s_0 = 0$ (which offers the best performance in the single dye case) and s_1, s_2 evenly spaced over $[0, 10^{-2}]$ will offer better model performance than the choices in our figure, but we leave this analysis to future work.

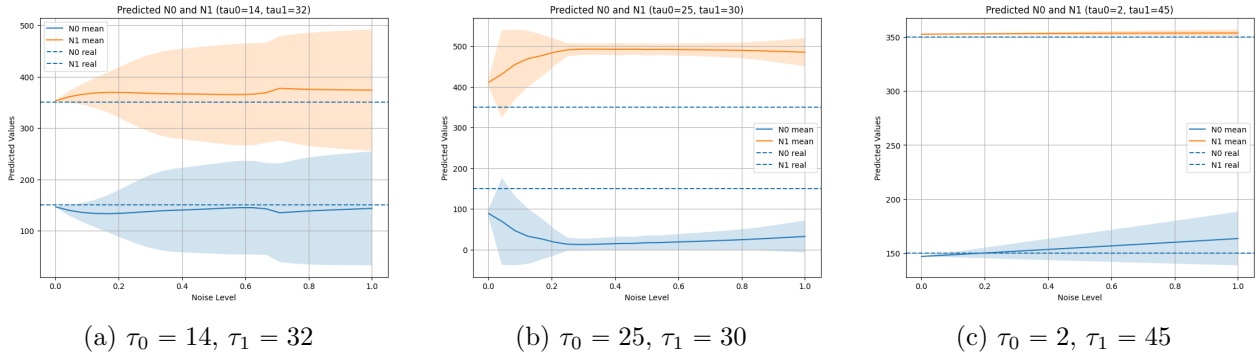


Figure 11: Effect of noise on predicted N_0 and N_1 of Laplace method, varying lifetimes τ_1, τ_2 .

Parameter: noise In Figure 11b, the standard deviation decreases as the noise increases to roughly 0.2. One would expect the standard deviation to increase as the noise increases. However, note that the mean prediction approaches $N_1 = 500$ and $N_0 = 0$. So as the noise increases, our method collapses to a single dye prediction. This explains the decrease in the standard deviation.

Parameter: cutoff value In Figures 13 and 14, we see that both methods behave quite poorly when the data has large truncation (cutoff value under 300). After this, the performance of both methods seem to stabilize. Note that 300 is a realistic cutoff value for our real world data.

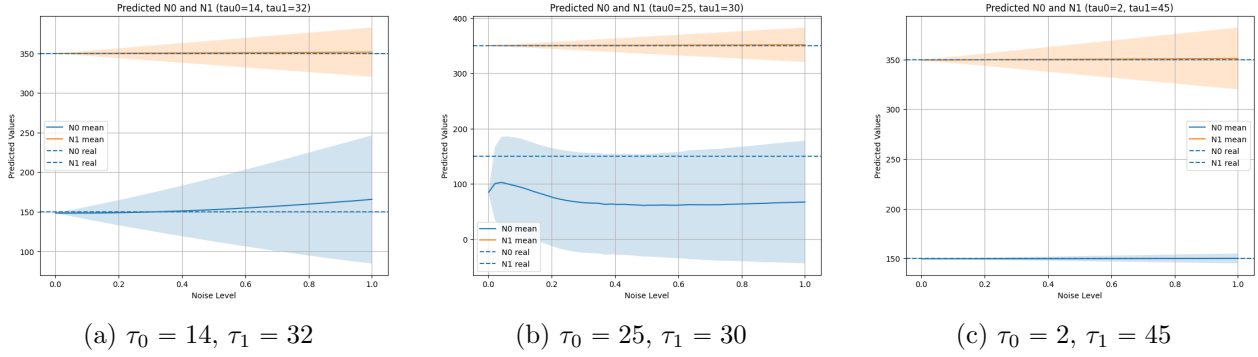


Figure 12: Effect of noise on predicted N_0 and N_1 for `curve_fit` method, varying lifetimes τ_1, τ_2 .

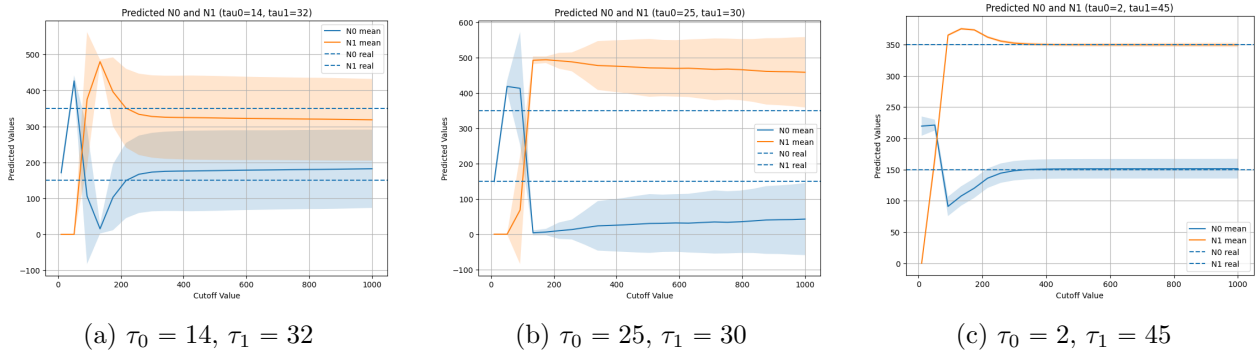


Figure 13: Effect of cutoff value on predicted N_0 and N_1 of Laplace method, varying lifetimes τ_1, τ_2 .

Parameter: number of samples In Figures 15 and 16, we see that the number of samples has very little impact on the performance of the methods. As expected, both methods behave poorly if the number of data points is very low (roughly below 50). This means that on real world data, the number of points will almost never be a limitation.

6.4.2.3 Overview of performance of methods for varying lifetimes

Overall, Figures 10 to 16 show that whether the methods work well depends heavily on the data's underlying parameters. If the τ values are similar, the Laplace method and the `curve_fit` method have trouble approximating the correct value. If the τ values are far apart, the method works quite well. If possible, biologists should choose dyes that are as far apart as possible when they expect the lifetimes to vary per pixel. In § 7 we investigate whether letting the lifetimes vary gives better predictions on real world data. The Laplace method is very dependent on the quality of the numerical solver. Future work could focus on further enhancing the numerical solver. We expect that this would yield better predictions. Allowing the τ values to vary means the Laplace transform needs to be combined with a numerical solver applied to a system of quadratic equations, which increases computational complexity.

In Figures 11 to 16, we see that the `curve_fit` method, in general, gives better predictions than the Laplace Method when lifetimes are allowed to vary.

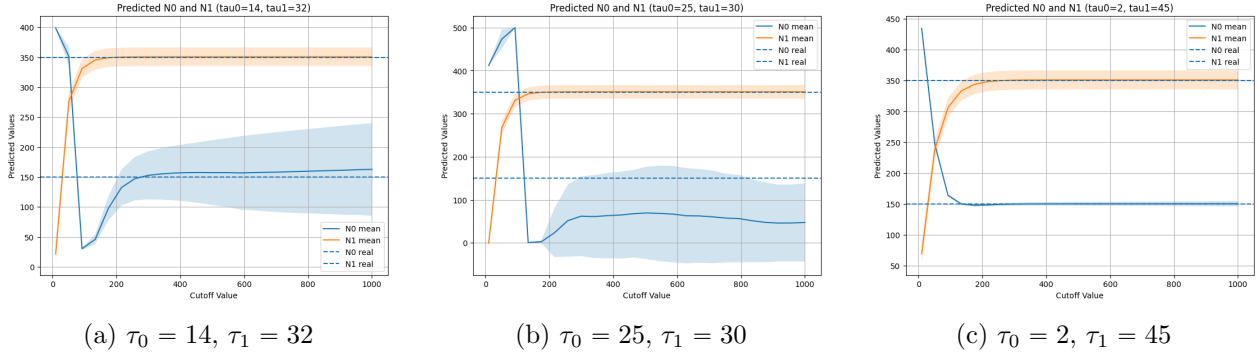


Figure 14: Effect of cutoff value on predicted N_0 and N_1 for `curve_fit` method, varying lifetimes τ_1, τ_2 .

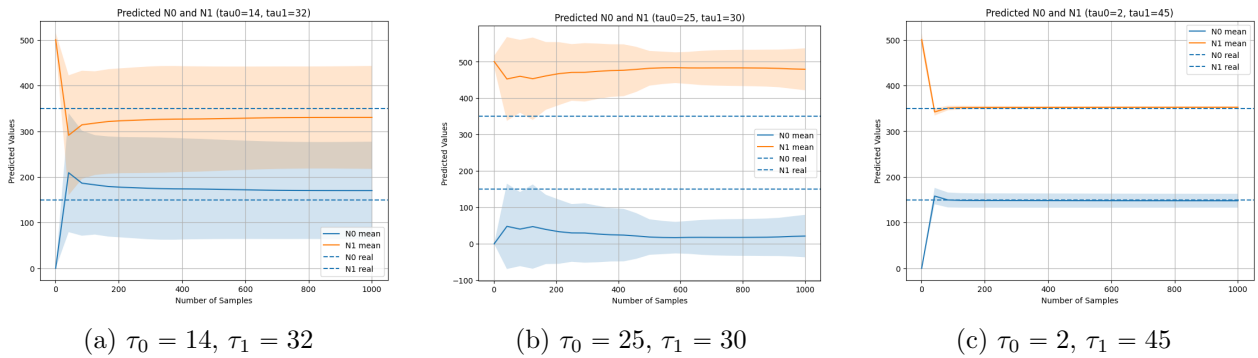


Figure 15: Effect of number of samples on predicted N_0 and N_1 of Laplace method, varying lifetimes τ_1, τ_2 .

7 Analysis on real data

So far, we have only tested the Fourier and Laplace methods on synthetic data. In this section, we analyse how the methods considered in this report perform on real data. The data we use was very generously provided to us by the team at SVI. This data was provided to us in raw form, thus we first pre-process the data at each pixel as discussed in § 5.1.2 to obtain the observed decay functions.

Firstly, we fit a single exponential curve to each pixel in the image. As discussed in § 6.1, we assume that pixels where the fit has a mean squared error below some cutoff `max_mse` contain one dominant dye and therefore give a good estimate of the true lifetime of said dye. The first method we use is the curve fit method, which fits a single exponential to the decay curve obtained from each pixel using least squares. The second and third methods we use are the Fourier and Laplace transforms discussed in § 6.4.2.1. For the Laplace method, we set the frequency parameter to $s = 0.01$, which is heuristically the best value we found (see Figure 10). To compute the Fourier transform, we set the frequency parameter to $s = 0.01i$. The Fourier and Laplace methods ran in 50 seconds as opposed to 5 minutes for the curve fitting method. Surprisingly, the results obtained are very similar. In Figure 17 we plot the MSE and lifetime heatmaps for each of the three models. In Figure 18 the histograms qualitatively look like bimodal distributions, thus we apply a Gaussian mixture model to the τ histogram (see § 6.3) to obtain estimates for the average τ values. Table 1 contains the corresponding lifetime values representing the two means of the bimodal distribution up to their respective standard deviations. From this table, we see τ values predicted by the three

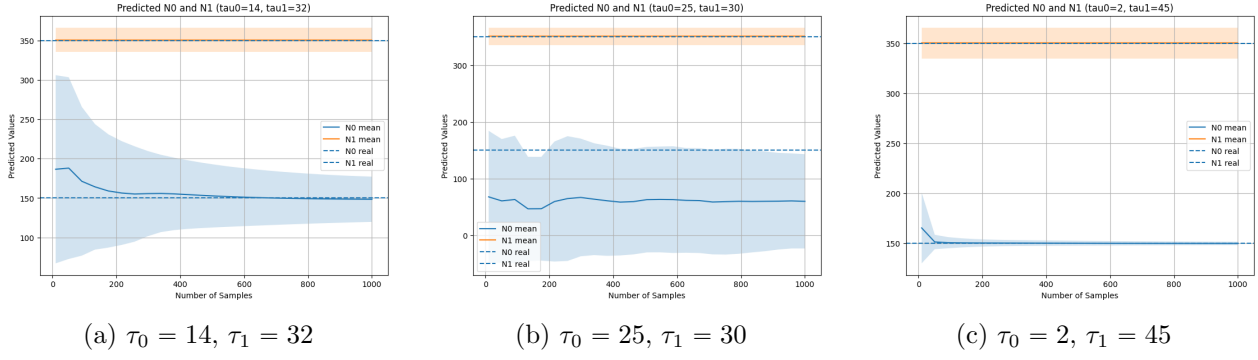


Figure 16: Effect of number of samples on predicted N_0 and N_1 for `curve_fit` method, varying lifetimes τ_1, τ_2 .

models only differ by a small amount. Finally, using the obtained lifetime values from each model and assuming the lifetimes remain constant across the image, we compute the relative concentrations of the dyes across the image as shown in Figure 19. We see that the models give virtually the same images, which is likely since the predicted τ 's are very similar.

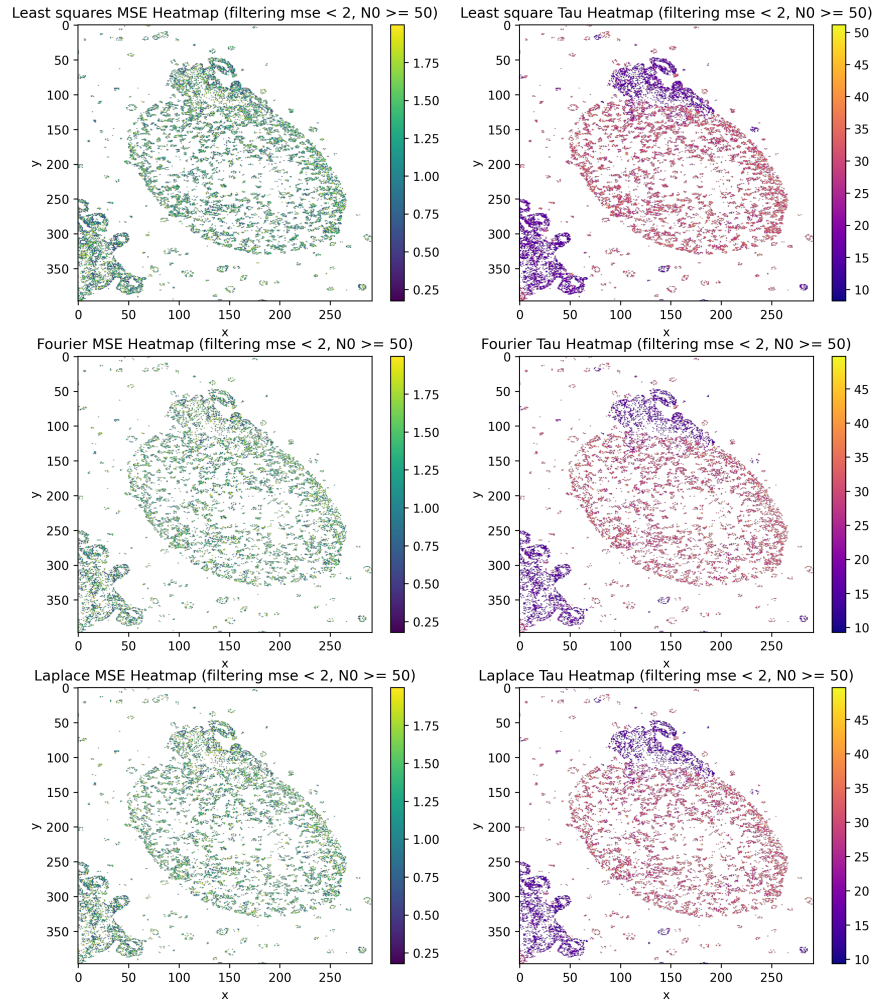


Figure 17: The per pixel MSE's and τ predictions for the three models.

	Least squares	Fourier	Laplace
τ_1	15.35 ± 2.06	15.22 ± 2.00	15.16 ± 2.00
τ_2	28.6 ± 4.64	28.39 ± 4.64	28.29 ± 4.6

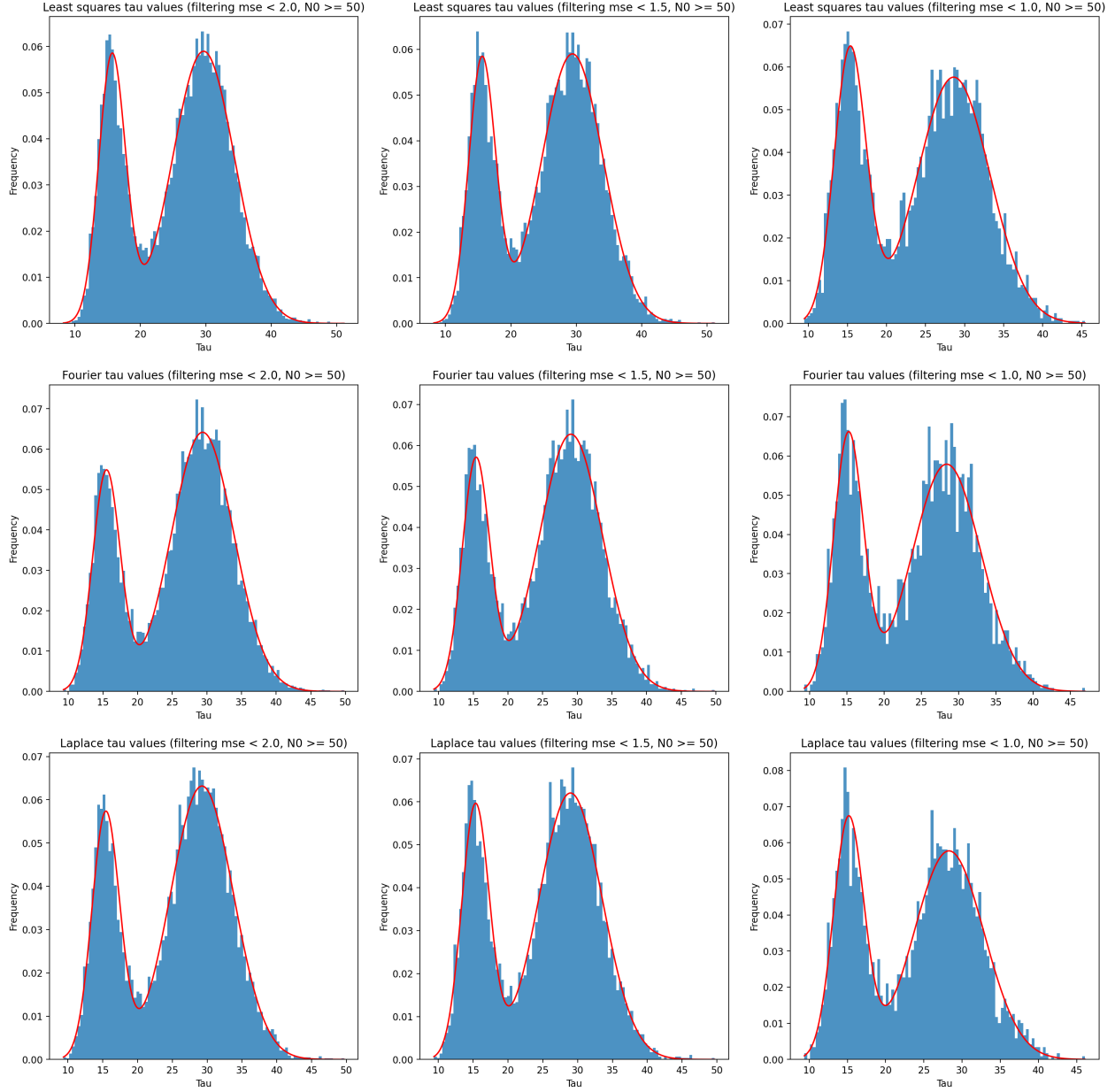
Table 1: The τ values we obtain from the bimodal fit, including standard deviations.

Figure 18: Bimodal fit on the histograms for each model.

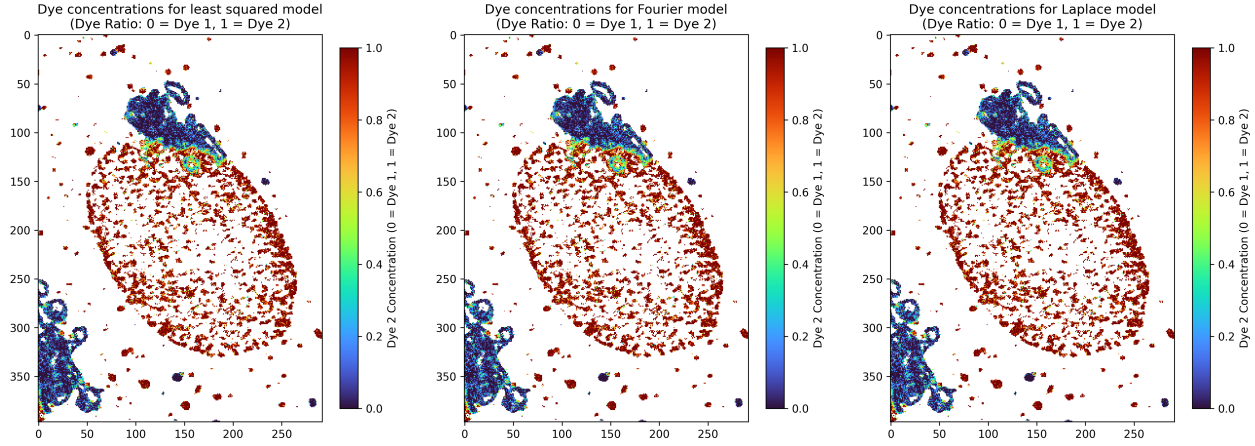


Figure 19: The dye concentration plots for the three single dye methods.

Finally, we plot the results obtained from the two dye Laplace method (see § 6.4). For this method we restrict the solution space of the τ 's. In Figure 21, we again plot the concentrations of the two dyes per pixel and a heat map of the MSE values for pixels with $\text{MSE} < 5$ (in Figure 20 we found that most pixels have an MSE of < 5 and therefore this is a reasonable cutoff). Here the solution space for the τ was taken to be $\tau_0 \in [10, 20]$ and $\tau_1 \in [25, 35]$. To compute the MSE we take the MSE of the difference between the double exponential curve obtained from the predicted values and the real data at every pixel. We see that the per-pixel fit is relatively good while, unfortunately, the predicted τ values accumulate around the bounds. This issue persists if we restrict the bounds further, in fact imposing stricter bounds for this method leads to a higher average MSE as shown in Figure 22 in which the error is plotted for different bounds where $|\tau_1 - 15| \leq \text{bound}$ and $|\tau_2 - 30| \leq \text{bound}$.

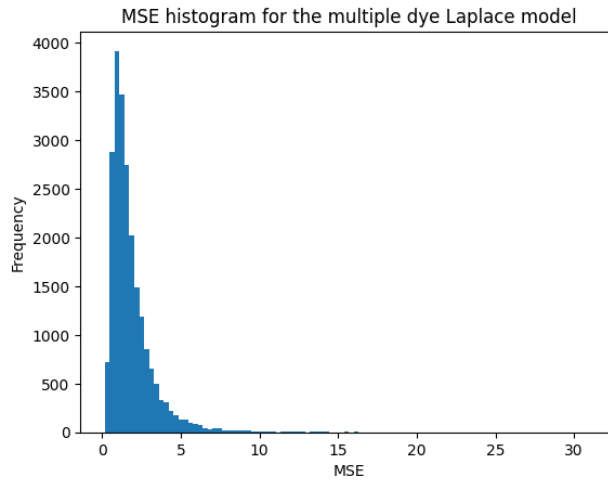


Figure 20: The MSE histogram for the varied τ two-dye Laplace method.

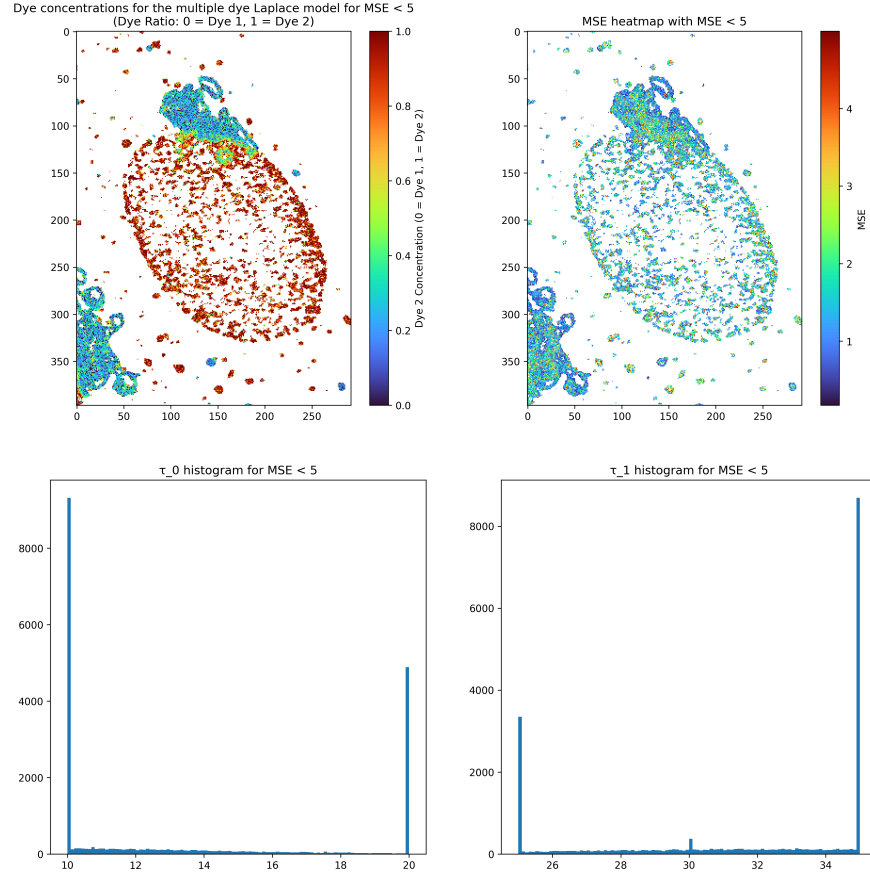


Figure 21: The dye concentration plot for the varied τ two-dye Laplace method, along with the τ_0 and τ_1 histograms.

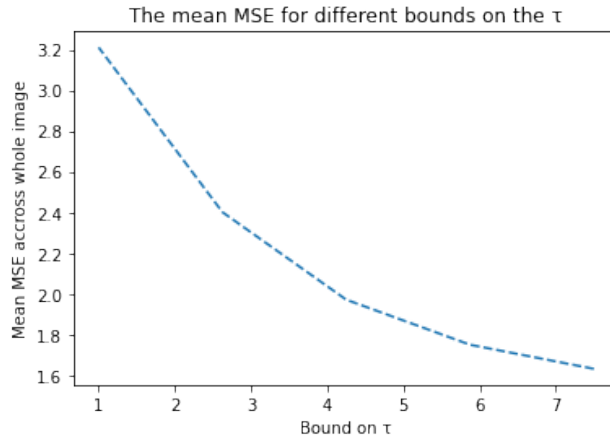


Figure 22: The average MSE of the varied τ two-dye Laplace method for different bounds where $15 - \text{bound} \leq \tau_1 \leq 15 + \text{bound}$ and $30 - \text{bound} \leq \tau_2 \leq 30 + \text{bound}$.

8 Conclusion

The analysis of our experimental pipeline using synthetic data highlights the suitability of the Fourier transform and the Laplace transform to quickly identify underlying lifetimes in an image with two fluorophores. For pixels where only a single dye is present, choosing frequencies smaller than 10^{-2} yields accurate estimators for a wide variety of parameters of synthetic data, with as few as 100 samples recorded over with a cutoff of at least $T \approx 200$. We only found a failure of the Laplace transform method to converge when τ was large ($\tau \approx 50$ and higher), but this failure was also observed for our baseline model which used Least Squares Regression. While across many trials the Laplace transform method returns an accurate mean prediction, one drawback of this method compared to our baseline method is its higher variance. Since global lifetime estimates are obtained by considering data from many pixels in the image when applied in our experimental pipeline, we do not expect this higher variance to be an issue experimentally. Indeed, in Figure 18 our pipeline applied to experimental data shows very similar predicted lifetimes for the baseline model and the Laplace/Fourier transform models. Furthermore, the Laplace transform method is much faster than our baseline, since it is linear with respect to the number of data points in the sample.

We also verify that in pixels where the concentration of both dyes is nonzero, modelling the data as if it came from a single dye yields poor fits, confirming that our approach of using the fit of the single dye model to select pixels in which only one dye is present is appropriate Appendix A.3.

In pixels where the concentration of both dyes is nonzero, we further wish to find the relative concentration of each dye. We first analyse the performance of the Laplace and Fourier transform methods when the lifetime values are known at a pixel and the concentrations are unknown. This method is fast compared to the baseline, because it does not use a numerical solver to solve a system of equations. We again obtain excellent results even for high noise in the data with the Laplace transform method, although when the underlying lifetimes of the fluorophores are similar, the model’s variance increases substantially (see Figure 4). Applying our pipeline to the experimental data provided by SVI, we find that the methods do not qualitatively differ when the lifetimes are assumed known from global estimates.

However, in our analysis of experimental data, we find that the assumption that the lifetime estimates obtained per pixel are constant across the image may be naive for our dataset (see § 7). We find that the lifetimes are normally distributed around two means. Further analysis is needed to estimate how much of this variation comes from underlying differences in lifetime and how much of it comes from model variance. When τ variation is introduced to the data at a pixel, the Laplace and Fourier models need to be adjusted to also predict the τ values. In particular, the Laplace/Fourier transform needs to be computed at multiple frequencies to set up a system of quadratic equations. We attempt to solve these resulting quadratic equations analytically, but find that they are extremely sensitive to noise in the data (Appendix B) and are therefore not suitable for our pipeline. We thus consider numerical solvers for the quadratic equation, which are slower. For most parameter combinations, we find that there is a region of frequency values for which the Fourier and Laplace methods with numerical solvers converge to the correct lifetimes and concentrations. When the lifetimes or concentrations are too similar, the methods failed to correctly predict the underlying model. However, a baseline model using least squares regression is also not accurate in this setting.

In the real dataset, allowing the lifetimes to vary far away from the lifetimes estimated globally yields overfitted solutions at pixels where the concentration of one of the dyes is low. We hypothesize that in this situation, the numerical solver can exploit variation of the corresponding lifetime to fit to the noise in the data. While setting bounds on the τ values decreases model performance as measured by mean square error, we think it may help prevent overfitting and the prediction of

nonsensical lifetimes. Such bounds can be obtained by considering the standard deviation of the obtained τ values from the single dye analysis, see § 6.3.

8.1 Recommendations and further work

We recommend using the Laplace and Fourier transform methods to obtain global estimates for the underlying lifetimes in an image, since they offer good accuracy across many samples and are much faster than models such as least squares.

We also recommend analysing the variation in the predicted lifetimes and using it to inform the assumption that lifetimes computed globally are the lifetimes at each pixel. In the case where variation of the predicted lifetimes is suspected to come from underlying biology, we recommend allowing lifetimes to vary in the model predicting concentration. While, when compared to assuming fixed lifetimes, this method is computationally more costly and works for a smaller subset of parameters, in situations where the lifetimes are actually varying, assuming they are fixed can yield nonsensical results.

Further work could focus on optimizing the numerical solvers we applied in our Fourier and Laplace transform methods. Due to time constraints, we were unable to finetune the solver, and we suspect that varying the constraints as well as the loss function of the solvers can yield increased speed and accuracy. This should begin with an analysis of computational complexity and speed of the current models. Secondly, further work could investigate adding a component in the processing pipeline that infers from the predicted lifetimes whether there is underlying lifetime variation across the image. If variation in lifetime comes from biological processes, it makes sense to predict lifetimes per pixel and use these predictions to predict concentrations. Otherwise, fixed lifetime estimates can be used to predict concentrations.

9 Acknowledgements

We would like to thank SVI for posing the question about fluorescence microscopy and kernel methods we have researched and implemented. Additionally, we would like to thank them for their hospitality and their expertise on fluorescence lifetime imaging microscopy, which helped us realize this final product.

Moreover, we would like to thank the Studygroup Mathematics with Industry (SWI) of 2025 for giving us an insight in their provisional results and notes. As their project is closely related to ours, we were able to use their progress to check the correctness and functionality of our ideas and interpretations.

In particular, we would like to thank Dr. Palina Salanevich for supervising our project. We found our discussions to not only be fruitful but also very enjoyable.

References

- [1] Rupsa Datta et al. “Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications”. English. In: *Journal of Biomedical Optics* 25.7 (July 2020), p. 71203. URL: <https://www.proquest.com/scholarly-journals/fluorescence-lifetime-imaging-microscopy/docview/2859140516/se-2>.
- [2] Rupsa Datta et al. “Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications”. In: *Journal of Biomedical Optics* 25.07 (May 2020), p. 1. ISSN: 1083-3668. DOI: 10.1117/1.jbo.25.7.071203. URL: <http://dx.doi.org/10.1117/1.JBO.25.7.071203>.
- [3] UC Davis Health. Apr. 2021. URL: <https://health.ucdavis.edu/cancer/news/headlines/new-imaging-diagnostic-tool-lights-up-tumors-during-brain-cancer-surgery/2021/04>.
- [4] Nadine A. Yehya. Dec. 2022. URL: <https://health.ucdavis.edu/news/headlines/uc-davis-health-surgeons-use-flim-imaging-to-detect-head-and-neck-cancer-during-surgery-/2022/12>.
- [5] Julien Bec et al. “Label-Free Visualization and Quantification of Biochemical Markers of Atherosclerotic Plaque Progression Using Intravascular Fluorescence Lifetime”. In: *JACC: Cardiovascular Imaging* 14.9 (Sept. 2021), pp. 1832–1842. ISSN: 1936-878X. DOI: 10.1016/j.jcmg.2020.10.004.
- [6] Dan Elson et al. “Time-domain fluorescence lifetime imaging applied to biological tissue”. In: *Photochemical & Photobiological Sciences* 3.8 (Aug. 2004), pp. 795–801. ISSN: 1474-9092. DOI: 10.1039/b316456j.
- [7] Yuansheng Sun, Richard N. Day, and Ammasi Periasamy. “Investigating protein-protein interactions in living cells using fluorescence lifetime imaging microscopy”. In: *Nature Protocols* 6.9 (Aug. 2011), pp. 1324–1340. ISSN: 1750-2799. DOI: 10.1038/nprot.2011.364.
- [8] Frederik Görlitz et al. “Mapping Molecular Function to Biological Nanostructure: Combining Structured Illumination Microscopy with Fluorescence Lifetime Imaging (SIM + FLIM)”. In: *Photonics* 4.3 (July 2017), p. 40. ISSN: 2304-6732. DOI: 10.3390/photonics4030040. URL: <http://dx.doi.org/10.3390/photonics4030040>.
- [9] Ari Gafni, Robert L. Modlin, and Ludwig Brand. “Analysis of fluorescence decay curves by means of the Laplace transformation”. In: *Biophysical journal* 15.3 (1975), pp. 263–280.

A Analysis of methodology

In this appendix, we describe in some more detail parts of our analysis of the proposed method.

A.1 Truncation and Laplace method

In § 5.1.3, we discuss the problem of truncation: since photons are observed until some time T , there might be photons emitted after time T that will not be observed. Our conclusion is that this does not pose a problem for the experimental dataset we received, since the number of photons we expect to emit after observations take place is multiple order of magnitudes smaller than 1; therefore, an ‘untruncated’ model suffices.

Here, we discuss how the model should be changed if truncation cannot be ignored. To incorporate truncation, we would estimate τ based on a model of

$$\tilde{N}(t) = \begin{cases} N(t) - N(T) = N_0 e^{-t/\tau} - N_0 e^{-T/\tau} & \text{if } t \in [0, T], \\ 0 & \text{if } t \in (T, \infty), \end{cases}$$

but this is not straightforward in general. For example, the Laplace transform of $\tilde{N}$ is

$$\begin{aligned} \mathcal{L}[\tilde{N}](s) &= \int_0^\infty e^{-st} \tilde{N}(t) dt = \int_0^T e^{-st} (N(t) - N(T)) dt \\ &= \frac{N_0}{s + 1/\tau} - \frac{N_0 e^{-T(s+1/\tau)}}{s + 1/\tau} - \frac{N_0 e^{-T/\tau}}{s} + \frac{N_0 e^{-T(s+1/\tau)}}{s}, \end{aligned}$$

which does not provide a polynomial equation for τ . As noted in [9], iterative methods may be employed to solve the obtained equation (or system of equations in the multi-dye case). However, numerical analysis shows that the difference $\mathcal{L}[N](s) - \mathcal{L}[\tilde{N}](s)$ between the Laplace transform of the non-truncated model and the Laplace transform of the truncated model is actually negligible with parameters with the order of magnitudes mentioned in § 5.1.3; this agrees with the conclusion we reached there, that the quantity $N(T)$ is negligible for our purposes. We conclude that the error introduced by using a non-truncated model for the phasor method via the Laplace transformation is not substantial.

A.2 Generation of synthetic data

Since the concentrations and mean lifetimes of the fluorophores are not known at each pixel in the experimental image under consideration, throughout our analysis we use computationally generated data with varying levels of noise to simulate observations at individual pixels and thus evaluate the models we consider.

We generate data simulating the exponential decay function

$$N(t) = N_0 e^{-t/\tau} - N_0 e^{-T/\tau},$$

where the constant term $N_0 e^{-T/\tau}$ is used to model the problem of truncation described in § 5.1.3. We further allow the specification of the number of samples to generate. To simulate variations in the data due to random processes in the sample, we need to be able to generate decay curves with noise synthetically. Following Gafni et al., we add Gaussian noise to the decay function at each time point [9, p. 272]. More specifically, Gaussian noise is added at each time point t to obtain

$$\tilde{N}(t) = N(t) + N_{\text{scale}} \cdot \epsilon_t$$

where $\epsilon_t \sim \mathcal{N}(\mu: 0, \sigma^2: N_0 e^{-t/\tau})$ and N_{scale} is a scaling parameter. Since we are additionally truncating, we may obtain negative values in the tail. Since this cannot happen in practice, we set these values to zero. To generate data that decays as a multi-exponential, synthetic data from single exponential decay is generated for multiple pairs (τ, N_0) and summed.

640 A.3 Accuracy of identifying single-dye pixels

Our approach to estimating the lifetimes rests on the assumption that the single-dye model is indeed accurate for pixels where only one dye is present, and *moreover*, that it is inaccurate for pixels where two different dyes are present with a significant concentration (see § 6.2). In this section, we assess the accuracy of this latter assumption.

645 To that end, we generate synthetic data for a two-dye model $N(t) = N_0 e^{-t/\tau_0} + N_1 e^{-t/\tau_1}$ for various values of the parameters N_0 , N_1 , τ_0 and τ_1 . We use the least squares method to fit a single-dye model $N(t) = N_0 e^{-t/\tau}$ to the generated data, and compute the mean squared error of this fit. We expect the error of fitting a single-dye model to data from two dyes with different lifetimes to be small when the concentration of one of the dyes is around 1, and large when the concentrations
650 are closer to 1/2. Moreover, when the difference between the lifetimes is smaller, we expect the fits to be better in general (if the lifetimes are equal, we cannot distinguish the dyes): even when the dyes have the same concentration, we expect a single-dye model will be more accurate.

In Figure 23, we plot the mean square error of the single-dye model's fit to synthetic two-dye data. We generated data fifty times and plotted the mean and standard deviation of the errors over
655 the fifty experiments. In Figure 23a, we vary the concentration of the dyes and put the concentration of dye 0 on the x -axis; we do this for multiple values of the lifetimes τ_0 and τ_1 . We see that the error is smallest for data that is actually from one dye, when the concentration of one of the dyes is around 1, and largest for data that is from two dyes that are present with the same concentration. Moreover, we see that the fits are overall better when the values of lifetimes are closer to each other,
660 as in the middle plot.

In Figure 23b, we vary the lifetimes τ_0 and τ_1 and put the value of τ_0 on the x -axis; we do this for multiple values of N_0 and N_1 , although in each case with $N_0 + N_1 = 300$. We see that the error is smaller when the values of the lifetimes are closer, but the minimum seems to be just above the value of τ_0 where $\tau_1 = \tau_0$. We also see that the fits are overall better when the values of N_0 and N_1
665 are further apart from each other, as in the left-hand plot.

A.4 Choosing a truncation parameter

Recall that the function that generates noisy truncated exponential data is given by

$$N(t) = N_0 e^{-t/\tau} - N_0 e^{-T/\tau}, \quad (\text{A.1})$$

where the constant term $N_0 e^{-T/\tau}$ is used to model the problem of truncation described in § 5.1.3.

670 This means we need to pick a truncation time T . We can investigate the effect of the truncation parameter on the accuracy of our model by plotting the cutoff value against the predicted values for N_0 and N_1 . This can be seen in Figure 24. As the plot shows, convergence to the real values of N_0 and N_1 occurs around $T \approx 200$. This implies that any value $T \geq 200$ would suffice as the truncation parameter. Since our experimental data is cut off at $T = 300$, we use $T = 250$ as our
675 truncation time throughout the report. This truncation time is used in our double-dye model as well, since we synthesize noisy double-dye data simply by generating two sets of noisy-double dye decays and summing them.

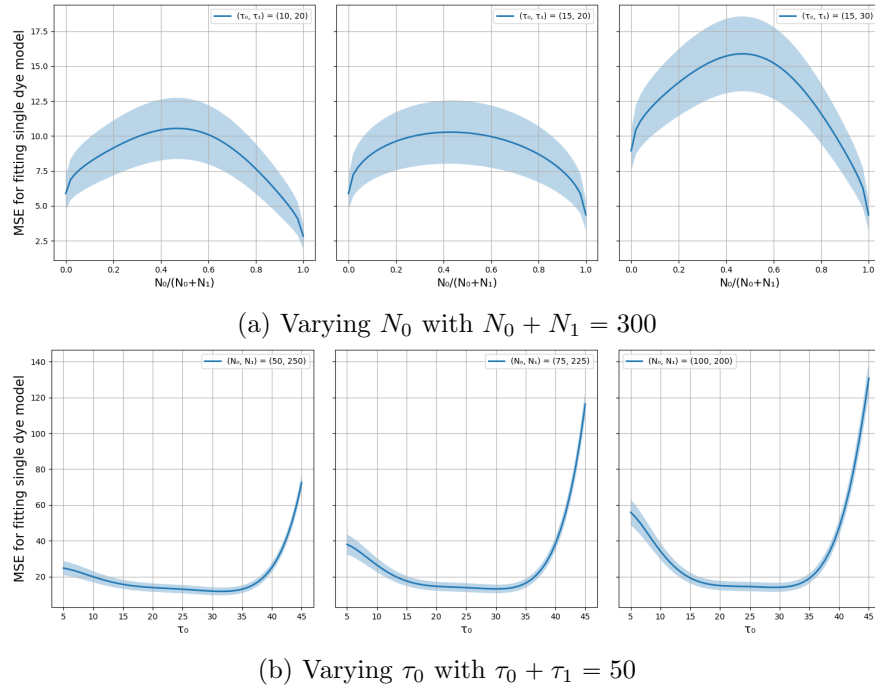


Figure 23: **Fitting a single-dye model to synthetic two-dye data.** Truncated noisy exponential decay data with truncation cutoff $T = 250$, noise coefficient $N_{\text{scale}} = 1/2$. Solid lines are mean values across fifty trials. Shaded areas show standard deviation across trials.

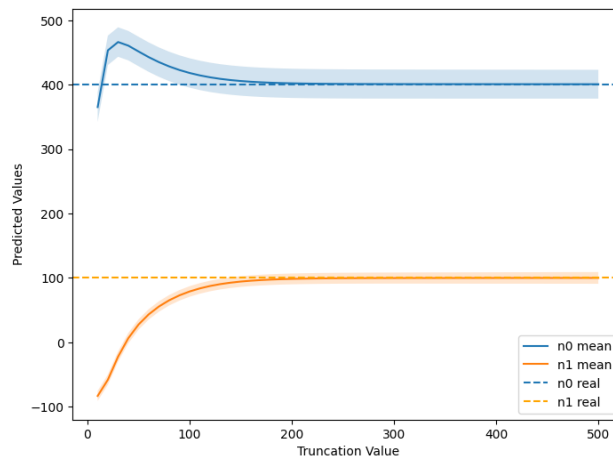


Figure 24: The effect of increasing the truncation parameter on Laplace model fit. $N_0 = 400$, $N_1 = 100$.

A.5 Parameter degeneracy

In practice, the problem described at the end of § 6.4.2.1 will be hard to solve because of the following reason: one can pick λ_0 , λ_1 , M_0 and M_1 that are very different from the correct values of τ_0 , τ_1 , N_0 and N_1 such that

$$M(t) = M_0 e^{-t/\lambda_0} + M_1 e^{-t/\lambda_1}$$

is a very close approximation of $N(t)$. The table below shows the mean-squared error between $M(t)$ and $N(t)$ for various values of $M_0, M_1, \lambda_1, \lambda_2, N_0, N_1, \tau_0, \tau_1$. The MSE is computed on 500 points with a truncation of 250 in (10^{-1} ns). In Figure 26, we plot specific instances of $N(t)$ and $M(t)$ for values from Figure 25 to show that these are almost the same. Note that in the case of $N_0 = 100$, $N_1 = 400$, $\tau_0 = 5$, $\tau_1 = 45$, the MSE between $N(t)$ and $M(t)$ increases more quickly when $M_0, M_1, \lambda_0, \lambda_1$ differ from N_0, N_1, τ_0, τ_1 compared to the other two cases. In this case, the problem is easier to solve. This is reflected in the graphs that show the performance of our methods.

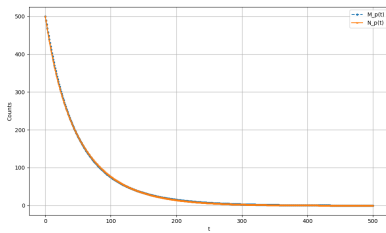
M_0	M_1	λ_0	λ_1	MSE	M_0	M_1	λ_0	λ_1	MSE	M_0	M_1	λ_0	λ_1	MSE
122	378	11	31	0.22561764	285	215	28	9	0.3565755	102	398	5	45	0.25062675
132	368	12	31	0.27981311	285	215	11	32	0.3727549	91	409	4	44	0.91448818
153	347	13	32	0.43082911	275	225	29	9	0.49092963	112	388	6	46	1.17993935
81	419	9	29	0.43424331	295	205	11	33	0.92158532	91	409	5	44	1.7663317
154	346	13	32	0.43978809	316	184	12	34	1.37522629	112	388	7	46	2.41667386
143	357	12	32	0.68744968	307	193	27	8	1.43956638	122	378	7	47	3.16853817
163	337	14	32	0.69130668	317	183	12	34	1.46164884	91	409	4	45	3.34873225
164	336	14	32	0.69167416	326	174	12	36	1.5249583	81	419	4	43	4.140005
142	358	12	32	0.77290251	306	194	27	8	1.59857432	82	418	3	43	4.19028464
215	285	16	34	1.3823572	307	193	12	33	1.67451498	81	419	3	43	4.2855166

(a) $N_0 = 100$, $N_1 = 400$, $\tau_0 = 10$, $\tau_1 = 30$

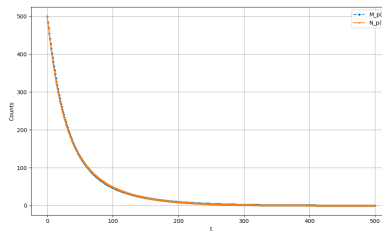
(b) $N_0 = 250$, $N_1 = 250$, $\tau_0 = 10$, $\tau_1 = 30$

(c) $N_0 = 100$, $N_1 = 400$, $\tau_0 = 5$, $\tau_1 = 45$

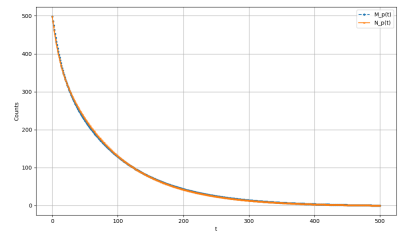
Figure 25: Difference between $M(t)$ and $N(t)$ for different values



(a) $N_0 = 100$, $N_1 = 400$, $\tau_0 = 5$, $\tau_1 = 45$, $M_0 = 215$, $M_1 = 285$, $\lambda_0 = 16$, $\lambda_1 = 34$



(b) $N_0 = 250$, $N_1 = 250$, $\tau_0 = 10$, $\tau_1 = 30$, $M_0 = 307$, $M_1 = 193$, $\lambda_0 = 12$, $\lambda_1 = 33$



(c) $N_0 = 100$, $N_1 = 400$, $\tau_0 = 5$, $\tau_1 = 45$, $M_0 = 122$, $M_1 = 378$, $\lambda_0 = 7$, $\lambda_1 = 47$

Figure 26: Examples of $M(t)$ and $N(t)$

B Analytic solution

It is suggested to solve the fluorescent lifetimes analytically using several Laplace transforms computed from different frequencies, as is done in [9, Appendix A(d)]. In this appendix, we consider

the efficacy of this approach for the two dye case. We symbolically solved the system of equations below using a computer algebra system. After a cursory analysis, we found the analytic solution to be quite sensitive to noise. We therefore decided not to investigate this method any further beyond this section.

For the two dye case, we consider the system of equations:

$$\begin{cases} L_0 &= \frac{N_0}{s_0 + \frac{1}{\tau_0}} + \frac{N - N_0}{s_0 + \frac{1}{\tau_1}} \\ L_1 &= \frac{N_0}{s_1 + \frac{1}{\tau_0}} + \frac{N - N_0}{s_1 + \frac{1}{\tau_1}} \\ L_2 &= \frac{N_0}{s_2 + \frac{1}{\tau_0}} + \frac{N - N_0}{s_2 + \frac{1}{\tau_1}}. \end{cases} \quad (\text{B.1})$$

as seen in § 6.4.2.1. Recall that $s_0, s_1, s_2, N, L_0, L_1$ and L_2 are known to us and that we want to solve for τ_0, τ_1, N_0 and N_1 . Firstly, note that since $N_0 = N - N_1$ we indeed only need three equations to solve for all the unknowns. Expressing a general closed solution for these equations is tedious. Therefore, we resorted to the **Sagemath** library in Python to find one for us. In doing this, we noticed that expressing the closed form solution for τ_0, τ_1 and N_0 in terms of all the variables ($s_0, s_1, s_2, N, L_0, L_1$ and L_2) cannot be done by this computer algebra system in a reasonable amount of time. To circumvent this issue, we employ a useful trick, namely, we additionally require $s_1 = 2s_0, s_2 = 3s_1$, so that only s_0 remains free.

It turns out that under these conditions, the **Sagemath** library can easily symbolically solve this system of equations. In our case, we obtained the solution:

$$\begin{aligned} N_0 &= \frac{1}{2} \left(N + \frac{X_4}{\sqrt{N^2 X_1 + 2N s_0 X_2 + s_0^2 X_3}} \right) \\ \tau_0 &= \frac{1 - V_1 + \sqrt{N^2 X_1 + V_3 + s_0^2 X_3}}{2 V_2} \\ \tau_1 &= \frac{1 - V_1 - \sqrt{N^2 X_1 + V_3 + s_0^2 X_3}}{2 V_2}, \text{ where} \\ V_1 &= (L_0 - 2L_1 + L_2)N + (3L_0 L_1 - (8L_0 - 5L_1)L_2)s_0 \\ V_2 &= s_0(N(L_0 - 4L_1 + 3L_2) + 2s_0(L_0 L_1 - 3(L_0 - L_1)L_2)) \\ V_3 &= 2N s_0(L_0^2 L_1 + 2L_0 L_1^2 + (4L_0 - L_1)L_2^2 - 2(4L_0^2 L_1 + L_0 L_1^2)L_2) \\ X_1 &= (L_0 - 2L_1 + L_2)^2 \\ X_2 &= (L_0^2 L_1 + 2L_0 L_1^2 + L_2^2(4L_0 - L_1) - 2L_2(2L_0^2 + L_1^2)) \\ X_3 &= L_0^2 L_1^2 + L_2^2(4L_0 - L_1)^2 - 2L_0 L_1 L_2(4L_0 + L_1) \\ X_4 &= 4L_0 L_1 L_2 s_0^2 + N^3(L_0 - 2L_1 + L_2)N^2 - 3N s_0 L_1(L_0 - L_2) \end{aligned}$$

In reality **Sagemath** predicts two solutions, but this naturally arises from a symmetry in the system of equations, i.e., we may interchange τ_0 and τ_1 if we also interchange N_0 and $N - N_0$.

Evaluating this method's sensitivity to Gaussian noise By construction, these equations give us τ_0, τ_1 and N_0 if we input the corresponding L_0, L_1, L_2 and s_0 . By calculating the Laplace transforms of our experimental data at $s_0, 2s_0$ and $3s_0$, we obtain approximates $\widehat{L}_0, \widehat{L}_1$ and $\widehat{L}_2$ of the respective L_i 's. Thus, it is of interest to us to consider how these predicted values for τ_0, τ_1 and N_0 change if we artificially add small errors to L_0, L_1 and L_2 . In this way, we obtain a proxy for the sensitivity of this method to noise. For the sake of simplicity, we model the $\widehat{L}_i$'s by sampling

from a Gaussian distribution centred around the true L_i 's with standard deviation $\sigma \times |L_i|$ for some (very small) $\sigma > 0$. Then we solve the system of equations using these $\hat{L}_i$ along with N and s_0 to obtain τ'_0 , τ'_1 and N'_0 . When doing this, we have to decide which of the two analytic solutions we consider. We calculate the relative error in both cases and simply consider the solution where the average relative error for τ_0 , τ_1 , N_0 and $N - N_0$ is the smallest. That is, if

$$\frac{|\tau'_0 - \tau_0|}{\tau_0} + \frac{|\tau'_1 - \tau_1|}{\tau_1} + \frac{|N'_0 - N_0|}{N_0} + \frac{|(N - N'_0) - (N - N_0)|}{N - N_0} > \frac{|\tau'_0 - \tau_0|}{\tau_0} + \frac{|\tau'_1 - \tau_1|}{\tau_1} + \frac{|(N - N'_0) - N_0|}{N_0} + \frac{|N'_0 - (N - N_0)|}{N - N_0} \quad (\text{B.2})$$

we switch the values of τ'_0 and τ'_1 and instead of N'_0 we use $N - N'_0$. By repeatedly computing these noise terms and then calculating the expected solutions, we thus get some estimate of how badly the analytic solution performs on noisy data.

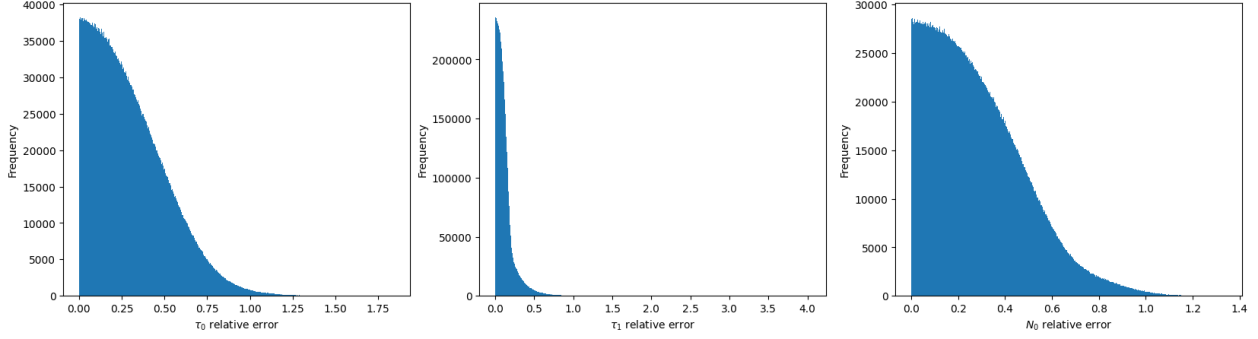


Figure 27: Relative error over 10^7 iterations of the predictions for τ_0 , τ_1 and N_0 , for $\tau_0 = 10$, $\tau_1 = 40$, $N_0 = 200$, $N = 500$, $s_0 = 0.01$ and $\sigma = \frac{1}{1000}$.

It turns out that even a relatively small amount of noise gives a wide spread of predictions for τ_0 , τ_1 and N_0 . In Figure 27 we see what the spread looks like for τ_0 , τ_1 and N_0 , for $\tau_0 = 10$, $\tau_1 = 40$, $N_0 = 200$, $N = 500$, $s_0 = 0.01$ and $\sigma = \frac{1}{1000}$. Even with this small amount of noise, we get quite a spread. Another thing to note is that the three graphs in Figure 27 look like the right side of a normal distribution. We can also consider how this standard deviation changes when we let the parameter σ vary. In Figure 28 we do this for $\frac{1}{1000} \leq \sigma \leq \frac{1}{100}$ and $\tau_0 = 10$, $\tau_1 = 40$, $N_0 = 200$, $N = 500$ and $s_0 = 0.01$. We note that the method is extremely unstable for predicting the τ_1 in this case. For τ_0 and N_0 we get the expected result that the standard deviation increases as σ increases.

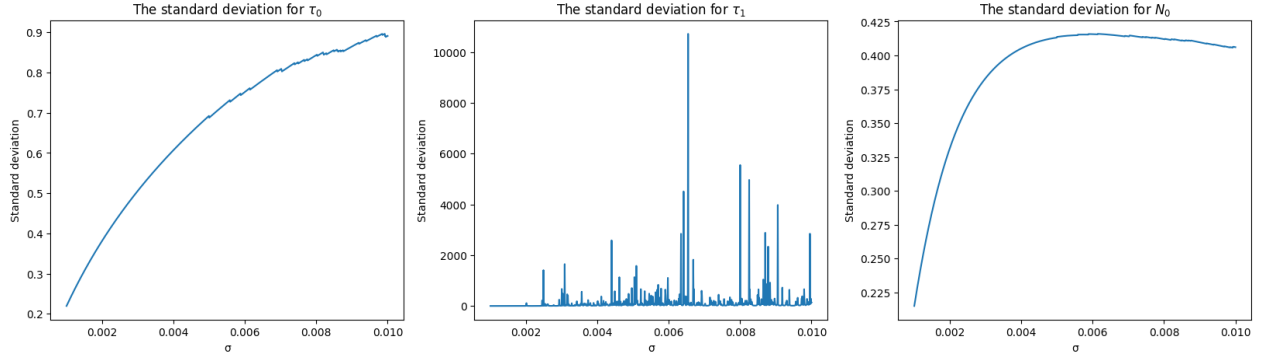


Figure 28: Standard deviation of the relative error of the predicted τ_0 , τ_1 and N_0 for 10^3 values of σ between $\frac{1}{100}$ and $\frac{1}{1000}$. For each σ , we sample 10^4 times to get the standard deviation. The underlying parameters are $\tau_0 = 10$, $\tau_1 = 40$, $N_0 = 200$, $N = 500$ and $s_0 = 0.01$.