



**HAL**  
open science

# An explicit Estimated Baseline Model for Robust Estimation of Fluorophores using Multiple-Wavelength Excitation Fluorescence Spectroscopy

Arthur Gautheron, M. Sdika, M. Hébert, B. Montcel

► **To cite this version:**

Arthur Gautheron, M. Sdika, M. Hébert, B. Montcel. An explicit Estimated Baseline Model for Robust Estimation of Fluorophores using Multiple-Wavelength Excitation Fluorescence Spectroscopy. *IEEE Transactions on Biomedical Engineering*, 2023, 71 (1), pp.295-306. 10.1109/TBME.2023.3299689 . hal-04188757

**HAL Id: hal-04188757**

**<https://hal.science/hal-04188757>**

Submitted on 7 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An explicit Estimated Baseline Model for Robust Estimation of Fluorophores using Multiple-Wavelength Excitation Fluorescence Spectroscopy

A. Gautheron, M. Sdika, M. Hébert, B. Montcel

**Abstract**—Spectroscopy is a popular technique for identifying and quantifying fluorophores in fluorescent materials. However, quantifying the fluorophore of interest can be challenging when the material also contains other fluorophores (baseline), particularly if the emission spectrum of the baseline is not well-defined and overlaps with that of the fluorophore of interest. In this work, we propose a method that is free from any prior assumptions about the baseline by utilizing fluorescence signals at multiple excitation wavelengths. Despite the nonlinearity of the model, a closed-form expression of the least squares estimator is also derived.

To evaluate our method, we consider the practical case of estimating the contributions of two forms of protoporphyrin IX (PpIX) in a fluorescence signal. This fluorophore of interest is commonly utilized in neuro-oncology operating rooms to distinguish the boundary between healthy and tumor tissue in a type of brain tumor known as glioma. Using a digital phantom calibrated with clinical and experimental data, we demonstrate that our method is more robust than current state-of-the-art methods for classifying pathological status, particularly when applied to images of simulated clinical gliomas. To account for the high variability in the baseline, we are examining various scenarios and their corresponding outcomes. In particular, it maintains the ability to distinguish between healthy and tumor tissue with an accuracy of up to 87%, while the ability of existing methods drops near 0%.

**Index Terms**—Estimation Methods, Source Separation, Non-Linear Spectral Unmixing, Fluorescence optical imaging, Fluorescence spectroscopy, Optical Signal, PpIX, Glioma, Neurosurgery

This work was funded by LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007), operated by the French National Research Agency (ANR); Cance'ropôle Lyon Auvergne Rhône Alpes (CLARA) within the program "OncoStarter", Infrastructures d'Avenir en Biologie Santé (ANR-11-INBS-000), within the program "Investissements d'Avenir" operated by the French National Research Agency (ANR) and France Life Imaging (ANR-11-INBS-0006).

A. Gautheron, M. Sdika, and B. Montcel are with the Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69621, LYON, France (e-mail: gautheron@creatis.insa-lyon.fr, sdika@creatis.insa-lyon.fr, montcel@creatis.insa-lyon.fr).

M. Hébert is with the Université Jean Monnet Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France (e-mail: mathieu.hebert@institutoptique.fr).

## I. INTRODUCTION

**F**LUORESCENCE is a non-linear optical process in which a chemical compound absorbs light at a specific wavelength and subsequently emits light at different wavelengths. In many materials including biological tissues, fluorescence involves a variety of fluorescent molecules. Some of these compounds may serve as can be relevant biomarkers in biomedical applications, by analyzing their time response after excitation or emission spectrum. Studying the temporal fluorescence response of biological tissues enables the estimation of fluorophore concentrations fitting decay models [1], or the estimation of neural activation spikes [2]. Fluorescence spectroscopy analyzes the spectral signal emitted by fluorophores. This technique offers several advantages for characterizing molecular reactions or interactions. For example, the spectral fluorescence response of biological samples can enable the detection of to detect breast cancer [3] or residual tumors in the vicinity of the surgical site [4]. However, as in [5] and [1], and in most practical cases, it is necessary to extract the temporal or spectral response of the fluorophores of interest from a global signal may include the fluorescence signal of other compounds. The base component signals can either be assumed to be known [6], assumed to have a specific signal shape (such as a Gaussian) [7], [8], or estimated through blind unmixing routines, such as the nonnegative matrix factorization algorithm [9]–[12]. In other words, a measured signal is of a linear combination of variable signals from fluorophores that are located near the measurement probe. We are interested in characterizing the specific contribution of a fluorophore of interest. Blind unmixing algorithms are generic and can be applied to various domains [1], [10], [13], [14] However, in our case, because they have limitations because they cannot accurately isolate contributions that are specific to a particular fluorophore of interest.

The sensitivity of fluorescence spectroscopy remains mainly limited by the presence of other fluorophores. The variability of the fluorescent response of these other fluorophores depends a lot on external factors [8], [15]–[23]. The high variability of these fluorophores can lead to significant crosstalk with the fluorophore of interest. Omitting a fluorophore, poorly modeling it, or using an incorrect reference spectrum affects

the detection of specific fluorescence signals from the fluorophores of interest. Crosstalk events cause an overestimation of the amplitude of the target fluorophores, leading to an overestimation of the contributions of the fluorophores of interest. This leads to a drop in specificity when interest biomarker contributions are utilized in a classification pipeline [19], [22].

To eliminate crosstalk, the general approach is to include a baseline signal in the model that encompasses all factors unrelated to the fluorophores of interest. Existing approaches are effective when the emission spectral band of the baseline is far from the one of the fluorophores of interest. In these cases, the analytical *a priori* on the baseline is an exponential shape function [15], [24]. When the emission spectral band of the baseline is close to or within one of the fluorophores of interest, the existing methods remain more limited and very often expert-dependent. In these cases, the baseline is modeled by a Gaussian function [8], or functions so-called expert, which contains a finite number of fluorophores of interest whose fluorescence signal shape is assumed known [22], [23], [25]. All these models have strong limitations in the face of the variability of baseline fluorophores or when some baseline fluorophores emit too close to the wavelength range of the fluorophore of interest. Despite these limitations, fluorescence spectroscopy has shown its relevance for several applications in neuro-oncology [26]–[31], food analysis [32] or biological studies at cell level [33].

Another approach suggested in [10] is to use the blind-source separation technique as a spectral unmixing routine. The main improvement of this technique is to estimate the baseline instead of modeling it. But this technique has existing limitations first as it requires an expert *a priori* to correctly tune hyperparameters like the number of sources. In addition, this technique is highly time-consuming as it relies on an iterative solver.

We propose in this study a novel approach to estimate the interest biomarkers' contributions. To relieve hyperparameters determination and the *a priori* on the spectral shape of the emission by the baseline, our method uses several excitation wavelengths, which is in contrast with the single excitation used in literature so far. This additional measurement allows us to estimate the baseline that must be fixed in a single measurement problem. However, considering multiple excitation wavelengths results in a non-linear estimation problem that in general does not have a closed-form solution. For the specific case of two distinct wavelengths, we were able to find the conditions for the problem to have a unique solution and to derive an analytic expression of the solution for all the variables. This makes our method computationally efficient, free from parameter tuning and cumbersome iterative estimation procedures.

To evaluate our method, we consider the practical case of the estimation of the contributions of two forms of Protoporphyrin IX (PpIX) in a fluorescence signal. This fluorophore of interest is widely used in the operating theatre in neuro-oncology to characterize the healthy/tumoral boundary of a brain tumor named glioma [34]–[36]. Glioma is the most frequent and aggressive primary brain tumor whose main therapy is com-

plete tumor removal surgery [37] followed by Stupp protocol [38]. Thus, the neurosurgeon uses this information on PpIX contributions to remove or not the measured tissue.

Various 5-aminolevulinic acid (5-ALA) induced PpIX fluorescence spectroscopy methods have been proposed to overcome sensitivity issues of fluorescence microscopy applied to low-density infiltrative parts of High-Grade Glioma (HGG) [25], [39] or to Low-Grade Glioma (LGG) [40]. However, the high levels of crosstalk due to the presence of other non-5-ALA-induced fluorophores largely impair specificity, because healthy samples are classified as tumoral. In section II, we first describe the new estimation method using multiple excitations and derive the closed-form solution of the associated estimation problem. In section III, we describe the numerical phantom calibrated on real HGG and LGG data [8], [41]. Then, numerical results show that our method improves the robustness of the estimation of PpIX contributions and thus the specificity of the classification. Especially in cases where existing methods fail, e.g. when other fluorophores emit in the same spectral domain as PpIX. Finally, we compare the performance of our method with other methods issued from the literature in terms of their ability to maintain significant specificity in detecting the healthy/margins boundary on synthetic images of LGG and HGG calibrated on a real experimental system [8]. Our method keeps a specificity of 87% while other models' specificity drops to 0% at best.

## II. METHODS

### A. Physical Model

The spectrally resolved signal of the fluorescence emission used in this study relies on multiple excitation wavelengths. The spectrally resolved signal of the fluorescence measured signal  $\mathbf{m}$  can be written as:

$$\mathbf{m}(\lambda, \lambda_e) = \sum_{i=1}^n \alpha_i \eta_i(\lambda_e) \mathbf{S}_i(\lambda) + \gamma(\lambda_e) \mathbf{m}(\lambda), \quad (1)$$

where  $\lambda_e$  is the fluorescence excitation wavelength,  $\mathbf{S}_i(\lambda) \in \mathbb{R}^p$  is the normalized emission spectrum of a fluorophore of interest,  $\eta_i(\lambda_e) \in \mathbb{R}$  the quantum yield at excitation wavelength  $\lambda_e$ ,  $\alpha_i \in \mathbb{R}$  the contribution of the  $i^{\text{th}}$  fluorophore of interest.  $\mathbf{b}(\lambda) \in \mathbb{R}^p$ , the baseline represents other endogenous fluorophore emission spectra whose amplitude concerning the excitation wavelength is modeled by  $\gamma(\lambda_e) \in \mathbb{R}$ . Each emission spectrum  $\mathbf{S}_i(\lambda)$  has been normalized by its integral over the entire wavelength range.

$\mathbf{S}_i(\lambda)$  refers to the signal of the fluorophores of interest and  $\eta_i$  corresponds to the amplitude variation due to the different excitation wavelengths.  $\gamma$  represents the baseline amplitude variation due to the different excitation wavelengths, and  $\mathbf{b}(\lambda)$  is the baseline i.e. the residual signal of all but the fluorophores of interest. Fluorophore estimation is the problem of estimating  $\alpha_i$ . In this problem,  $\mathbf{S}_i(\lambda)$  and  $\eta_i$  are supposed to be known,  $\alpha_i$ ,  $\gamma$  and  $\mathbf{b}(\lambda)$  are to be estimated.

One emphasizes that the main assumption in equation 1 and hereinafter is that the fluorescence baseline keeps the same emission spectral shape whatever the excitation wavelength is. This assumption is based on the fact that the emission

spectra are generally independent of the excitation wavelength [42]. Please note that in the presence of  $k$  distinct excitation wavelengths, the condition  $n + k - 1 \leq (k - 1)p$  is required for the problem to be either overdetermined or perfectly determined. We assume hereinafter that the previous condition is fulfilled. Moreover, it is assumed that  $p \gg k$ , e.g.  $p = 819$  and  $k = 2$ , as in Part III.

For the sake of concreteness, we consider only the case of  $n = 2$  fluorophores of interest and  $k = 2$  excitations wavelengths:

$$\begin{aligned} \mathbf{m} &= \alpha_1 \eta_1 \mathbf{S}_1 + \alpha_2 \eta_2 \mathbf{S}_2 + \mathbf{b} \\ \mathbf{m}' &= \alpha_1 \eta_1' \mathbf{S}_1 + \alpha_2 \eta_2' \mathbf{S}_2 + \gamma \mathbf{b}. \end{aligned} \quad (2)$$

Defining contributions and the estimated measurements vector as  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T \in \mathbb{R}^2$  and  $\hat{\mathbf{m}} = (\mathbf{m}^T, \mathbf{m}'^T)^T \in \mathbb{R}^{2p}$ , equation 2 can be written in a matrix form as

$$\begin{aligned} \hat{\mathbf{m}} &= \begin{pmatrix} S & I \\ SE & \gamma I \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{pmatrix} \\ &= M_\gamma \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{pmatrix} \end{aligned} \quad (3)$$

where  $S = (\eta_1 \mathbf{S}_1 \quad \eta_2 \mathbf{S}_2) \in \mathbb{R}^{p \times 2}$ ,  $E = \text{diag}(\rho_1, \rho_2) \in \mathbb{R}^{2 \times 2}$  and  $\rho_i = \frac{\eta_i}{\eta_i'} \in \mathbb{R}$ .

## B. Estimation Model

1) *Estimated Baseline Model (EB)*: Our estimator of the fluorophore contributions, called the Estimated Baseline model (EB), is the least square solution of the equation 3:

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \left\| \hat{\mathbf{m}} - M_\gamma \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{pmatrix} \right\|_2^2 \quad (4)$$

where  $\hat{\mathbf{m}} = (\mathbf{m}^T, \mathbf{m}'^T)^T \in \mathbb{R}^{2p \times 1}$  is the experimental measurements vector. In this optimization problem,  $\boldsymbol{\alpha}$ ,  $\mathbf{b}$  and  $\gamma$  are the variables to estimate and all other parameters are given. We stress that, as opposed to the state-of-the-art, the Estimated Baseline model is considered as a semi-blind method by assuming that the spectral shape of the baseline, which is estimated directly from the data, is the same at two excitation wavelengths.

2) *Analytic Resolution of Estimated Baseline Model*: The interaction between  $\gamma$  and  $\mathbf{b}$  make the Estimated Baseline Model non-linear. However, we will show in this section that it is still possible to find a closed-form expression for  $\boldsymbol{\alpha}$ . The problem is treated in two steps: the least-square solution  $(\boldsymbol{\alpha}_\gamma, \mathbf{b}_\gamma)$  for a given  $\gamma$  is first derived and then the optimal  $\gamma$  is found:

$$\begin{aligned} \min_{\boldsymbol{\alpha}, \gamma, \mathbf{b}} \left\| \hat{\mathbf{m}} - M_\gamma \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{pmatrix} \right\|^2 &= \min_{\gamma} \min_{\boldsymbol{\alpha}, \mathbf{b}} \left\| \hat{\mathbf{m}} - M_\gamma \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{b} \end{pmatrix} \right\|^2 \\ &= \min_{\gamma} \left\| \hat{\mathbf{m}} - M_\gamma \begin{pmatrix} \boldsymbol{\alpha}_\gamma \\ \mathbf{b}_\gamma \end{pmatrix} \right\|^2. \end{aligned} \quad (5)$$

For a fixed  $\gamma$ , matrix  $M_\gamma$  is constant and the problem is linear. The optimal  $(\boldsymbol{\alpha}, \mathbf{b})$  is given by  $(\boldsymbol{\alpha}_\gamma, \mathbf{b}_\gamma) = M_\gamma^+ \hat{\mathbf{m}} = (M_\gamma^T M_\gamma)^{-1} M_\gamma^T \hat{\mathbf{m}}$  where  $M_\gamma^+$  is the Moore Penrose pseudo inverse of  $M_\gamma$ .

An expression for  $M_\gamma^+$  can be obtained using a formula for the inverse of block matrices, equation 2.3 of [43].  $M_\gamma^T M_\gamma$  is given by:

$$M_\gamma^T M_\gamma = \begin{pmatrix} S^T S + (SE)^T SE & (I + \gamma E) S^T \\ S(I + \gamma E) & (1 + \gamma^2) I \end{pmatrix} \quad (6)$$

and if  $E_g = \gamma I - E$  and  $\bar{S} = (S^T S)^{-1}$ , the Schur complement  $S_u$  of  $M_\gamma^T M_\gamma$  is:

$$\begin{aligned} S_u &= \bar{S}^{-1} + E^T \bar{S}^{-1} E - (I + \gamma E) S^T \frac{I}{1 + \gamma^2} S (I + \gamma E) \\ &= \frac{E_g \bar{S}^{-1} E_g}{1 + \gamma^2}. \end{aligned} \quad (7)$$

Defining  $E_{gg} = (I + \gamma E) E_g^{-1}$  and  $\bar{S} = (S^T S)^{-1}$ ,  $(M_\gamma^T M_\gamma)^{-1}$  has the following expression:

$$(M_\gamma^T M_\gamma)^{-1} = \begin{pmatrix} (1 + \gamma^2) E_g^{-1} \bar{S} E_g^{-1} & -E_g^{-1} \bar{S} E_{gg} S^T \\ -S E_{gg} \bar{S} E_g^{-1} & \frac{(I + S E_{gg} \bar{S} E_{gg} S^T)}{1 + \gamma^2} \end{pmatrix} \quad (8)$$

and finally, the pseudo inverse  $M^+$  is given by:

$$M_\gamma^+ = \begin{pmatrix} \gamma E_g^{-1} S^+ & -E_g^{-1} S^+ \\ \frac{I - \gamma \bar{S} E_{gg} S^+}{1 + \gamma^2} & \frac{\gamma I + S E_{gg} S^+}{1 + \gamma^2} \end{pmatrix} \quad (9)$$

where  $S^+ = \bar{S} S^T = (S^T S)^{-1} S^T$ .

Note that this expression is valid only when  $S$  is full rank and  $E_g$  is invertible, i.e. when  $\gamma \notin \{\rho_1, \rho_2\}$ .

These two conditions are in agreement with the physics of the problem:  $S$  is full-ranked when the fluorophores of interest are not correlated, invertibility of  $E_g$  means that the baseline and the fluorophores of interest are not correlated. The two excitation wavelengths should be chosen to ensure that  $\gamma$  at these wavelengths is different from the ratios of the fluorophores of interest. From a signal processing point of view, this condition creates specific hot points for  $\gamma$  where  $\boldsymbol{\alpha}$  cannot be estimated. Noises in the measurements  $\mathbf{m}$  and  $\mathbf{m}'$  enlarges these hot points into instability regions. It is assumed hereinafter that these conditions are satisfied.

3) *Analytical estimation of the parameters*: Using expression 9 for  $M_\gamma^+$ , the least-square solution for  $\boldsymbol{\alpha}$  and  $\mathbf{b}$  for a given  $\gamma$  are given by:

$$\boldsymbol{\alpha}_\gamma = E_g^{-1} S^+ (\gamma \mathbf{m} - \mathbf{m}') \quad (10)$$

$$\begin{aligned} \mathbf{b}_\gamma &= \frac{\mathbf{m} + \gamma \mathbf{m}'}{1 + \gamma^2} + \frac{S E_{gg} S^+}{1 + \gamma^2} (\mathbf{m}' - \gamma \mathbf{m}) \\ &= \frac{1}{1 + \gamma^2} (\mathbf{m} + \gamma \mathbf{m}' - S E_{gg} E_g \boldsymbol{\alpha}) \end{aligned} \quad (11)$$

4) *Residual calculation and  $\gamma$  estimation*: Equation 5 can now be written:

$$\min_{\gamma} \phi(\gamma) = \mathbf{r}^T \mathbf{r} + \mathbf{r}'^T \mathbf{r}' \quad (12)$$

where

$$\begin{pmatrix} \mathbf{r} \\ \mathbf{r}' \end{pmatrix} = M_\gamma \begin{pmatrix} \boldsymbol{\alpha}_\gamma \\ \mathbf{b}_\gamma \end{pmatrix} - \begin{pmatrix} \mathbf{m} \\ \mathbf{m}' \end{pmatrix} = \begin{pmatrix} S \boldsymbol{\alpha}_\gamma + \mathbf{b}_\gamma - \mathbf{m} \\ S E \boldsymbol{\alpha}_\gamma + \gamma \mathbf{b}_\gamma - \mathbf{m}' \end{pmatrix}.$$

Using equations 10 and 11, one obtains:

$$\begin{aligned}
 (1 + \gamma^2)\mathbf{r} &= (1 + \gamma^2)S\boldsymbol{\alpha}_\gamma + \mathbf{m} \\
 &\quad + \gamma\mathbf{m}' - S(I + \gamma E)\boldsymbol{\alpha}_\gamma - (1 + \gamma^2)\mathbf{m} \\
 &= \gamma SE_g\boldsymbol{\alpha}_\gamma - \gamma(\gamma\mathbf{m} - \mathbf{m}') \\
 &= \gamma SE_g E_g^{-1} S^+ (\gamma\mathbf{m} - \mathbf{m}') - \gamma(\gamma\mathbf{m} - \mathbf{m}') \\
 &= \gamma(SS^+ - I)(\gamma\mathbf{m} - \mathbf{m}'), \tag{13}
 \end{aligned}$$

and

$$\begin{aligned}
 (1 + \gamma^2)\mathbf{r}' &= (I - SS^+)(\gamma\mathbf{m} - \mathbf{m}') \tag{14} \\
 -\gamma(1 + \gamma^2)\mathbf{r}' &= (1 + \gamma^2)\mathbf{r},
 \end{aligned}$$

from which the following expression of  $\phi$  can be deduced:

$$\phi(\gamma) = \frac{\gamma^2 \|\mathbf{q}\|^2 - 2\gamma\mathbf{q}^T\mathbf{q}' + \|\mathbf{q}'\|^2}{1 + \gamma^2}$$

where  $\mathbf{q} = (I - SS^+)\mathbf{m}$  and  $\mathbf{q}' = (I - SS^+)\mathbf{m}'$ .

As  $\phi$  is a rational function with no real pole and goes to  $+\infty$  at  $\pm\infty$ , the optimal  $\gamma$  is a zero of:

$$\begin{aligned}
 \frac{d\phi}{d\gamma} &= 2 \frac{(\gamma\mathbf{q}^T\mathbf{q} - \mathbf{q}^T\mathbf{q}')(1 + \gamma^2) - \gamma(\gamma^2\mathbf{q}^T\mathbf{q} - 2\gamma\mathbf{q}^T\mathbf{q}' + \mathbf{q}'^T\mathbf{q}')}{(1 + \gamma^2)^2} \\
 &= \frac{2}{(1 + \gamma^2)^2} (\gamma(\mathbf{q}^T\mathbf{q} - \mathbf{q}'^T\mathbf{q}') + \gamma^2\mathbf{q}^T\mathbf{q}' - \mathbf{q}^T\mathbf{q}'). \tag{15}
 \end{aligned}$$

As the discriminant of the numerator

$$\Delta = (\mathbf{q}^T\mathbf{q} - \mathbf{q}'^T\mathbf{q}')^2 + 4(\mathbf{q}^T\mathbf{q}')^2 \geq 0 \tag{16}$$

is positive, there are two distinct roots:

$$\gamma_\epsilon = \frac{\mathbf{q}'^T\mathbf{q}' - \mathbf{q}^T\mathbf{q} + \epsilon\sqrt{\Delta}}{2\mathbf{q}^T\mathbf{q}'} \quad \text{with } \epsilon \in \{-1, 1\} \tag{17}$$

The second derivative of  $\phi$  is positive for the optimal  $\gamma$  and:

$$\begin{aligned}
 \frac{d^2\phi}{d\gamma^2} &= \frac{d}{d\gamma} \left( \frac{2\mathbf{q}^T\mathbf{q}'}{(1 + \gamma^2)^2} (\gamma - \gamma_{opt}^+) (\gamma - \gamma_{opt}^-) \right) \\
 &= \frac{2\mathbf{q}^T\mathbf{q}'}{(1 + \gamma^2)^4} [(\gamma - \gamma_{opt}^+ + \gamma - \gamma_{opt}^-)(1 + \gamma^2)^2 \\
 &\quad - (\gamma - \gamma_{opt}^+) (\gamma - \gamma_{opt}^-) 4\gamma(1 + \gamma^2)] \tag{18}
 \end{aligned}$$

and consequently

$$\frac{d^2\phi}{d\gamma^2}(\gamma_\epsilon) = \frac{\epsilon 2\sqrt{\Delta}}{(1 + (\gamma_{-1})^2)^2} \tag{19}$$

and we conclude that  $\epsilon = 1$  i.e.  $\gamma^+$  is the optimal  $\gamma$ .

### III. VALIDATION ON SIMULATIONS

In this section, we consider the practical case of the estimation of contributions of PpIX in a fluorescence signal. This fluorophore of interest is widely used in the operating theatre in neuro-oncology to characterize the healthy/tumoral boundary of a brain tumor named glioma [34], [35]. Glioma is the most frequent and aggressive primary brain tumor whose main therapy is complete tumor removal surgery [37]. Thus, the neurosurgeon uses this information on PpIX contributions to remove or not the measured tissue. Now, we compare the proposed method with state-of-the-art methods, including

blind sources separation method when possible. We consider the two different forms of PpIX described in [15]. Besides the well-known PpIX fluorescence emission appearing with a maximum at 634 nm under excitation at 405 nm, another fluorescence band can be observed with a maximum at 620 nm under specific conditions [15]. These two bands, hereinafter labeled PpIX620 and PpIX634, can be interpreted as bands corresponding to two different forms of PpIX [8], [15], [20], both ones having some interest. Parameters with subscript 1 denote quantities related to PpIX620 and parameters with subscript 2 denote quantities related to PpIX634.

We compare our method with multiple excitation models extrapolated from those of the literature. We define a digital phantom calibrated on real High-Grade Gliomas (HGG) and Low-Grade Gliomas (LGG) data to get a realistic simulation environment.

We now distinguish LGG with two pathological statuses, LGG and Healthy, from HGG with four pathological statuses: Solid Part, Dense Margins, Sparse Density Margins, and Healthy.

The baseline is defined as the non-5-ALA induced fluorescence, meaning the fluorescence that would be emitted in the same tissue without prior administration of 5-ALA to the patient. We assume that a pre-processing of the measurements has already taken place to subtract autofluorescence away from the spectral band of interest as proposed in the literature [15], [24]. Thus, the resulting baseline considers only non-5-ALA-induced fluorophores whose emission is almost overlapping with the one of PpIX. The performance of the EB method is estimated for various conditions of baseline fluorophores whose emission spectrum overlaps the one of PpIX. Performance metrics that we use are the error of estimated PpIX contributions and the accuracy, sensitivity, and specificity of classification in the pathological status. Finally, the clinical diagnostic performance is assessed using glioma images synthesized by an experimentally calibrated digital phantom.

#### A. Models for comparison purpose

Using the same notations as in section II-B, we now describe multiple excitation models extrapolated from those of the literature.

1) *Baseline Free Model (BF)*: In the Baseline Free model (BF), the baseline contribution is discarded. In the existing literature [15], [24] after having subtracted an exponential shape function, a least-square regression is used to estimate the fluorescence contributions of PpIX into the fluorescence signal. BF<sub>1</sub> refers to the Baseline Free Model with one excitation wavelength, which can be written in matrix form as follows:

$$\mathbf{m} = S\boldsymbol{\alpha} . \tag{20}$$

BF<sub>2</sub> refers to the Baseline Free Model with two excitation wavelengths. It can be written in matrix form as follows:

$$\tilde{\mathbf{m}} = \begin{pmatrix} S \\ SE \end{pmatrix} \boldsymbol{\alpha} . \tag{21}$$



In both cases, the parameters are estimated using a linear least square regression.

2) *Gaussian Baseline Model (GB)*: In the Gaussian Baseline model (GB) [8], the contribution of the baseline spectrum into the fluorescence signal is a Gaussian function and a least-square method is used to estimate PpIX contributions  $\alpha_1, \alpha_2$ . GB<sub>1</sub> refers to the Gaussian Baseline Model with one excitation wavelength, which is defined in equation 22. GB<sub>2</sub> refers to the Gaussian Baseline Model with two excitation wavelengths, which can be written as follows:

$$\mathbf{m} = \mathbf{S}_{\lambda_e} = \alpha_1 \mathbf{S}_1 + \alpha_2 \mathbf{S}_2 + A \cdot \mathcal{G}(\mu, \sigma) \quad (22)$$

$$\mathbf{m}' = \mathbf{S}'_{\lambda_e} = \alpha_1 \eta'_1 \mathbf{S}_1 + \alpha_2 \eta'_2 \mathbf{S}_2 + A' \cdot \mathcal{G}(\mu, \sigma) \quad (23)$$

where  $\alpha_1, \alpha_2, A, A', \mu, \sigma$  are estimated by a bounded least squares curve fitting method with  $\alpha_1 \in [0, 10^6]$ ,  $\alpha_2 \in [0, 10^6]$ ,  $A \in [0, 10^6]$ ,  $A' \in [0, 10^6]$ ,  $\mu \in [580, 605]$  nm,  $\sigma \in [5, 100]$  nm. Bounds have been defined from [8].

3) *Non-negative matrix factorization (NMF)*: We compared our results with a blind spectral unmixing method, known as non-negative matrix factorization (NMF) [44]–[46], using the *nnmf* internal routine of MATLAB 2022b. In the same trend as the equation 2 in [12], the 1-by-2*p* matrix  $\tilde{\mathbf{m}}^T$  of the measured signal is factored into two nonnegative matrix  $W$  (1-by- $n'$ ) and  $H$  ( $n'$ -by-2*p*) such as  $\tilde{\mathbf{m}}^T \approx W \times H$ . For a small set of  $n'$  nonnegative representative vectors concatenated in  $H$ ,  $W$  represents their weights of  $\tilde{\mathbf{m}}^T$ . In the case of spectral unmixing,  $n'$  can also be seen as the number of different sources sought. For the rest of this article, we will set  $n' = n + 1$  to 3: PpIX620, PpIX634, and the baseline. On a practical aspect, the stopping criteria and initial values of  $W$  and  $H$  were set to default values. **The initial value of  $W$  was a randomly generated array of numbers between 0 and 1. The initial value of  $H$  was set to known reference spectra and a 1-dimensional randomly generated array of numbers between 0 and 1 for the baseline. Experiments have also been performed using randomly generated arrays of numbers between 0 and 1 as the initial values of  $H$  and  $W$ . They lead to slightly poorer classification results which do not change the conclusion.** The stopping criteria were either a residual value lower than or equal to  $1e^{-4}$  or a relative change in the elements of  $W$  and  $H$  lower than or equal to  $1e^{-4}$ .

4) *Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS)*: Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) is a group of techniques for recovering pure response profiles [47] (spectra, pH profiles, time profiles, ...) of chemical constituents or species from unresolved mixtures obtained in chemical processes when no prior information is available on the nature and composition of these mixtures. MCR-ALS methods have been extended to the analysis of many types of experimental data [48], including multi-channel data and non-evolutionary processes, such as abundance extraction for Multispectral Fluorescence Lifetime Imaging Microscopy [49]. Practically, we used the toolbox described hereafter [50], **using known reference spectra and a 1-dimensional randomly generated array of numbers between 0 and 1 for the baseline.** The precision was set to 0.01% and no constraints were set on spectra and concentration, as for

the EB method. The maximum iterations boundary was set to 100, which was considered as high enough because the iteration process stopped at iteration 22 caused by reaching the set precision. **Experiments have also been performed using the pure method of extraction of initial spectra. In that case, the iteration process stopped at iteration 52 caused by reaching the set precision. They lead to slightly poorer classification results which do not change the conclusion.**

### B. Fluorescence digital phantom

To make a fair comparison of the new EB method with existing ones, we use a fluorescence digital phantom. This phantom includes PpIX and also additive fluorophores which compose the baseline. The digital phantom is parameterized from in vivo data [8] and experimental data obtained on liquid PpIX phantom made as in [20], [51].

We consider two excitation wavelengths  $\lambda_e = 405$  nm and  $\lambda'_e = 385$  nm which are the respective maxima of the absorption spectrum of the two forms of PpIX considered here; the quantum yield ratio for PpIX620 is  $\eta'_1/\eta_1 = 1.62$ ; for PpIX634, it is  $\eta'_2/\eta_2 = 0.76$ , referring to [41]. The measurements and reference spectra used in the generation and the processing related to this part have a dimension  $p=819$ . To make the baseline fluorescence spectrum neutral towards the studied models, the latter is modeled as a Lorentzian function with various parameters: the quantum yield  $\gamma(\lambda_e)$ , central emission wavelength  $\lambda_{para}$ , standard deviation  $\sigma_{para}$ , and amplitude  $A$ .

A two-part noise calibrated on experimental data is included in the fluorescence simulation phantom: a Poisson noise  $\mathcal{P}$  mimics photon noise and an additive Gaussian white noise  $\mathcal{N}$  mimics electronic noise. This leads to the following function representing the generated spectrum:

$$\mathbf{m} = \mathcal{P} [\alpha_1 \eta_1 \mathbf{S}_1 + \alpha_2 \eta_2 \mathbf{S}_2 + \mathbf{b}(\lambda_e)] + \mathcal{N}, \quad (24)$$

where

$$\mathbf{b}(\lambda_e) = \gamma(\lambda_e) \cdot A \cdot \frac{2}{\pi \sigma_{para}} \frac{1}{1 + \left( \frac{\lambda - \lambda_{para}}{\frac{\sigma_{para}}{2}} \right)^2}.$$

All results given in this work have been computed over  $10^5$   $\alpha$  random draws equally split into 100x100 bins to map the region  $\alpha \in [0, 2] \times [0, 2]$ . Each region of the  $\alpha$  plane is associated with a pathological status (cf. Fig. 2). We focus on the boundary between healthy and tumoral tissues, which corresponds to an area of this plane. All random draws were performed with Gaussian distributions linked with pathological status, whose parameters are extracted from previous in vivo works [8].

### C. Sets of Baseline parameters

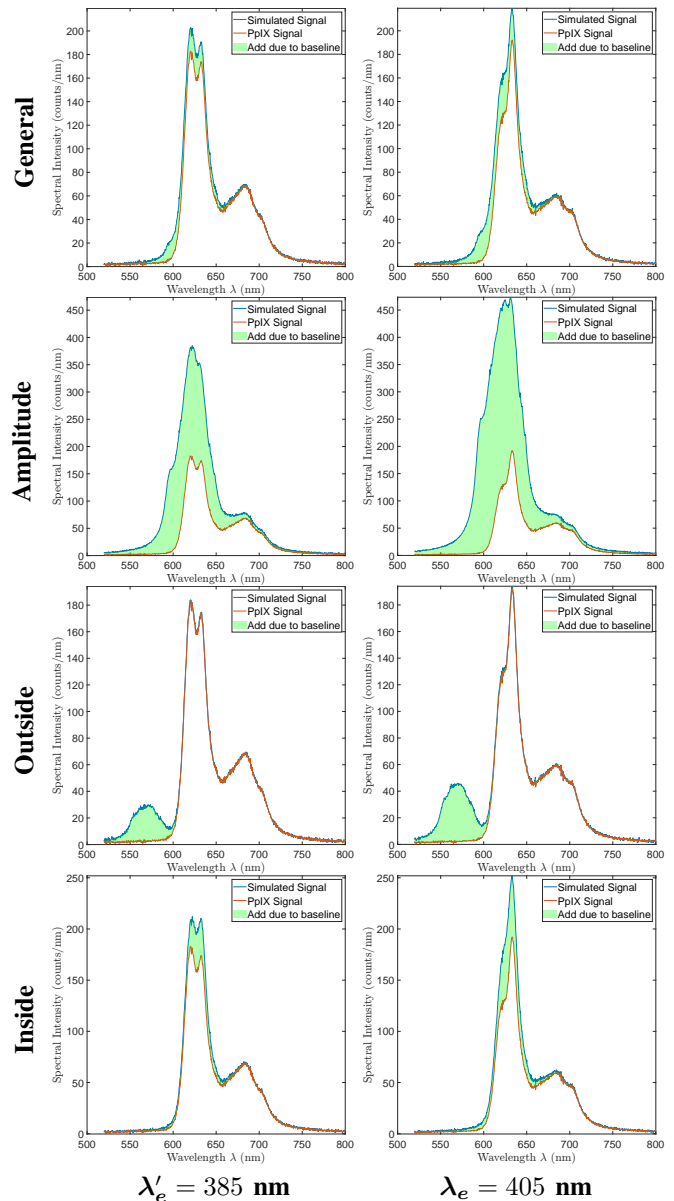
Seven baseline datasets have been simulated. These datasets correspond to different scenarios for the application of our method in clinical practice. The simulation parameters are given in Table I where  $\mathcal{U}(a, b)$  is a uniform distribution between  $a$  and  $b$ , and  $\Gamma(k, \beta)$  is a gamma distribution with shape factor  $k$  and scale factor  $\beta$ .

Name of the dataset	Parameters	Description
<b>General</b>	$A = \Gamma(4.102, 82.6840)$ u.a. $\lambda_{para} = \mathcal{U}(590, 650)$ nm $\sigma_{para} = \mathcal{U}(0, 25)$ nm $\gamma = \mathcal{U}(0.1, 1.125)$	General case with all the existing problems due to an endogenous fluorophore of low amplitude compared to PpIX
<b>Amplitude</b>	$A = \Gamma(4.102, 826.684)$ u.a.	Study of the influence of endogenous fluorophore's** amplitude. For example, the amplitude of the lipopigments fluorescence signal increases with aging [52], [53].
<b>Width</b>	$\sigma_{para} = \mathcal{U}(0, 50)$ nm	Study of the influence of the spectral width of the endogenous fluorophore**. For example, the spectral width of the lipofuscin classically extends on 80 nm [54] but in certain studies, it is extended well beyond [23].
<b>Outside</b>	$\lambda_{para} = \mathcal{U}(550, 590)$ nm $\gamma = \mathcal{U}(0.1, 3.00)$	Study of the influence of the central wavelength of the endogenous fluorophore** located in the emission range of the PpIX. For example, water-soluble porphyrins have an emission maximum near the PpIX emission range [17], [19], [25], [55].
<b>Inside</b>	$\lambda_{para} = \mathcal{U}(620, 640)$ nm $\gamma = \mathcal{U}(0.1, 3.00)$	
<b>Inner</b>	$\gamma = \mathcal{U}(1, 1.2)$	Study of the influence of the quantum yield of the endogenous fluorophore**. For example, endogenous fluorophores such as water-soluble porphyrins may have quantum yields close to those of PpIX620 and PpIX634.
<b>Across</b>	$\gamma = \mathcal{U}(0.6, 0.8)$	

**TABLE 1:** Parameters used for the simulation of the seven baseline datasets. Only differences from "General" are specified for the following six datasets. \*\*In this Table, endogenous fluorophore corresponds to a non-5-ALA-induced fluorophore.

The first three datasets of parameters are defined to study the robustness regarding the amplitude or width parameters of the fluorescence spectrum attached to the baseline. The Outside and Inside datasets study the influence of the position of the baseline central wavelength regarding the spectral emission band of PpIX. The Inner and Across datasets are defined to study the impact of the baseline fluorescence quantum yield. This parameter is particularly critical regarding the instability areas of EB (see II-B.2). The Inner dataset models baseline quantum yields ranging out of the PpIX instability spots,

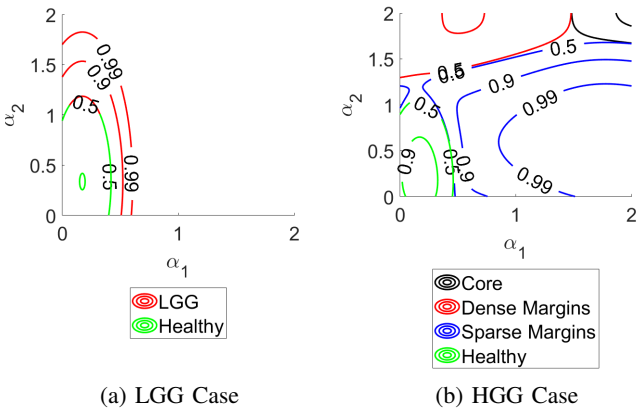
whereas the Across set models the intersection of the baseline and PpIX quantum yields ranges. An example of fluorescence spectrum for some of the baseline sets is plotted in Fig. 1.



**Fig. 1:** Measurement signal (in blue) and PpIX signal (in orange) for each baseline dataset in row and for each excitation wavelength in column. The green area corresponds to the signal addition due to the baseline. Each signal  $m$ , *resp.*  $m'$ , corresponds to a vector such as  $m^T$ , *resp.*  $m'^T$ ,  $\in \mathbb{R}^{p \times 1}$  with  $p = 819$ .

#### D. Classification method

To reach the final aim of pathological status prediction, one evaluated the ability to classify tissues from the estimated contributions of PpIX inside the fluorescence spectrum. The same simulated measurements  $\tilde{m}$  have been independently processed by each estimation model, and estimated PpIX contributions  $\alpha^*$  have been classified with a naive Bayesian



**Fig. 2:** Display level set of probability map to be in the class of the given color. In a), LGG class corresponds to the red color and the Healthy class to the green one. In b), the Core class corresponds to the black color, Dense Margins to the red color, Sparse Margins to the blue color, and the Healthy class to the green one. In each figure, the horizontal axis corresponds to  $\alpha_1$  and the vertical axis to  $\alpha_2$ .

classifier [56]. Knowledge of classes given in [8] has been used to define *a priori* probability laws associated with each class. Ground truth PpIX contributions  $\alpha$  used in the phantom have also been classified to serve as a reference for classification performance. For each bin, we compute the accuracy  $acc$  by using the formula below where  $c$  is a class,  $c^*$  an estimated class,  $C$  the set of classes,  $s$  a sample,  $S_c$  the set of samples in class  $c$ ,  $c_s^*$  the estimated class for the sample  $s$ ,  $N$  the total number of samples in a bin, and  $N_c$  the number of samples of the bin that belong to the class  $c$ . For each bin of any map, the color is related to the accuracy value, the greener the higher accuracy.

$$acc = \sum_{c \in C} \left( \frac{\sum_{s \in S_c} (c_s^* = c)}{N_c} \frac{N_c}{N} \right) = \sum_{c \in C} p(c^* = c) \cdot p(c) \quad (25)$$

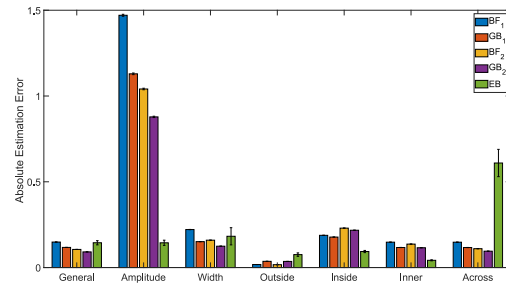
$p(c)$  follows a gaussian probability law whose parameters are defined in [8]. The accuracy is computed in LGG or HGG cases. For an HGG, the estimated parameter pair of each measurement may be associated with a most likely class between Core, Dense Margins, Sparse Margins, or Healthy. For an LGG, the same classification may occur for the two following classes: Low Grade, i.e. tumoral, or Healthy. In Fig. 2, we computed the probability map of classes of the HGG case, resp. LGG, to help interpret the accuracy maps. This allows us to associate a most likely class to each region of the  $(\alpha_1, \alpha_2)$  plane. The region in the bottom left corner of each map corresponds to the healthy tissue region. One notices that the probability 0.5 represents the border between two different classes of the  $\alpha$  plane region.

We also compare our results with blind spectral unmixing methods such as non-negative matrix factorization (NMF) and the Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS). Since the latter do not necessarily provide unmixing results in  $(\alpha_1, \alpha_2)$  space, the comparison was performed after the final output of the spectral unmixing

followed by classification process, in sections III-E.2 and III-E.3. We chose not to proceed with output space adaptation of the spectral unmixing method to keep a fair comparison between all methods. To perform the classification, it was necessary to train separate Bayesian classifiers for HGG and LGG on the general dataset for the NMF, resp. MCR-ALS results. The training of the four separate Bayesian classifiers (two for NMF and two for MCR-ALS) is conducted using 45 items for each class. To compare the classification results of each state-of-the-art method, we opted to generate receiver operating characteristic (ROC) curves and calculate the area under the curve (AUC). The higher the AUC, the better the classifier. The AUC was computed after using NMF, MCR-ALS, or estimation models.

### E. Results and Discussion

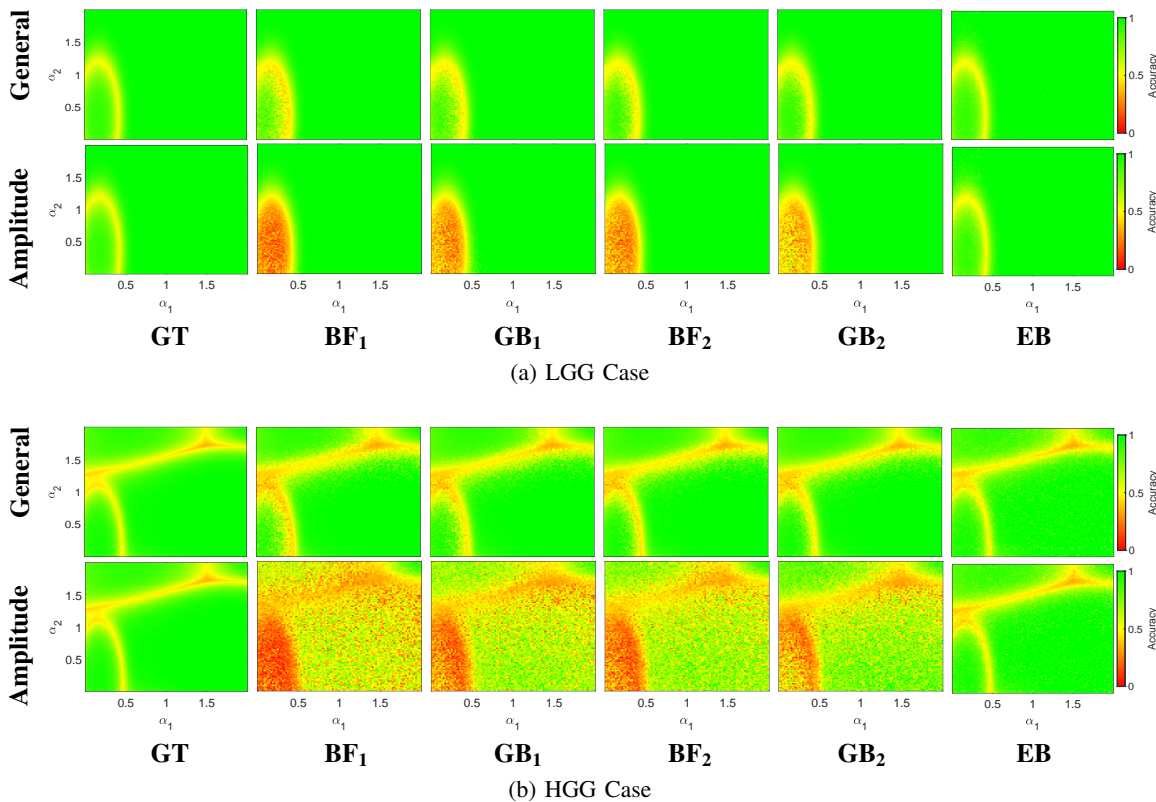
Fluorescence signal by non-5-ALA-induced fluorophores varies a lot, which leads to important crosstalk with PpIX. Each baseline set of parameters represents one example among all possible spectra. We now look at the estimation error on the parameters of interest with every model and for all these sets.



**Fig. 3:** Mean and standard error of Absolute estimation Error of  $\alpha$ . Each color bar corresponds to an estimation method, and each group corresponds to a set of baseline parameters (refer Table I).  $BF_1$ , resp.  $GB_1$ ,  $BF_2$ ,  $GB_2$  refers to the Baseline Free Model with one excitation wavelength, resp. the Gaussian Baseline Model with one excitation wavelength, the Baseline Free Model with two excitation wavelengths, and the Gaussian Baseline Model with two excitation wavelengths. The Estimated Baseline Model (EB) is our proposed method.

1) *Estimation Error on  $\alpha$* : In Fig. 3 are presented the mean value and the standard error of the absolute estimation error between the ground truth  $\alpha$  used to generate spectra with the numerical phantom and  $\alpha^*$  estimated using a given method on a specific baseline set of parameters (see Table I). One sees that with every model, the error increases a lot except for the EB for which the error remains rather constant. Concerning General, Amplitude, and Width datasets, we can notice that existing models are all quite insensitive to Width but very sensitive to Amplitude and that  $EB_2$  remains insensitive to Width while solving the Amplitude issue. Considering the Outside dataset, existing models work well because the emission band of the baseline is far from the PpIX emission spectrum: crosstalk is low, and baseline and PpIX are decorrelated.





**Fig. 4:** Accuracy Map for each estimation method in column : GT, *resp.*  $BF_1$ ,  $GB_1$ ,  $BF_2$ ,  $GB_2$  refers to Ground Truth, *resp.* the Baseline Free Model with one excitation wavelength, the Gaussian Baseline Model with one excitation wavelength, the Baseline Free Model with two excitation wavelengths, the Gaussian Baseline Model with two excitation wavelengths. The Estimated Baseline Model (EB) is our proposed method. We classify on one side Low Grade Glioma (a) and on the other side High Grade Glioma (b). Columns give the set name defined in Table I : the baseline's amplitude increased when going from General to Amplitude dataset.

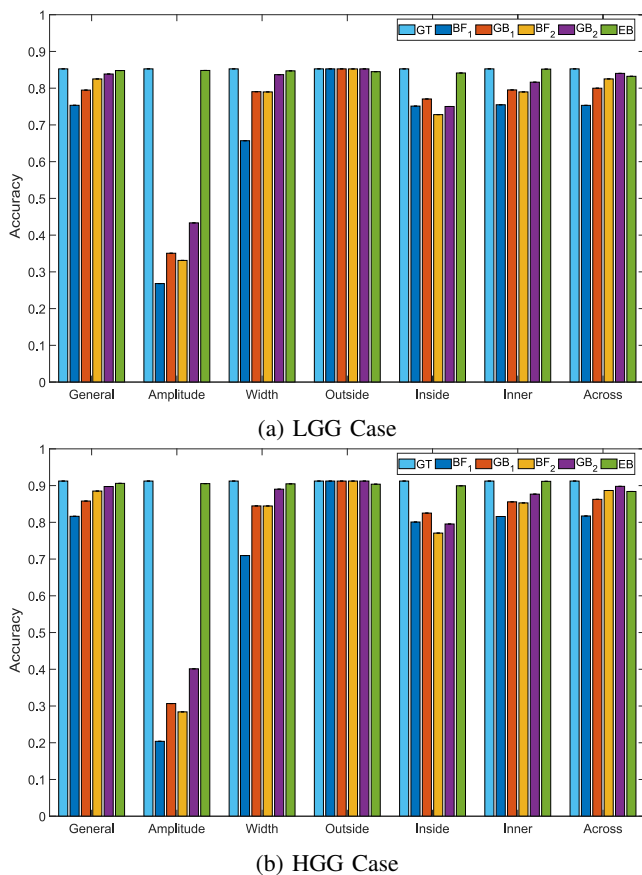
Comparing the Inside dataset with the Outside one, we notice that the error made by state-of-the-art models is increased on average by at least a factor of 5. At the same time, the EB model varies very little: the error is on average increased by a factor of 1.24. EB is not very sensitive to Inside/Outside datasets since it has no a priori on the spectrum of the baseline while existing methods are very sensitive to the baseline emission band.

For Inner Dataset, EB has the lowest error and, at the same time, other models are not better than the General case. The Inner Dataset is chosen to be favorable for EB since the baseline quantum yield  $\gamma$  is decorrelated from those of PpIX. Because existing models do not consider it, they are insensitive to changes of  $\gamma$ . On the contrary, Across Dataset is chosen to be unfavorable for EB. One notices the significant increase in EB's error when the quantum yield of the baseline approaches those of PpIX. This is a direct consequence of specific hot areas explained in the previous Section II-B.2. It seems important to point out that the case modeled by the Across dataset is not very plausible physically, because it considers a fluorophore with a quantum yield  $\gamma$  close to PpIX without the spectral shape corresponding to the one of PpIX. Moreover, if this case happens, one solution to tackle

this issue can be to change  $\lambda_e$ , so the hot points will be outside the area of interest.

EB is not always the model that has the lowest error, especially when the *a priori* are in favor of existing models (see for example the Outside dataset). In any case, it maintains a low error even when the amplitude increases. The Across case is not very plausible physically. A more plausible physical case where fluorophores are very close to PpIX would be a merge of Across and Inside datasets. We assume that all the models would become bad because the case is very difficult for all models, even EB would not solve this case.

**2) Impact of  $\alpha^*$  on classification:** In Fig. 4, one sees with both LGG and HGG that with the Amplitude dataset, EB is the only model able to correctly classify Healthy samples when the baseline amplitude increases. Comparing single- and two- excitation wavelength models for the Amplitude dataset, one can notice that two excitation wavelength models ( $GB_2$ ,  $BF_2$ ) improve a little the correct classification of Healthy class samples located in the center of the region. It is confirmed by an increase of  $BF_2$ , *resp.*  $GB_2$ , accuracy compared to  $BF_1$ , *resp.*  $GB_1$  in Fig. 5. Thus, the success of EB is more due to the estimated baseline than the multiple excitation wavelengths. In the case of HGG, the border between Low-Density Margins,



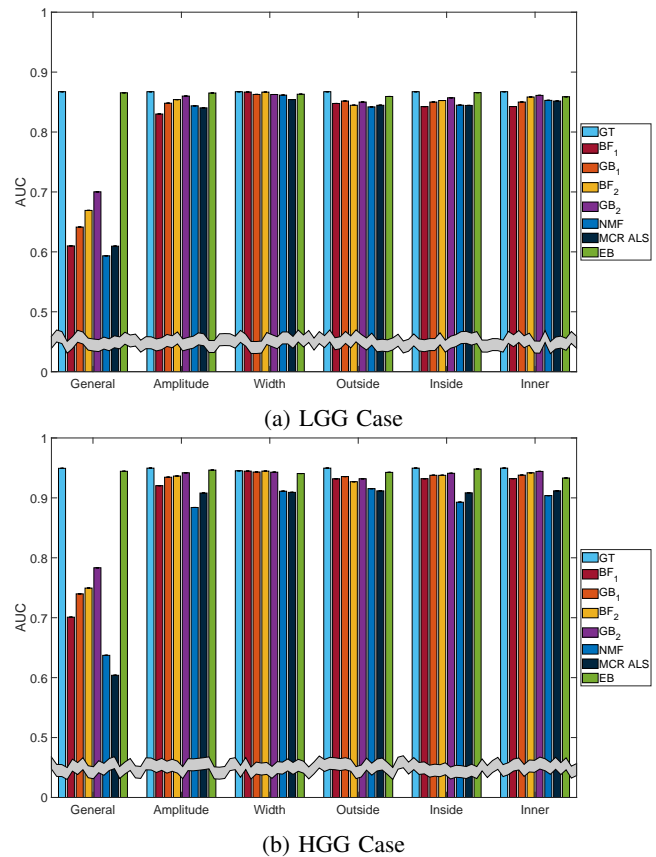
**Fig. 5:** Mean Accuracy of the region of interest ( $0 \leq \alpha_1 \leq 0.25$ ,  $0 \leq \alpha_2 \leq 0.75$ ). Each color corresponds to the same model and each group to a specific set of baseline parameters. GT, *resp.*  $BF_1$ ,  $GB_1$ ,  $BF_2$ ,  $GB_2$  refers to Ground Truth, *resp.* the Baseline Free Model with one excitation wavelength, the Gaussian Baseline Model with one excitation wavelength, the Baseline Free Model with two excitation wavelengths, the Gaussian Baseline Model with two excitation wavelengths. The Estimated Baseline Model (EB) is our proposed method.

High-Density Margins, and Core is thicker for  $GB_1$ ,  $GB_2$  and  $BF_1$ ,  $BF_2$  models than for EB which is a direct consequence of the estimation error by these models, refer to Fig 3. In addition, the shading effect in Fig. 4 reveals a loss in classification precision for core and margins classes. In the case of LGG, the ability to classify is lost because there are only two classes. For the Width dataset, we notice that EB is the only model whose accuracy remains constant with both LGG and HGG. Thus, EB seems to be the most relevant model if the baseline amplitude or width has an important variability.

We now study the impact of the baseline central wavelength on accuracy (see Fig. 5). Considering the Outside dataset, all models have similar performance. One can notice for the Inside dataset that EB is the only model whose accuracy overpasses 75%. Despite the estimation error being higher in EB than in existing models, the classification performance is equivalent in the Outside dataset where there is no crosstalk between baseline and PpIX fluorescence spectra. We can thus consider using EB even when there is no crosstalk issue since

it is as good as the state-of-the-art methods. This result is striking because EB was not particularly expected to perform in this case. This result reveals the strength of EB compared to other models: it is less sensitive to the emission band of the baseline.

Let us assess the impact of baseline fluorescence quantum yield on accuracy. For the Inner dataset, the accuracy is the same as in the General dataset for all models. Thus, the drop of error with EB, in this case, does not affect the classification.



**Fig. 6:** Area under the receiver operating characteristic curve (AUC). Each color corresponds to the same model and each group to a specific set of baseline parameters. GT, *resp.*  $BF_1$ ,  $GB_1$ ,  $BF_2$ ,  $GB_2$  refers to Ground Truth, *resp.* the Baseline Free Model with one excitation wavelength, the Gaussian Baseline Model with one excitation wavelength, the Baseline Free Model with two excitation wavelengths, and the Gaussian Baseline Model with two excitation wavelengths. The AUC of the Non-Negative Matrix Factorization (NMF) spectral unmixing method, *resp.* the Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) method, is represented in each group by the royal blue bar, *resp.* by the dark blue bar. The Estimated Baseline Model (EB) is our proposed method.

Since the strength of the estimated baseline model is given by the estimated baseline, one now wants to compare it to blind spectral unmixing methods. We chose NMF as the comparison method. However, since the estimated output parameters of each model are not in the same space, we observe the receiver operating characteristic curve and the area under this curve for each method. Fig. 6b, *resp.* Fig. 6a, represents the area under

the curve (AUC) for the different methods and datasets in the HGG case, *resp.* LGG case.

The General dataset is the classifier training dataset for NMF. One notices that in terms of AUC on this dataset the NMF method performs worse than our Estimated Baseline model for both the LGG and HGG cases.

As the baseline's amplitude increases in the dataset Amplitude, the AUC of the NMF method drops as much as that of the state-of-the-art models. When the baseline's width increases in the dataset Width, the AUC of the NMF method is reduced a little compared to the Estimated Baseline Model's one.

Considering the Outside dataset, all models have similar performance for LGG. For HGG, NMF performs somewhat worse than all other methods, including EB. For the Inside dataset, the AUC of NMF and peak methods are lower than GT's AUC whereas the EB and GT's AUC are similar for both LGG and HGG.

One focuses on the effect of baseline quantum yield on the AUC. For the Inner dataset, the AUC is the same as for the General dataset for all models: the AUC of NMF is lower than the one of EB. For the Across dataset associated with HGG, we note that the AUC of NMF remains lower than the one of EB. For LGG, NMF has an AUC in the same range as the one of EB.

Indeed, EB has a better accuracy and a higher AUC than the existing models including NMF when the baseline amplitude increases, and it keeps an accuracy and AUC equivalent to those of existing models in the unfavorable case of Across dataset where the baseline quantum yield  $\gamma$  is close to those of the PpIX. One can imagine using EB all the time even if in cases modeled by the Across dataset since at worst it will be equivalent to state-of-the-art performances.

3) *Clinical Practice performance on synthesized glioma image*: To measure the final impact of this study on clinical practice, we apply the "estimation + classification" pipeline to an image of glioma considering a real experimental system.

Each glioma is either LGG (Fig. 7a) or HGG (Fig. 7b). Clinically, the average size of the glioma is 3 cm, we set its size to 75% of image size so the pixel size corresponds to 130  $\mu\text{m}$  which corresponds to the typical scale at the tissue level. A single spectrum is generated in each pixel according to the pixel's ground truth class. Thus, the emission spectrum of the baseline and the contributions of the fluorophore of interest are different between the two neighboring pixels. The size of a measurement probe is 4 mm (30 pixels in the image) which corresponds to the typical size of a spectroscopic probe or hyperspectral imaging system in the literature [8]. Images of Fig. 7 account for the spatial point spread function of the probe: the initial  $300 \times 300 \times 2x_p$  hyperspectral images have been padded with replicates of the borders of the original content and then convolved with a Gaussian kernel of size  $30 \times 30$  pixels. Finally, the resulting  $300 \times 300 \times 2x_p$  hyperspectral images, called Clinical Synthetic Images (CSI), are clipped to the original size. For all images, the estimation is performed independently pixel by pixel. To numerically estimate the classification, we simplify the problem into two classes, healthy or tumor tissue. We thus calculate the sensitivity also called true positive rate (TPR) and specificity called true negative rate (TNR) of each classification for all datasets. These metrics are relevant because sensitivity refers to the model's ability to correctly detect tumor tissues, while specificity relates to the model's ability to correctly detect healthy tissues. Case General 1px of Tables II is equivalent to section III-E.2 because the convolution effect of the probe is not considered. One notice that the classification task is an ill-posed problem because the sensitivity and specificity of ground truth are unable to reach 1.0.

We first focus on the Amplitude and General Datasets. Corresponding sensitivity (TPR) and specificity (TNR) values are given in Tables IIa, IIb. We notice a strong impact of the measurement probe's convolution on the classification sensitivity (TPR) which goes from approximately 80% to

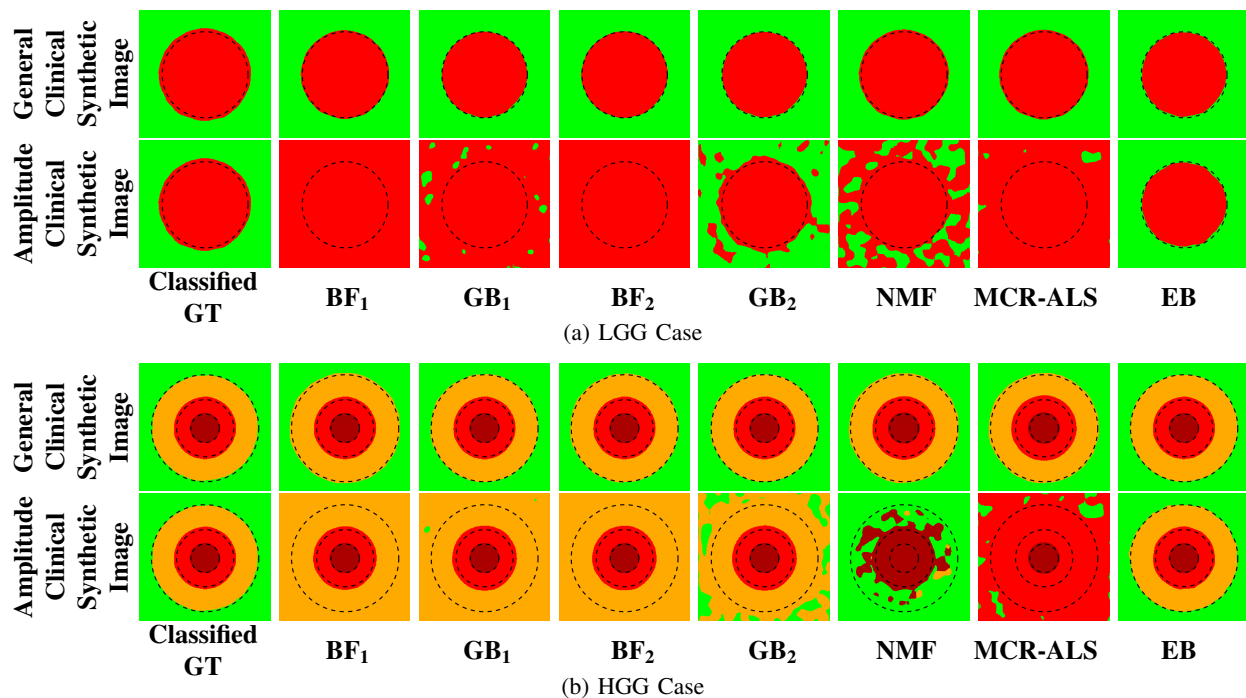
Model	General 1px		General CSI		Amplitude CSI		Outside CSI		Inside CSI		Inner CSI		Across CSI	
	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR
Ground Truth	0.58	0.79	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00	0.94	1.00
BF <sub>1</sub>	0.41	0.86	0.16	1.00	0.00	1.00	0.93	1.00	0.16	1.00	0.16	1.00	0.16	1.00
GB <sub>1</sub>	0.46	0.84	0.92	1.00	0.05	1.00	0.96	1.00	0.21	1.00	0.92	1.00	0.92	1.00
BF <sub>2</sub>	0.48	0.83	0.80	1.00	0.00	1.00	0.94	1.00	0.65	1.00	0.22	1.00	0.69	1.00
GB <sub>2</sub>	0.51	0.82	0.93	1.00	0.09	1.00	0.95	1.00	0.87	1.00	0.92	1.00	0.93	1.00
NMF	0.54	0.80	0.97	1.00	0.47	1.00	0.98	1.00	0.95	1.00	0.96	1.00	0.96	1.00
MCR ALS	0.57	0.79	0.97	1.00	0.18	1.00	0.99	1.00	0.96	1.00	0.96	1.00	0.97	1.00
<b>EB</b>	<b>0.57</b>	<b>0.79</b>	<b>0.94</b>	<b>1.00</b>	<b>0.93</b>	<b>1.00</b>	<b>0.94</b>	<b>1.00</b>	<b>0.94</b>	<b>1.00</b>	<b>0.94</b>	<b>1.00</b>	<b>0.94</b>	<b>1.00</b>

(a) LGG Case

Model	General 1px		General CSI		Amplitude CSI		Outside CSI		Inside CSI		Inner CSI		Across CSI	
	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR
Ground Truth	0.59	0.89	1.00	0.97	1.00	0.97	1.00	0.97	1.00	0.97	0.90	1.00	0.90	1.00
BF <sub>1</sub>	0.43	0.93	0.98	1.00	0.01	1.00	1.00	0.98	0.97	1.00	0.26	1.00	0.26	1.00
GB <sub>1</sub>	0.48	0.92	1.00	0.99	0.29	1.00	1.00	0.95	0.98	1.00	0.87	1.00	0.87	1.00
BF <sub>2</sub>	0.49	0.91	0.99	1.00	0.09	1.00	1.00	0.97	0.99	1.00	0.49	1.00	0.79	1.00
GB <sub>2</sub>	0.52	0.91	1.00	0.98	0.85	1.00	1.00	0.96	0.99	1.00	0.87	1.00	0.88	1.00
NMF	0.76	0.82	0.93	1.00	1.00	0.70	0.98	1.00	0.91	1.00	0.92	1.00	0.93	1.00
MCR ALS	0.79	0.81	0.95	1.00	0.30	1.00	0.99	1.00	0.95	1.00	0.94	1.00	0.95	1.00
<b>EB</b>	<b>0.59</b>	<b>0.89</b>	<b>1.00</b>	<b>0.97</b>	0.99	<b>0.97</b>	<b>1.00</b>	<b>0.97</b>	<b>1.00</b>	<b>0.97</b>	<b>0.90</b>	<b>1.00</b>	<b>0.90</b>	<b>1.00</b>

(b) HGG Case

TABLE II: Sensitivity (TPR) and Specificity (TNR) values for different datasets including clinical synthetic images (CSI) ones.



**Fig. 7:** Results on 300x300 clinical synthetic images of Low-Grade Glioma (a) and High Grade Glioma (b). Row name corresponds to an estimation model: GT, *resp.*  $BF_1$ ,  $GB_1$ ,  $BF_2$ ,  $GB_2$  refers to Ground Truth, *resp.* the Baseline Free Model with one excitation wavelength, the Gaussian Baseline Model with one excitation wavelength, the Baseline Free Model with two excitation wavelengths, and the Gaussian Baseline Model with two excitation wavelengths. The Estimated Baseline Model (EB) is our proposed method. Columns gives the set name defined in Table I: the baseline's amplitude increased when going from General to Amplitude dataset. The red color refers to the tumor area while the green color refers to the healthy area. In HGG Case (b), the orange color refers to sparse tumor margins.

100%. In the same way, the specificity (TNR) increases from 59% to 73 % at best. This is relevant in terms of clinical practice since the neurosurgeon removes millimeter-sized pieces that therefore give very high sensitivity. Moreover, previous results are confirmed on simulated glioma: only EB can correctly distinguish healthy tissues from tumoral ones when the amplitude of the baseline increases.

One can see a loss of specificity when comparing the Inside with the Outside dataset for all models except EB. This goes along with previous results. In addition, the specificities of the Outside dataset are higher than in the General dataset. This is due to the difference in baseline positions between the two datasets. Moreover, in the HGG case, we notice that the specificity of  $GB_1$  and  $GB_2$  is better than the one of Ground Truth. This bias of the method has no physical sense and is because the Lorentzian is narrower than the Gaussian.

Thus, when the Lorentzian is positioned where the Gaussian of the GB models can be (i.e. the outside Dataset), then the Gaussian baseline of GB models has higher values than the Lorentzian at the level of the PpIX peaks. As a result,  $\alpha$  is systematically underestimated and we have this effect of a fall in the number of true healthy identified as tumoral and therefore an increase in specificity.

Looking at Inner and Across datasets, we notice that the specificity of each model is not impacted by the change related to these datasets. EB has the best specificity in the unfavorable

case of Across dataset where the baseline quantum yield  $\gamma$  is close to those of the PpIX. From a performance point of view, the EB method is 1.06 times faster than the blind methods which are finally not very accurate, and 1706 times faster than the accurate state-of-the-art methods ( $GB_2$ ). From a clinical point of view, EB is the best of all models in all studied cases.

#### IV. CONCLUSION

In this work, we proposed a novel approach using several excitation wavelengths to estimate the PpIX fluorescence-related biomarkers contributions which make our method free from any *a priori* on the fluorescence emission spectrum of the baseline.

Furthermore, we derived analytical expressions for the least square estimator of the proposed model. This allows our method to be computationally efficient and free from the use of iterative non-linear estimation procedures. We have shown that this new model better estimates the pathological status of the measured tissue when baseline amplitude increases: it keeps a specificity almost equal to that of the ground truth while the one of existing models drops to 0. In addition, this new model is as accurate as existing models in unfavorable cases, e.g., when baseline quantum yield is near PpIX ones. This new model reveals the potential of multiple excitation wavelengths to increase the classification sensitivity of fluorescence measurements.



Future work may focus on hot spot management and extend this study to more than two excitation wavelengths. In addition, complementary aims will be explored regarding the application of the model to real clinical data and the correlation between the baseline and the fluorophores of interest. Furthermore, the characterization of the execution speed depending on the image size and the size  $p$  of the spectra could be explored in relation to clinical constraints. This work assumes the independence of the baseline spectral shape with the excitation wavelength, which is consistent with the physics of fluorescence. However, complex microenvironments could induce mechanisms that alter this assumption, and the EB model's robustness to such effect could be explored. **As described in [36], 5-ALA-induced PpIX is the clinical gold standard for HGG despite the associated financial constraints. In neurosurgery, compared with other fluorescent agents, 5-ALA-induced PpIX is an agent that specifically labels tumor cells. For LGG, the use of 5-ALA as a surgical adjunct remains more limited mainly by limited PpIX visible fluorescence, which reduces the signal-to-noise ratio of the measurements, and the post-processing applied gives poorer results. The EB model presented and characterized in this article aims at better differentiating the PpIX 620 and 634 forms, which are present together in cases where the fluorescence signal is weak. Finally, the EB model could be applied to other fluorophores of clinical interest, particularly to try to determine the relative quantities of NADH protein-bound form and free one.**

## REFERENCES

- [1] O. Gutierrez-Navarro *et al.*, "A fully constrained optimization method for time-resolved multispectral fluorescence lifetime imaging microscopy data unmixing," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 6, pp. 1711–1720, 2013.
- [2] J. Sebastian *et al.*, "Spike estimation from fluorescence signals using high-resolution property of group delay," *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 2923–2936, 2019.
- [3] G. Palmer *et al.*, "Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer (march 2003)," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 11, pp. 1233–1242, 2003.
- [4] A. M. Mohs *et al.*, "An integrated widefield imaging and spectroscopy system for contrast-enhanced, image-guided resection of tumors," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 5, pp. 1416–1424, 2015.
- [5] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral Unmixing With Spectral Variability Using a Perturbed Linear Mixing Model," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 525–538, Oct 2015.
- [6] S. Kaneko *et al.*, "Fluorescence-Based Measurement of Real-Time Kinetics of Protoporphyrin IX After 5-Aminolevulinic Acid Administration in Human In Situ Malignant Gliomas," *Neurosurgery*, vol. 85, no. 4, pp. E739–E746, Oct. 2019.
- [7] A. C. Croce and G. Bottiroli, "Autofluorescence Spectroscopy and Imaging: A Tool for Biomedical Research and Diagnosis," *Eur. J. Histochem.*, vol. 58, no. 4, Oct 2014.
- [8] L. Alston *et al.*, "Spectral complexity of 5-ala induced ppix fluorescence in guided surgery: a clinical study towards the discrimination of healthy tissue and margin boundaries in high and low grade gliomas," *Biomed. Opt. Express*, vol. 10, no. 5, pp. 2478–2492, May 2019.
- [9] A.-S. Montcuquet *et al.*, "In vivo fluorescence spectra unmixing and autofluorescence removal by sparse nonnegative matrix factorization," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2554–2565, 2011.
- [10] E. Barash *et al.*, "Multiplexed analysis of proteins in tissue using multispectral fluorescence imaging," vol. 29, no. 8, pp. 1457–1462, 2010.
- [11] F. Woolfe *et al.*, "Autofluorescence removal by non-negative matrix factorization," *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1085–1093, 2011.
- [12] A.-S. Montcuquet *et al.*, "Nonnegative matrix factorization: a blind spectra separation method for in vivo fluorescent optical imaging," *Journal of biomedical optics*, vol. 15, no. 5, pp. 056009–056009, 2010.
- [13] X. Li *et al.*, "Adaptive relationship preserving sparse nmf for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.
- [14] J. Gu, B. Yang, and B. Wang, "Nonlinear unmixing for hyperspectral images via kernel-transformed bilinear mixing models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.
- [15] B. Montcel *et al.*, "Two-peaked 5-ALA-induced PpIX fluorescence emission spectrum distinguishes glioblastomas from low grade gliomas and infiltrative component of glioblastomas," *Biomed. Opt. Express*, vol. 4, no. 4, pp. 548–558, Apr 2013.
- [16] M. Zanella *et al.*, "Multimodal optical analysis discriminates freshly extracted human sample of gliomas, metastases and meningiomas from their appropriate controls." *Sci. Rep.*, vol. 7, p. 41724, Feb 2017.
- [17] W. Dietel *et al.*, "5-aminolaevulinic-acid-induced formation of different porphyrins and their photomodifications," *Lasers Med. Sci.*, vol. 12, no. 3, pp. 226–236, Oct 1997.
- [18] H. C. K. and H. S. M., "Evaluating the effect of local pH on fluorescence emissions from oral bacteria of the genus *Prevotella*." *J. Biomed. Opt.*, vol. 21, no. 8, p. 84003, Aug 2016.
- [19] B. Montcel *et al.*, "5-ALA-induced PpIX fluorescence in gliomas resection: spectral complexity of the emission spectrum in the infiltrative compound," in *Neurophotonics*. International Society for Optics and Photonics, Jun 2013, vol. 8804, p. 880409.
- [20] L. Alston *et al.*, "Nonlinear relation between concentration and fluorescence emission of protoporphyrin IX in calibrated phantoms," *J. Biomed. Opt.*, vol. 23, no. 9, p. 097002, Sep 2018.
- [21] N. Haj-Hosseini *et al.*, "Optical touch pointer for fluorescence guided glioblastoma resection using 5-aminolevulinic acid," *Lasers Surg. Med.*, vol. 42, no. 1, pp. 9–14, Jan 2010.
- [22] D. A. Haidar *et al.*, "Spectral and lifetime domain measurements of rat brain tumors," *Biomed. Opt. Express*, vol. 6, no. 4, pp. 1219–1233, Apr 2015.
- [23] D. Black *et al.*, "Characterization of autofluorescence and quantitative protoporphyrin IX biomarkers for optical spectroscopy-guided glioma surgery - Scientific Reports," *Sci. Rep.*, vol. 11, no. 20009, pp. 1–12, Oct 2021.
- [24] T. Ando *et al.*, "Precise comparison of protoporphyrin IX fluorescence spectra with pathological results for brain tumor tissue identification," *Brain Tumor Pathol.*, vol. 28, no. 1, pp. 43–51, Feb 2011.
- [25] J. J. Bravo *et al.*, "Hyperspectral data processing improves PpIX contrast during fluorescence guided surgery of human brain tumors," *Sci. Rep.*, vol. 7, no. 9455, pp. 1–13, Aug 2017.
- [26] Y. Li *et al.*, "Intraoperative fluorescence-guided resection of high-grade gliomas: a comparison of the present techniques and evolution of future strategies." *World Neurosurg.*, vol. 82, no. 1-2, pp. 175–185, Jul 2013.
- [27] J. T. C. Liu, D. Meza, and N. Sanai, "Trends in Fluorescence Image-Guided Surgery for Gliomas," *Neurosurgery*, vol. 75, no. 1, pp. 61–71, Jul 2014.
- [28] P. A. Valdés *et al.*, "Optical technologies for intraoperative neurosurgical guidance," *Neurosurg. Focus*, vol. 40, no. 3, p. E8., Mar 2016.
- [29] R. Maugeri *et al.*, "With a Little Help from My Friends: The Role of Intraoperative Fluorescent Dyes in the Surgical Management of High-Grade Gliomas," *Brain Sci.*, vol. 8, no. 2, p. 31., Feb 2018.
- [30] E. D. Bander, R. Magge, and R. Ramakrishna, "Advances in glioblastoma operative techniques," *World neurosurgery*, vol. 116, pp. 529–538, 2018.
- [31] G. A. Wagnières, W. M. Star, and B. C. Wilson, "In vivo fluorescence spectroscopy and imaging for oncological applications," *Photochem. Photobiol.*, vol. 68, no. 5, pp. 603–632, Nov 1998.
- [32] I. Dos Santos *et al.*, "Direct quantification of red wine phenolics using fluorescence spectroscopy with chemometrics," *Talanta*, vol. 236, p. 122857., Jan 2022.
- [33] Y. Chen, Y. Tang, and H. Wang, "Quantification of ATP in cell by fluorescence spectroscopy based on generalized ratio quantitative analysis model," *Spectrochim. Acta, Part A*, vol. 263, p. 120170., Dec 2021.
- [34] W. Stummer *et al.*, "Technical principles for protoporphyrin-IX-fluorescence guided microsurgical resection of malignant glioma tissue," *Acta Neurochir.*, vol. 140, no. 10, pp. 995–1000, 1998.

- [35] —, “Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase iii trial,” *Lancet Oncol.*, vol. 7, no. 5, pp. 392–401, May 2006.
- [36] R. Sun, H. Cuthbert, and C. Watts, “Fluorescence-guided surgery in the surgical treatment of gliomas: Past, present and future,” *Cancers*, vol. 13, no. 14, p. 3508, 2021.
- [37] D. N. Louis *et al.*, “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary,” *Acta Neuropathol.*, vol. 131, no. 6, pp. 803–820, Jun 2016.
- [38] R. Stupp *et al.*, “Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma,” *New England Journal of Medicine*, vol. 352, no. 10, pp. 987–996, Mar. 2005.
- [39] C. Hadjipanayis, G. Widhalm, and W. Stummer, “What is the Surgical Benefit of Utilizing 5-Aminolevulinic Acid for Fluorescence-Guided Surgery of Malignant Gliomas?” *Neurosurgery*, vol. 77, no. 5, pp. 663–673, Nov 2015.
- [40] M. Jaber *et al.*, “Is Visible Aminolevulinic Acid-Induced Fluorescence an Independent Biomarker for Prognosis in Histologically Confirmed (World Health Organization 2016) Low-Grade Gliomas?” *Neurosurgery*, vol. 84, no. 6, pp. 1214–1224, Jun 2019.
- [41] C. Jonin *et al.*, “Two photon excited fluorescence and hyper rayleigh scattering of protoporphyrin ix,” *Journal of Photochemistry and Photobiology A: Chemistry*, vol. 402, p. 112812, 2020.
- [42] “Introduction to Fluorescence,” in *Principles of Fluorescence Spectroscopy*, J. R. Lakowicz, Ed. Boston, MA: Springer US, 2006, pp. 1–26.
- [43] T.-T. Lu and S.-H. Shiou, “Inverses of  $2 \times 2$  block matrices,” *Computers & Mathematics with Applications*, vol. 43, no. 1-2, pp. 119–129, Jan. 2002.
- [44] M. W. Berry *et al.*, “Algorithms and applications for approximate nonnegative matrix factorization,” *Computational Statistics & Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.
- [45] S. Sra and I. Dhillon, “Generalized nonnegative matrix approximations with bregman divergences,” *Advances in neural information processing systems*, vol. 18, 2005.
- [46] Z. Yang *et al.*, “Blind spectral unmixing based on sparse nonnegative matrix factorization,” *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1112–1125, 2011.
- [47] A. d. Juan, J. Jaumot, and R. Tauler, “Multivariate Curve Resolution (MCR). Solving the mixture analysis problem,” *Analytical Methods*, vol. 6, no. 14, pp. 4964–4976, Jun. 2014, publisher: The Royal Society of Chemistry.
- [48] A. de Juan and R. Tauler, “Multivariate Curve Resolution: 50 years addressing the mixture analysis problem – A review,” *Analytica Chimica Acta*, vol. 1145, pp. 59–78, Feb. 2021.
- [49] O. Gutierrez-Navarro *et al.*, “Blind End-Member and Abundance Extraction for Multispectral Fluorescence Lifetime Imaging Microscopy Data,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 606–617, Mar. 2014, conference Name: IEEE Journal of Biomedical and Health Informatics.
- [50] J. Jaumot, A. de Juan, and R. Tauler, “MCR-ALS GUI 2.0: New features and applications,” *Chemometrics and Intelligent Laboratory Systems*, vol. 140, pp. 1–12, Jan. 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169743914002111>
- [51] T. B. Melø and G. Reisaeter, “The physicochemical state of protoporphyrin IX in aqueous solution investigated by fluorescence and light scattering,” *Biophys. Chem.*, vol. 25, no. 1, pp. 99–104, Nov 1986.
- [52] E. A. Porta, “Pigments in Aging: An Overview,” *Annals of the New York Academy of Sciences*, vol. 959, no. 1, pp. 57–65, 2002.
- [53] R. D. Jolly, D. N. Palmer, and R. R. Dalefield, “The analytical approach to the nature of lipofuscin (age pigment),” *Archives of Gerontology and Geriatrics*, vol. 34, no. 3, pp. 205–217, Jun. 2002.
- [54] A. C. Croce *et al.*, “Fatty liver oxidative events monitored by autofluorescence optical diagnosis: Comparison between subnormothermic machine perfusion and conventional cold storage preservation,” *Hepatology Research: The Official Journal of the Japan Society of Hepatology*, vol. 47, no. 7, pp. 668–682, Jun. 2017.
- [55] J. S. Dysart and M. S. Patterson, “Photobleaching kinetics, photoproduct formation, and dose estimation during ALA induced PpIX PDT of MLL cells under well oxygenated and hypoxic conditions,” *Photochem. Photobiol. Sci.*, vol. 5, no. 1, pp. 73–81, Jan 2006.
- [56] “AutoClass: A Bayesian Classification System,” in *Machine Learning Proceedings 1988*. Morgan Kaufmann, Jan 1988, pp. 54–64.