



**HAL**  
open science

# In vivo fluorescence spectra unmixing and autofluorescence removal by sparse Non-negative Matrix Factorization

Anne-Sophie Montcuquet, Lionel Herve, Fabrice Navarro, Jean-Marc Dinten,  
Jerome I. Mars

► **To cite this version:**

Anne-Sophie Montcuquet, Lionel Herve, Fabrice Navarro, Jean-Marc Dinten, Jerome I. Mars. In vivo fluorescence spectra unmixing and autofluorescence removal by sparse Non-negative Matrix Factorization. IEEE Transactions on Medical Imaging, 2011, 58 (9), pp.2554-2565. 10.1109/TBME.2011.2159382 . hal-00625349

**HAL Id: hal-00625349**

**<https://hal.science/hal-00625349v1>**

Submitted on 21 Sep 2011

**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.

# *In vivo* fluorescence spectra unmixing and autofluorescence removal by sparse Non-negative Matrix Factorization

Anne-Sophie Montcuquet<sup>1,2</sup>, Lionel Hervé<sup>1</sup>, Fabrice Navarro<sup>1</sup>, Jean-Marc Dinten<sup>1</sup> and Jérôme I. Mars<sup>2</sup>

**Abstract**—Fluorescence imaging locates fluorescent markers that specifically bind to targets, as tumors: markers are injected to a patient, optimally excited with near infrared light, and located thanks to emitted back fluorescence analysis. To investigate thick and diffusive media, as the fluorescence signal decreases with the light travel distance, the autofluorescence of biological tissues comes to be a limiting factor. To remove autofluorescence and isolate specific fluorescence, a spectroscopic approach, based on Non-negative Matrix Factorization (NMF), is explored. To improve results on spatially sparse markers detection, we suggest a new constrained NMF algorithm which takes sparsity constraints into account. A comparative study between both algorithms is proposed on simulated and *in vivo* data.

**Index Terms**—Source separation, fluorescence optical imaging, autofluorescence.

## I. INTRODUCTION

Fluorescence spectroscopy and Fluorescence diffuse optical tomography (FDOT) are promising non-invasive, user-friendly and cost effective methods to serve medical diagnostic systems. Fluorescent compounds, called markers, are injected to a patient, and specifically bind to targeted tumors [1], [2], [3], [4]. This method of targeting allows researchers to detect and localize cancer cells in patients. Once injected, markers are excited by near infrared (NIR) light – between 600 and 900 nm – where the tissue absorption is lower. The excitation wavelength is chosen to ease the biological tissue penetration, and to optimally excite the injected markers. The fluorescence signal emitted back by markers is measured, and the cancer cells may then be localized. So far, NIR fluorescence imaging is mainly used on small animals where some markers are available for injection, and where the layer of biological tissues to explore does not exceed a few centimeters. In such case, a biological tissues intrinsic fluorescence - called *autofluorescence* [5] - exists, but is insignificant compared to the fluorescent marker signal and does not prevent localization of the cancer cells. For medical diagnostic application on thick media (human prostate, or breast for example, that is around 4 cm), the autofluorescence is worth considering: as the fluorescence signal gets exponentially weak with the light travel distance, the autofluorescence signal remains constant and turns into a limiting factor. The analysis of fluorescence

signals impaired by autofluorescence may lead to a wrong localization of the markers and consequently to a wrong localization of cancer cells: the signal needs to be preprocessed in order to remove autofluorescence.

Several methods to unmix fluorescence spectra and to filter autofluorescence have already been developed and tested on small animal examination equipments. Among those methods are the non linear least squares [6], principal component analysis (PCA), independent component analysis (ICA) and singular value decomposition (SVD) [7], [8], [9], [10] methods. The ICA method requires sources to be statistically independent, SVD considers orthogonal sources, and many other *a priori* knowledges about the nature of the sources are taken into account in those methods. But a principal one remains missing: non-negativity [11]. Many real-world data are non-negative, among are the fluorescence data, and the unmixed fluorescence spectra provided by a separation source method have a physical meaning only when non-negative. In light of this observation, some studies began to consider Non-negative Matrix Factorization (NMF) to separate fluorescence spectra.

NMF was notably investigated by Paaetero and Tapper, and gained popularity in 2001, through the works of Lee and Seung [12]. In 2004, Gobinet used the NMF decomposition on spectroscopic data to unmix several pure fluorescence spectra and get a fluorescence chemical mapping of wheat grain sections [13]; in 2007, they used NMF to dewax Raman signals of human skin biopsies [14]. NMF was also applied to fluorescence spectroscopy, to unmix several markers by Xu *et al.*[15],[16]. Their work is the closest one from our research, but is still applied only to small animal *in vivo imaging*, while we try to focus on models and tumor/healthy skin ratios close to deep embedded markers problems. For *in vivo* fluorescence spectroscopy, the unmixing problem is referred to as a blind source separation problem since the spectra may vary according to the fluorescent dye biological environment. Fluorescence spectra to separate are also supposed statistically dependent, which filters out many methods (such as ICA). Finally, the NMF algorithm seems to be, in many ways, more suitable in blind positive spectra separation than other separation methods. Several NMF algorithms exist, based on diverse criteria to minimize and optimization methods: we present in the paper a classical NMF algorithm that deals with multiplicative update rules [12]. But as all blind source separation methods, it is impossible to find a unique NMF decomposition [17]. Without any constraint, there

<sup>1</sup>CEA, LETI, MINATEC, 17 rue des martyrs, 38054 Grenoble cedex 9, France.

<sup>2</sup>GIPSA-Lab, Rue de la houille blanche BP46, 38402 Saint Martin d'Hères, France.

is an infinity of decomposition solutions. By considering only non-negative solutions, which is inherent to the NMF method, we already restrain the solution set. To relax more the non-uniqueness issue and to refine the solution set, we suggest to apply appropriate regularizations and *a priori* knowledge considerations. Such additional constraints usually deal with sparsity and smoothness constraints [18], [19]. Kim and Park got for example interested in sparse NMFs algorithms [20] by  $L_1$ -norm constraint term minimization, Cichocki et al. have presented cost functions based not anymore on the Kullback-Leibler divergence but on Csiszár's  $\varphi$ -divergence [21] while other approaches use alternative cost functions formulations [22], [23]. For fluorescence imaging application, we propose a regularized NMF algorithm that takes spatial sparsity constraints into account to improve NMF decomposition and tested on spectroscopic simulated and *in vivo* data.

In this article, we first introduce the NMF method, and one of the most popular associated algorithm, proposed by Lee and Seung in 2001[12]. From that classical algorithm, to specifically detect spatially sparse fluorescent sources, we defined a new constrained algorithm with sparsity constraints. A comparative study is then run in a second part to compare NMF results obtained with or without taking sparsity constraints into account. To illustrate that study, a simulated example on which several tests are run is proposed. Finally, we present experimental *in vivo* data on the last part. Up to three fluorescent sources with overlapping emission spectra have to be unmixed: two fluorescent markers – ICG loaded into nanoparticules (ICG-LNP) and Alexa 750 – plus the autofluorescence signal. Once more, unmixing results obtained with the NMF algorithm and the sparse NMF algorithm are compared.

## II. THEORY

### A. Non-Negative Matrix Factorization

For initial non-negative mixed data  $M$ , NMF proposes to find a couple of matrices  $(A, S)$  with non-negatives coefficients, whose product optimally approaches  $M$ . The classical NMF definition says[12]:

Given a non-negative matrix  $M \in \mathbb{R}^{X \times Y}$ , find non-negative matrices  $A \in \mathbb{R}^{X \times P}$  and  $S \in \mathbb{R}^{P \times Y}$  such that :

$$M \simeq AS \quad (1)$$

where non-negative matrices are matrices whose all factors are nonnegative and  $P$  stands for the number of sources to unmix.

Applied to spectroscopy, matrix  $A$  is considered as the weights matrix, and  $S$  as the spectra matrix; both matrices  $A$  and  $S$  contain respectively as much columns and lines  $P$  as fluorescent sources to separate.

To find the best matrices  $A$  and  $S$  that satisfy equation 1, two distinct steps must be considered. First, a criterion

that links  $M$  with  $A$  and  $S$  has to be defined. Then, in a second step, the criterion has to be optimized. Optimizing this criterion – generally by minimizing it – under the non-negativity constraint would lead to the best couple of solutions  $(A, S)$ . Several criteria (the Euclidean distance, the Kullback-Leibler divergence [12]...) and different optimization methods (Alternating Least Square, multiplicative update rules) may suit to NMF method.

### B. Sparsity constraints

Besides the two basic steps – criterion choice and optimization – that lead to the NMF decomposition, additional constraints may be considered. Those constraints may directly modify criterion to optimize, or would imply a new optimization step.

Several studies got interested in sparsity constraints, applied to NMF. Among them we should name Hoyer *et al.* [24] and Stadlthanner *et al.* [25] who defined new sparse NMF algorithms. Both research groups got interested in sparsity of matrix  $S$ , while no assumption was made on  $A$ . But we can easily extend these algorithms in order to constrain matrix  $A$ . A difference remains between both algorithms: as Hoyer constrained all rows of  $S$  to have a common sparsity value, Stadlthanner extended Hoyer's algorithm by providing a different sparsity coefficient to each row of matrix  $S$ . This change makes the extended sparse NMF (esNMF [25]) algorithm more suitable for blind source separation (BSS) problems, where sources may have different sparsenesses. We implemented these algorithms in order to test them on our data. In our case, we got interested in sparseness of matrix  $A$ , and considered that sparseness of all columns of  $A$  could differ from a source to another. Sparse columns of  $A$  in spectroscopy is the expression of spatially sparse fluorescence signals. Indeed, for local specific markers distributions, we expect to get peaked and sparse weight columns in matrix  $A$ . *A contrario*, when unmixing is not complete, a residual autofluorescence baseline surrounds the specific marker peak. This property is depicted Figure 1.

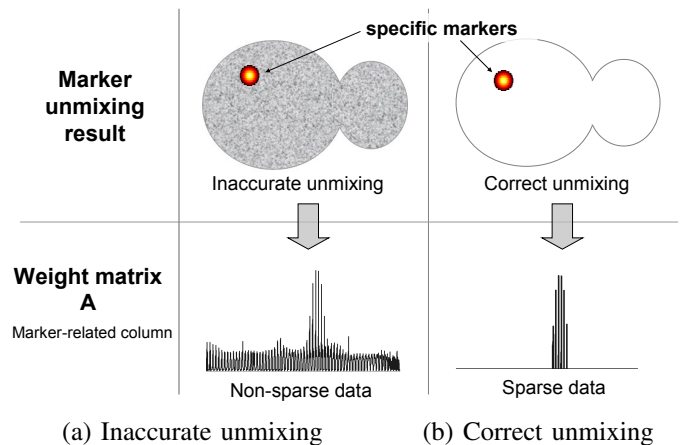


Fig. 1. Link between quality of unmixing result and weight matrix  $A$  obtained.

An intuitive reasoning to improve unmixing would be to smooth over the unwanted autofluorescence residuals on spe-

cific markers weight profiles, by thresholding the smallest values: in other words, we look for sparse  $A$  columns for specific markers contributions.

We introduce a sparsity value [24]: let us consider a weight matrix  $A$  of size  $X \times P$ , the sparsity of a given column  $A_p$  ( $p \in (1, P)$ ) of  $A$  is:

$$\text{sparsity}(A_p) = \frac{\sqrt{X} - \left( \sum_{x=1}^X |a_{xp}| / \sqrt{\sum_{x=1}^X a_{xp}^2} \right)}{\sqrt{X} - 1} \quad (2)$$

Sparsity value ranges from 0 for non-sparse results to 1 for extremely sparse results, as depicted Figure 2.

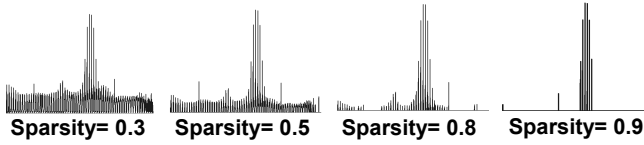


Fig. 2. Sparsity value ranges from 0 to 1, from non-sparse to highly sparse signals: example on weights vectors.

In the next section, we propose an NMF algorithm that finds the perfect threshold so that our matrix  $A$  sparsity remains as close as possible to a chosen sparsity value.

### III. MATERIAL AND METHODS

#### A. Simulated example

In order to test our unmixing method, we define a simulated example to run the NMF algorithms on. We propose to build a simulated phantom, composed of an autofluorescence part and of a specific fluorescence one (to compare to a tumor pointed out by an injected fluorescent marker). The specific fluorescence part is the product of a weight vector  $A_1$  by a fluorescence spectrum  $S_1$  (see Figure 3 - a). *A pari*, the autofluorescence part is the product of the weight vector  $A_2$  by the fluorescence spectrum  $S_2$  (see Figure 3 - b). Finally the total simulated phantom is obtained by adding the specific fluorescence and the autofluorescence parts (Figure 3 - c). In fluorescence optical imaging, signal ratio between healthy tissue and tagged tumor depends on biomarkers injected.

From bibliography, and from experience, ratios from 3 to 15 [26] (for more specific-to-tumor markers) are usual. New generation of activatable fluorescent markers reach ratios from 24 to 180 [27], [28], [29], depending on wavelength range observed, and conditions of experimentation: *ex vivo* or *in vivo*, and localization of tumor. For this simulation example, to reproduce deep embedded markers detection, we chose a ratio tumor/healthy tissue with highest intensity value approximately equal to 1.

#### B. Contrast definition

In order to evaluate the tumor detection, we introduce the contrast  $C_{T,N}$  which is measured between a tumorous area  $T$  and a normal (or healthy) tissues area  $N$ . The value of  $C_{T,N}$  characterizes the tumor detection after autofluorescence removal, on simulation and experimental results. Average

intensity of fluorescence signal is measured on both concerned ROIs:  $\bar{T}$  and  $\bar{N}$  are respectively the average intensities in photons per pixel of areas  $T$  and  $N$  on analyzed images (see Figure 3-c).

$$c_{T,N} = \frac{\bar{T} - \bar{N}}{\bar{T} + \bar{N}} \quad (3)$$

The closer to one the contrast value gets, the better the detection will be.

#### C. Classical NMF algorithm

In 2001, NMF popularity increased after Lee and Seung published two new NMF algorithm, based on the use of multiplicative update rules that minimize specific criteria.

The NMF decomposition looks for the best solution couple  $(A, S)$  whose product best approaches the initial data  $V$ . Classically, to find matrices  $A$  and  $S$ , a chosen criteria is iteratively minimized.

Different criteria to minimize, or *cost functions*, can be used: we may cite the square of the euclidean distance between  $V$  and  $AS$ , and the Kullback-Leibler divergence criterion [12]. Here, we define the cost function  $F$  to minimize as the square of the Euclidean distance between  $V$  and  $AS$  [12], lower bounded by 0:

$$F = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \left( v_{xy} - \sum_{p=1}^P a_{xp} s_{py} \right)^2 = \|V - AS\|_2^2 \quad (4)$$

Here,  $\|\cdot\|_2$  denotes the standard Euclidean norm.

In order to get solutions  $(A, S)$ , the following *optimization* problem is thus considered:

**Problem 1:** Find couple  $(A, S)$  such as  $(A, S) = \underset{(A,S) \geq 0}{\operatorname{argmin}} \|V - AS\|_2^2$

A classical gradient descent method may be used to solve Problem 1. Nevertheless, multiplicative update rules that minimize  $F$  have been developed [12]. They offer a good compromise between speed and ease of implementation to solve Problem 1:

**Theorem 1** *The distance  $\|V - AS\|_2^2$  is non increasing under the update rules:*

$$S_{xp} \leftarrow S_{xp} \frac{(A^t V)_{xp}}{(A^t AS)_{xp}} \quad A_{py} \leftarrow A_{py} \frac{(V S^t)_{py}}{(AS S^t)_{py}} \quad (5)$$

The proof of this theorem is given in Lee and Seung's publication [12].

We precisely got interested in those update rules because of their ease of implementation and speed, for which they were initially created. Moreover, for initial non-negative matrices  $A$  and  $S$ , the constraint of non-negativity is inherent to the method: thanks to multiplicative rules, results remain non-negative all along the iterations.

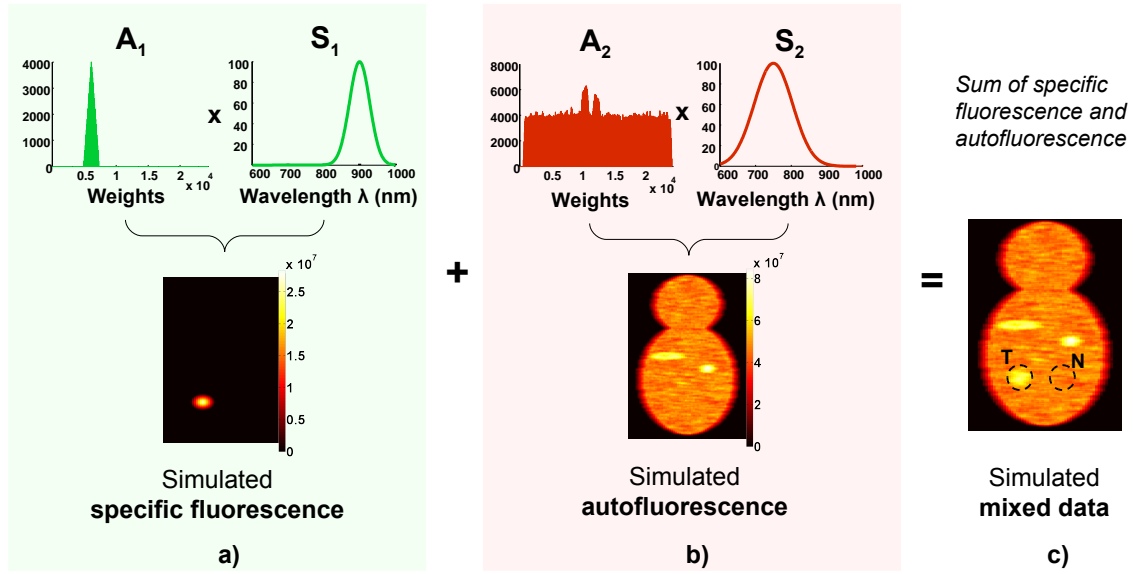


Fig. 3. By summing of a specific fluorescence signal a) with an associate fluorescence spectrum and an autofluorescence signal b) with a different fluorescence spectrum, we obtain simulated mixed data c).

The cost function definition, followed by the optimization step compose the classical NMF algorithm; an intrinsic initialization step chose non-negative matrices to start the algorithm [30]. In our case, since a few information is available on fluorescence spectra, and to restrain the solution set, we chose to initialize matrix  $S$  with calibration spectra.

The NMF algorithm thus defined is described in this section:

---

**Algorithm: NMF**

For  $i=1$  to  $N_{iter}$

- 1) Initialize  $A$  and  $S$  with respectively positive constants and positive spectra models
  - 2) Update  $S$ :  $S \leftarrow S \frac{(A^t V)}{(A^t A S)}$
  - 3) Update  $A$ :  $A \leftarrow A \frac{(V S^t)}{(A S S^t)}$
  - 4) restart steps 2 and 3 until stopping criterion is obtained (for example when  $\|r_{i-1} - r_i\|^2 < \epsilon$ , for  $r_i = \|V - AS\|^2$  and for a chosen  $\epsilon$ ).
- 

**D. Constrained NMF**

For our fluorescence imaging problem, since columns  $A_p$  of  $A$  that refer to the fluorescent markers distribution in the medium are expected to be sparse, bring an *a priori* information on matrix  $A$  would help restrain the NMF solution set, and attenuate the non-uniqueness ambiguity. We implemented a sparse NMF algorithm that would help promoting sparsity of matrix  $A$  thanks to a thresholding step which at each iteration of algorithm after the optimization part that solves Problem 1 would favor sparse updates. Indeed both minimizing  $F$  and dealing with sparsity of columns  $A_p$  of matrix  $A$  is usually simultaneously resolved applying a thresholding on data  $A_p$  that brings a “sparsity promoting” behavior to the algorithm [31].

Working out an appropriate threshold value at each iteration comes to find a number which can be comprised between 0 and the maximum of  $A_p$  values. In order to relax the thresholding value choice, we prefer to select an initial sparsity value for concerned vector  $A_p$  thus comprised between 0 and 1 as described in previous section; at each iteration our algorithm will find the best threshold value to eliminate background values and make sparsity of vector  $A_p$  remain close from the initial chosen sparsity value.

The last iteration will end with a classical updating of matrices  $A$  and  $S$  (without thresholding step) in order to balance the fluorescence intensities between all fluorescence sources.

In this section, taking the classical algorithm as a basis, we propose to describe our sparse NMF algorithm implementation. The first NMF algorithm steps do not change from the classical one, but a sparsity step is added for each iteration in order to remain close to a wanted sparsity value  $\varphi$  defined by the user:

---

**Algorithm: NMF with sparsity constraints**

For  $i=1$  to  $N_{iter}$

- 1) Initialize  $A$  and  $S$  with respectively positive constants and positive spectra models
- 2) Update  $S$ :  $S \leftarrow S \frac{(A^t V)}{(A^t A S)}$
- 3) Update  $A$ :  $A \leftarrow A \frac{(V S^t)}{(A S S^t)}$
- 4) **Each column  $A_p$  of  $A$  referring to weights of specific markers**, for a wanted sparsity value  $\varphi$  for coefficients of column  $A_p$ , is changed into  $\tilde{A}_p$ :

$\forall x \in (1, X)$ ,

$$\tilde{a}_{xp} = \begin{cases} a_{xp} & \text{if } a_{xp} \geq \frac{\max(a_{xp})}{\beta_{pmin}} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

with  $\beta_{pmin}$  the threshold value such as:

$$\beta_{pmin} = \underset{\beta_p}{\operatorname{argmin}} \left| \operatorname{sparsity} \left( \tilde{A}_p(\beta_p) \right) - \varphi \right|$$

- 5) restart steps 2 to 4 until stopping criterion is obtained (for example when  $\|r_{i-1} - r_i\|^2 < \epsilon$ , for  $r_i = \|V - AS\|^2$  and for a chosen  $\epsilon$ ).
- 6) if  $i = N_{iter}$  then run steps 2 and 3 before to end the algorithm.

As the previous algorithm is implemented with Matlab, we use the `fminsearch` function (Matlab) to solve the `argmin` problem. `Fminsearch` attempts to find a minimum of function  $\left| \operatorname{sparsity} \left( \tilde{A}_p(k) \right) - \varphi \right|$  where variable  $k$  is unknown, starting at an initial estimate of  $k$ . The threshold  $\beta_{pmin}$  which is selected is the value from all  $\beta_p$  tested that returns the minimum of the tested function.

Let us consider that the vector  $A_p$  has, at first iteration of the algorithm, a sparsity value which is smaller (for example  $\operatorname{sparsity}(A_p)=0.2$ ) than the chosen initial sparsity value  $\varphi$  (for example  $\varphi = 0.6$ ). Then in step (4) of the algorithm, the `fminsearch` function will select an optimal threshold value  $\beta_{pmin}$  from all  $\beta_p$  tested: when this threshold is applied on  $A_p$  data, all values in vector  $A_p$  that are less than the threshold  $\beta_{pmin}$  are set to zero, and new vector  $\tilde{A}_p$  will be sparser than before. In that step of the algorithm, the threshold is chosen so that the value  $\left| \operatorname{sparsity} \left( \tilde{A}_p(k) \right) - \varphi \right|$  is minimized, but the threshold does not assure that the obtained sparsity value of  $A_p$  is equal to  $\varphi$ .

*A contrario* if vector  $A_p$  has a sparsity value greater (for example  $\operatorname{sparsity}(A_p)=0.7$ ) than the chosen initial sparsity value  $\varphi$  (for example  $\varphi=0.6$ ), then the obtained threshold  $\beta_{pmin}$  at the end of step (4) of the algorithm will be equal to 0, since  $A_p$  is already sparser than expected.

Then in both cases, steps 2 and 3 of the algorithm are restarted, and matrices  $A$  and  $S$  are updated thanks to the multiplicative update rules that minimize this time the cost function  $\|V - AS\|^2$ . At that point, the sparsity value of vector  $A_p$  may have changed since  $A$  has been updated, and sparsity value could thus exceed the chosen initial sparsity value  $\varphi$ . In that case, threshold obtained at further iteration would be set to 0, as explained previously. Finally by repeating steps 2 to 4 in turns, the cost function is minimized with sparse solutions being favoured, but priority is still given to the cost function minimization.

*Sparsity value selection:* The constrained NMF algorithm asks for an initial sparsity value chosen by the user. When fluorescence sources distributions in media are sparse, looking for a sparse column  $A_p$  is appropriate; this *a priori* information helps restraining the solutions set. We empirically noticed that

a large choice of initial sparsity values, even far from the true expected sparsity value, help improving the classical NMF results. We propose in next section a set of simulation results that underline this idea.

For now no automatic selection of initial sparsity value has been implemented, and we empirically choose this parameter. NMF algorithm being quick to converge (a few seconds to one minute for a whole body mouse study), manual selection of this initial sparsity value does not prevent fast data processing.

#### IV. NMF VERSUS SPARSE NMF

In this section, we propose a comparative study between the NMF algorithm without spatial sparsity constraints and our new constrained algorithm that looks for spatially sparse solutions for the specific markers fluorescence.

##### A. Influence of sparsity value choice

A first comparison is made between both algorithms by simply running them on our simulated data (see Figure 4-a) where sparsity of simulated vector  $A_1$  is equal to 0.9017. Initialization for matrix  $S$  was the same for both tries, Gaussian were chosen, but lightly translated (50 nm) compared to true expected spectra. We test 11 initial sparsity values for the sparse NMF algorithm, from 0 to 1 to examine influence of sparsity initialization on results. Results of matrices  $A$  and  $S$  obtained after unmixing are presented Figure 4-b: the NMF algorithm without sparsity constraints did not manage to correctly unmix both sources. Taking sparsity constraints into account, here with an initial sparsity value equal to 0.8, led to an accurate result (see Figure 4-c). The imposed sparsity value of 0.8 finally reached 0.9027 at convergence of sparse algorithm when true sparsity value is 0.9017. We can also underline that a wrong initialization prevented the classical NMF algorithm to converge to the true spectra, but did not influence the sparse NMF algorithm. We propose in next part to study more carefully both algorithms' robustness to initialization.

Finally Figure 5 proposes sparsity values of  $A_1$  and contrast between tumoral and healthy tissues obtained after unmixing for all initial sparsity values tested. We note on graph 5-a that thresholding steps of sparse NMF can only improve the classical NMF results. Only a few values of initial sparsity from 0 to 0.2 would not make any change on unmixing results compared to classical NMF performances (sparsity of unmixed vector  $A_1$  is then equal to 0.2381). But for a large range of initial sparsity values, from 0.2 to 1, sparse NMF considerably improved classical NMF unmixing. In that range, initial sparsity values to bind to were at the origin of beneficial thresholding steps that helped the algorithm to converge to a sparse solution closer to the simulated  $A_1$ .

This result is comforted by detection results presented on graph 5-b. Again for a range of initial sparsity values from 0.2 to 1, contrast  $C_{T,N}$  between tumoral and healthy tissue areas is much closer to 1 than the contrast equal to 0.68 obtained with the classical NMF algorithm. The detection of the tumor is thus sensibly improved by the sparse NMF algorithm processing.

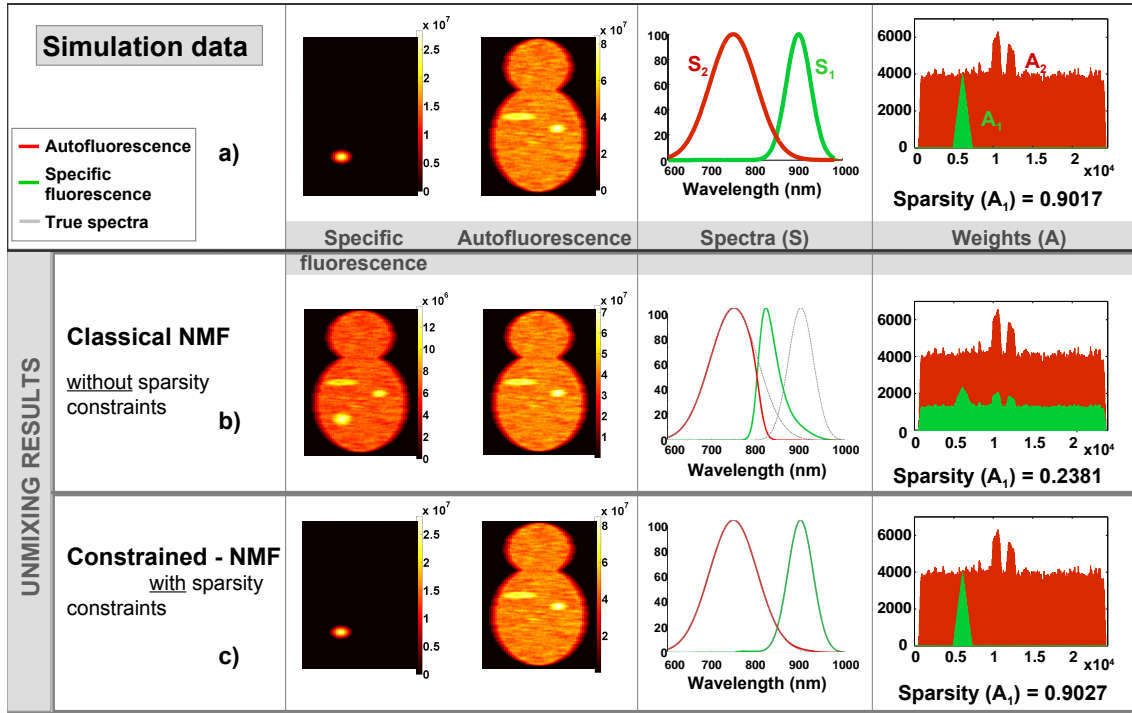


Fig. 4. First comparison of sparse and non-sparse NMF algorithms: a) simulated data, b) Results obtained with the classical NMF algorithm, c) Results with our sparse NMF algorithm (for an initial sparsity value equal to 0.8). Taking sparsity constraints into account leads to more accurate results.

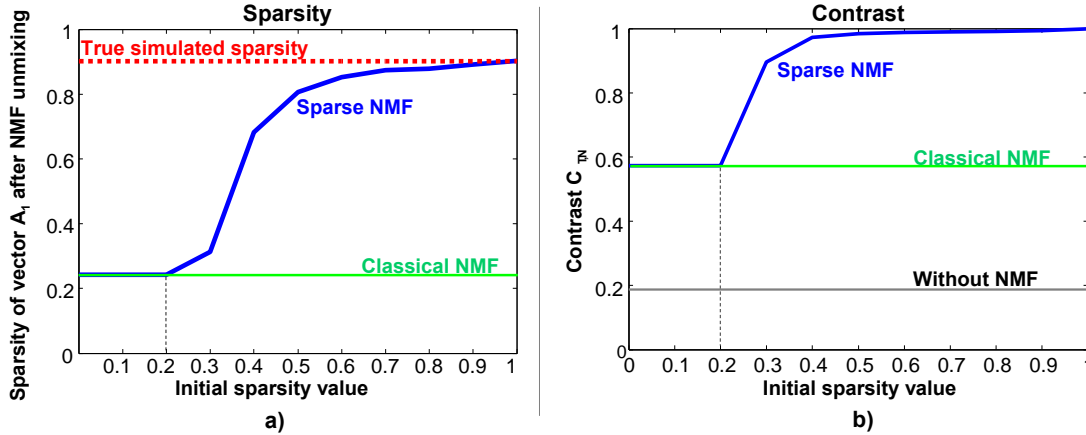


Fig. 5. a) Sparsity values of  $A_1$ , and b) contrast  $C_{T,N}$  between tumoral and healthy tissues obtained after unmixing for 11 initial sparsity values for  $A_1$  tested between 0 and 1.

### B. Comparison with other sparse NMF algorithms

As mentioned previously, several sparse NMF algorithms already exist; in this section we propose to test the algorithm developed by Hoyer [24] on the simulated data already presented in the previous paragraph. Hoyer's algorithm proposes a projected gradient descent algorithm for NMF with sparseness constraints that uses a projection operator to enforce a desired degree of sparseness. The main difference with our algorithm lies in that last property: as we propose a thresholding step to help the algorithm to converge to a more accurate solution, Hoyer's method forces the solution of matrix  $A$  (or  $S$  in his article) to reach a fixed sparsity value. The sparsity value choice is thus crucial if we want to get a pertinent solution while the initial sparsity value we chose in our algorithm does

not necessary has to be reached.

We implemented Hoyer's algorithm with the possibility to impose different sparsity values for distinct columns of  $A$  as suggested by Stadlthanner in [25]. We ran it on previous simulated data (cf. Figure 4-a). We first studied sparsity of  $A_1$  obtained after NMF, depending on the initial sparsity value chosen (see Figure 6-a) and compared it to our results: as expected, sparsity of vector  $A_1$  is equal to the desired sparseness at convergence. Thresholding steps are more adapted to our problem since a larger range of initial sparsity values lead to pertinent results. This is especially useful when the size of the tumor can not be estimated and that the initial sparsity value has to be empirically chosen. On Figure 6-b we then compared detection results obtained with Hoyer's algorithm or

ours on that same example by plotting contrast  $C_{T,N}$  obtained between tumoral and healthy areas after unmixing. Contrast obtained on that example (except for a unique initialization of the sparsity value at 0.1) is always greater than contrast found with Hoyers algorithm.

We compared both methods on simulated and *in vivo* data, and the thresholding method chosen in our algorithm is more robust to a larger choice of initial sparsity values than methods implemented by Hoyer and Stadlthanner.

### C. Robustness to initialization

We want to test the robustness to initialization of our sparse NMF algorithm and to compare it with the classical NMF one. By making the initialized matrix  $S$  vary, the algorithm may lead to different solutions for the NMF decomposition. The solutions are indeed very sensitive to the initialization [30]. On our simulation example, for a matrix  $S$  initialized exactly on the simulated spectra, the solutions obtained by both algorithms were the exact ones. We propose in this section to test the robustness to the initialization of both algorithms we compare, with or without sparsity constraints, on our simulated data. To create different initialization spectra for matrix  $S$ , we simply translate spectra we used for simulation (see Figure 7-a) on a chosen wavelength range, from around 100 nm lower from true spectra emission peaks, to 100 nm upper: obtained range of  $N$  initialization spectra for matrix  $S$  is presented on Figure 7-b.

We then run both NMF algorithms, with or without sparsity constraints,  $N$  times for each initialization spectra tested. Obtained solution sets are presented Figure 8. Solution set obtained with the non-sparse algorithm presents various solutions depending on translated initializations (see Figure 8 - a) while the algorithm that takes sparsity constraints into account leads to a more robust solution set, even with a wide range of initializations.

The sparse NMF algorithm appears to be much more robust to the choice of the initialization than the classical NMF algorithm: the range of solutions for matrix  $S$  obtained when taking sparsity constraints into account gives the right positions for the fluorescence peaks when the classical algorithm presents wrong solutions. Only the initializations the closest from the true expected spectra allow the classical NMF algorithm to converge to right spectra. *A contrario*, the sparse NMF algorithm is robust enough to give results all almost superimposed to the true spectra. In next part, a more detailed study is made on solution set for both algorithms.

### D. Range of admissible solutions to the NMF equation

As explained before, the NMF decomposition is not unique: without constraints, an infinity of solutions is admissible. Impose non-negativity to NMF results restrains the solution set.

To compare once more the sparse NMF algorithm and the classical one (without sparsity constraints), we propose to adapt a uniqueness study conducted by Moussaoui *et al.* in 2005 [32] to our data. This study defines – from a particular solution of NMF decomposition – the range of admissible

solutions that lead to a same result. We briefly present the theory proposed by Moussaoui in the next section. We then propose to compare the ranges of admissible solutions on our simulated data obtained after NMF decompositions, with or without sparsity constraints.

1) *Problem statement*: Let us assume a factorization of  $V$  by the product of matrices  $A$  and  $S$  exists. If we now consider an invertible matrix  $T$  of size  $P \times P$ , then a new couple  $(\tilde{A}, \tilde{S})$  of solutions is easily found:

$$X = \underbrace{(AT^{-1})}_{\tilde{A}} \underbrace{(TS)}_{\tilde{S}} \quad (7)$$

There is an infinity of factorizations of matrix  $V$ . The solution range may nevertheless be restricted by the non-negative constraint imposed to  $\tilde{A}$  and  $\tilde{S}$ .

2) *Case for two sources*: If two fluorescence sources are considered, for an initial set of data  $V$  of size  $m \times n$ , matrices  $A$  and  $S$  respectively of size  $m \times 2$  and  $2 \times n$  are:

$$A = (a_1 \quad a_2) \quad \text{et} \quad S = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad (8)$$

Let us consider an invertible matrix  $T$ , of size  $2 \times 2$ :

$$T = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix} \quad (9)$$

To avoid indeterminacy problems between lines of  $S$  (and by default columns of  $A$ ), the condition  $\alpha + \beta < 1$  is imposed. Invert matrix of  $T$  is given by:

$$T^{-1} = \frac{1}{1 - \alpha - \beta} \begin{pmatrix} 1 - \beta & -\alpha \\ -\beta & 1 - \alpha \end{pmatrix} \quad (10)$$

Matrices  $\tilde{A} = A \times T^{-1}$  and  $\tilde{S} = T \times S$  may thus be expressed as:

$$\forall k, \tilde{S} = \begin{pmatrix} (1 - \alpha) \times s_{1k} + \alpha \times s_{2k} \\ \beta \times s_{1k} + (1 - \beta) \times s_{2k} \end{pmatrix} \geq 0 \quad (11)$$

$\forall \ell$ ,

$$\tilde{A} = \frac{1}{1 - \alpha - \beta} \times \begin{pmatrix} (1 - \beta) \times a_{\ell 1} - \beta \times a_{\ell 2} & -\alpha \times a_{\ell 1} + (1 - \alpha) \times a_{\ell 2} \end{pmatrix} \geq 0 \quad (12)$$

a) *Admissible solutions range*: Solve the preceding equations for each term of matrices  $\tilde{A}$  and  $\tilde{S}$  leads to the admissible intervals in which  $\alpha$  and  $\beta$  have to be included [32]; with  $B_1 = \{k; s_{1k} < s_{2k}\}$  and  $B_2 = \{k; s_{1k} > s_{2k}\}$  :

$$\max_{k \in B_1} \left\{ \frac{s_{1k}}{s_{1k} - s_{2k}} \right\} \leq \alpha \leq \min \left\{ \frac{a_{\ell 2}}{a_{\ell 1} + a_{\ell 2}} \right\} \quad (13)$$

$$\max_{k \in B_2} \left\{ \frac{-s_{2k}}{s_{1k} - s_{2k}} \right\} \leq \beta \leq \min \left\{ \frac{a_{\ell 1}}{a_{\ell 1} + a_{\ell 2}} \right\} \quad (14)$$

The ranges of admissible solutions for the NMF equation are constrained by the values of parameters  $\alpha$  and  $\beta$ .



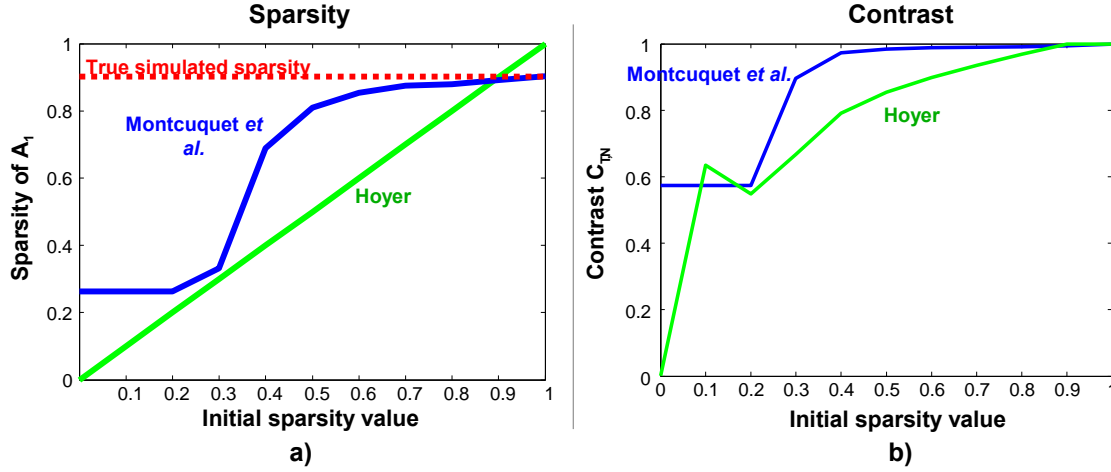


Fig. 6. Comparison between our sparse NMF algorithm and Hoyer's. a) Sparsity values of  $A_1$ , and b) contrast  $C_{T,N}$  between tumoral and healthy tissues obtained after unmixing for 11 initial sparsity values for  $A_1$  tested between 0 and 1 for both algorithms: Hoyer's and ours

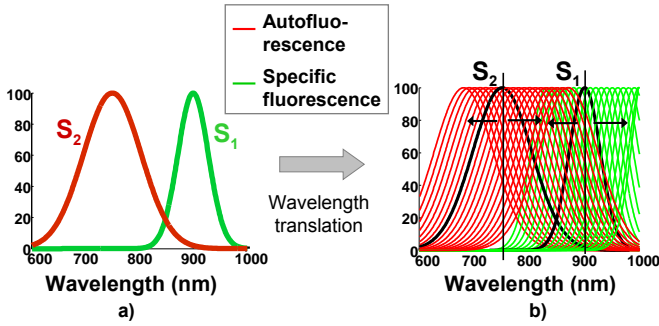


Fig. 7. the two true spectra  $S_1$  and  $S_2$  used to simulate data are translated on a wavelength range, from 100 nm lower from true spectra emission peaks, to 100 nm upper, to create a wide range of initialization spectra.

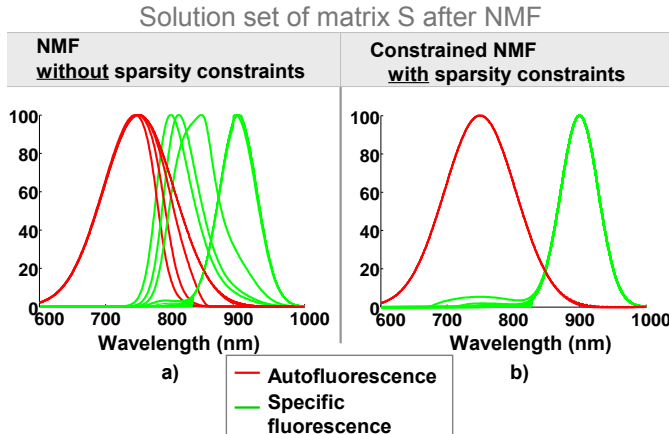


Fig. 8. Solution sets obtained for matrix  $S$  depending on the initialization of the algorithm, and depending on the tested algorithm: a) NMF without sparsity constraints, b) NMF with sparsity constraints.

b) *Application to simulated data:* We run both NMF algorithms (with or without sparsity constraints) on our simulated data. The initialization of  $S$  was perfectly chosen (that means equal to the expected spectra). We obtain a couple of solutions ( $A, S$ ) for both algorithms. From that couple, ranges

for parameters  $\alpha$  and  $\beta$  and matrices  $T$  are calculated, and ranges of solutions for  $S$  are plotted. We compare ranges of admissible solutions obtained, taking or not sparsity constraints into account; results are presented Figure 9, initial solution  $S$  is plotted in black, and true solution  $S$  in dotted black. Sparsity constraints lead to a more restricted range of solutions, close to true solution expected, and gives once more better results than classical NMF algorithm.

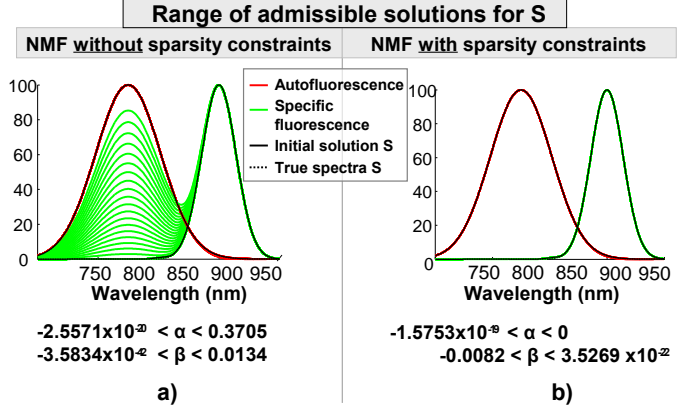


Fig. 9. (a) Range of admissible solutions with the classical NMF algorithm, without sparsity constraints (b) Range of admissible solutions, with the constrained NMF algorithm (sparsity constraints).

### E. Robustness to deep and multiple sites tumors detection

1) *Deep tumors detection:* To study detection of tumors deep embedded in tissues, we propose a simulation based on real breast data: many clinical experiments on this organ allows us to design a computer breast model with realistic optical properties of tissues. A simulated marked tumor is introduced to the model, and consistent modeling fluorescence acquisitions of the simulated breast are obtained with modified depth of the marked tumor. This simulation study was first introduced in [30]: the article details the optical parameters choice and the light propagation model.

On a specific example similar to previous ones, we simulate the intensity decrease of the fluorescence signal emitted by markers when they are moved in depth from 1 mm to 40 mm in tissues: results are presented Figure 10.

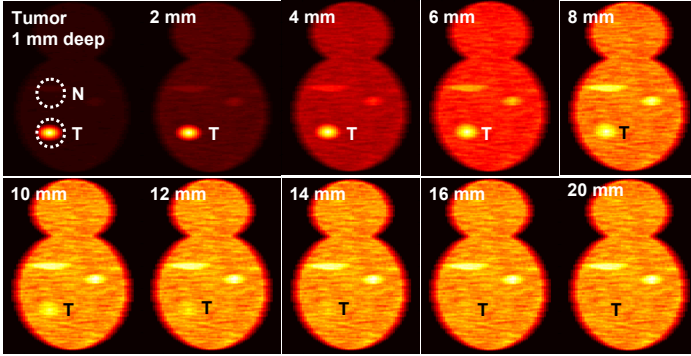


Fig. 10. A simulated tumor binded by fluorescent markers are moved in depth in breast tissues, from 1 mm to 40 mm. Specific signal emitted by tumor is quickly indistinguishable from autofluorescence signal (results of mixed signals from 20 mm to 40 mm are the same).

The signal of interest emitted by markers becomes quickly indistinguishable from autofluorescence (from around 20 mm deep). We run the classical NMF algorithm and our sparse NMF algorithm on the simulated data and calculate at each depth the contrast  $C_{T,N}$  between tumoral and healthy areas after unmixing. The initial sparsity value for our sparse NMF algorithm was chosen equal to 0.7, away from the true simulated sparsity of vector  $A_1$  (equal to 0.9017) on purpose. Results are presented Figure 11: when a contrast equal to 0.4 is reached for a tumor at 6 mm deep in tissues before unmixing (Figure 11-a), classical NMF processing allows to get a similar contrast for a tumor placed 12 mm deep (Figure 11-b). Finally even with no accurate sparsity initialization, contrast obtained with our sparse NMF algorithm is considerably improved compared to both prior cases (see Figure 11-c): tumor is detected with a contrast tumor/healthy tissue equal to 0.4 at 27 mm.

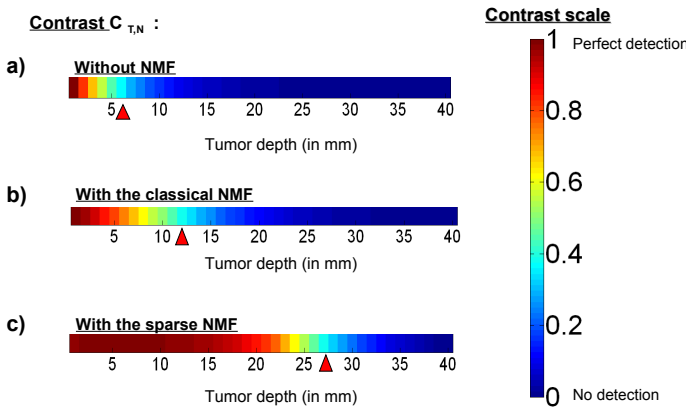


Fig. 11. Deep tumors detection results: a) contrast obtained before unmixing (contrast equal to 0.4 at 6 mm), b) contrast after classical NMF processing (contrast equal to 0.4 at 12 mm), c) contrast obtained with our sparse NMF algorithm (contrast equal to 0.4 at 27 mm).

Sparsity constraints and thresholding steps helped pushing

back the detection limits.

2) *Multiple sites tumors detection*: Tumors may also grow at multiple sites due to the metastasis, and the robustness of our NMF algorithm for multiple tumors detection should be examined. We propose on our previous example to simulate a group of 5 tumors; different intensity levels are chosen in order to mimic fluorescence signals emitted by tumors at several depths in tissues (see Figure 12-a and 12-b). Each tumor is associated with the same fluorescence emission spectra  $S_1$  as in previous examples. Depending of depth of tumor, spectra lightly vary (cf. [30]).

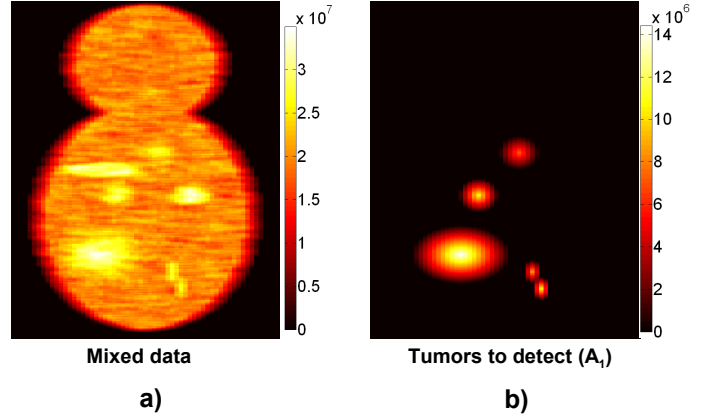


Fig. 12. Simulation of multiple sites tumors. a) Mixed data, b) 5 tumors to detect.

We run the sparse NMF algorithm with several initial sparsity values from 0 to 1: results of sparsity of vector  $A_1$  obtained after unmixing are presented Figure 13. Once more, for a large range of sparsity values, from 0.25 to 1, unmixing results are considerably improved compared to classical NMF results (see green line on Figure 13) even for multiple sites tumors. For smaller initial sparsity values (0 to 0.25), obtained results are similar to the classical NMF ones, and the thresholding step had no effect on data. Sparsity constraints could only improve NMF results but would not return less accurate unmixing results.

We propose in next section to test the sparse NMF algorithm on mice data.

## V. *in vivo* UNMIXING RESULTS

To test our algorithm this time on real data, an *in vivo* experiment is performed on a mouse. *In vivo* experiments imply that an autofluorescence signal is necessarily measured. We want to test the NMF algorithm on three sources unmixing: we thus chose two specific markers to separate from each other, and from autofluorescence signal.

### A. Feasibility experiment

The animal procedure was in compliance with the guidelines of the European Union (regulation n°86/609), taken in the French law (decree 87/848) regulating animal experimentation. All efforts were made to minimize animal suffering. The animal manipulation was performed with sterile

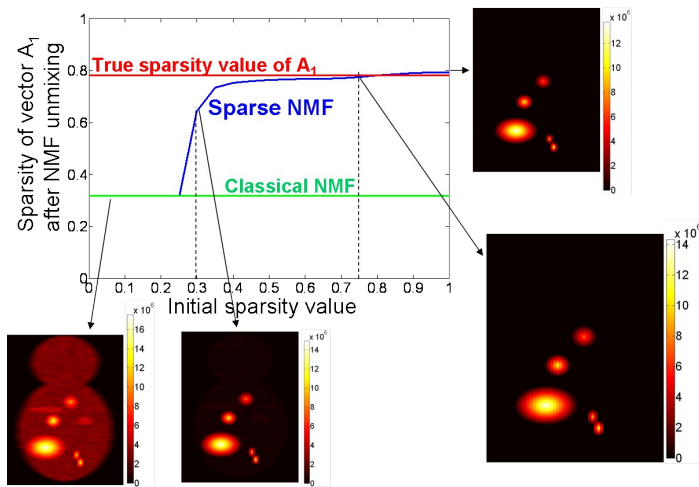


Fig. 13. Sparsity values of  $A_1$  obtained after sparse NMF unmixing, and associate tumors weights results (blue line). Comparison with true simulated sparsity of  $A_1$  (red line) and classical NMF results (green line).

techniques and approved by the Grenoble Animal Care and Use committee (France) (registration number 20\_iRTSV L  ti-FNG-02). An adult female nude mouse (Janvier, Le Genest saint-isle, France) was used throughout the experiments. It was housed in approved facilities, at  $21 \pm 1^\circ\text{C}$  under diurnal lighting conditions. The mouse arrived at the animal facility two weeks before the experiments start and had free access to food and water.

The animal is placed on a translation stage with  $N_y$  positions of travel range. To acquire spectrally resolved measurements and a whole scanning, the mouse is illuminated with a laser line at 690 nm, and the stage is translated  $N_y$  times. For each position of the stage, the emitted back fluorescence signal is collected along a line of  $N_x$  points by an imaging spectrometer coupled with a charge-coupled device camera (Andor Technologies): a  $N_x \times N_\lambda$  acquisition is measured (see Figure 14). At the end of the scan,  $N_y$  acquisitions are obtained, each of size  $N_x \times N_\lambda$ . For this experiment,  $N_x$  was equal to 255 and  $N_\lambda$  to 1024, which corresponds to a wavelength range around 590 to 973 nm. Before to run the NMF algorithm, mixed data of size  $(N_y \times N_x \times N_\lambda)$  are reordered as a 2-D array of size  $(N_y \times N_x, N_\lambda)$ .

To be able to quantify our results, especially the unmixed autofluorescence signal, a first acquisition is performed on the mouse, with no specific fluorescent markers: only the autofluorescence of the mouse tissues is detected. The resulting acquisition is presented Figure 15-a. Besides are acquired *ex vivo* spectra of both specific markers we use: Alexa 750 and Indocyanine Green loaded into Lipid Nanoparticles (ICG-LNP)[33] (see Figure 16).

To simulate marked tumors after circulating tumor-specific markers have been eliminated everywhere except in the tumor, two glass capillary filled with specific markers are inserted subcutaneously (see Figure 15-b)). Tubes are respectively filled with  $5 \mu\text{l}$  of ICG-LNP at  $0.35 \mu\text{mol/l}$  and  $5 \mu\text{l}$  of Alexa 750 at  $0.1 \mu\text{mol/l}$ . The three distinct fluorescent sources –

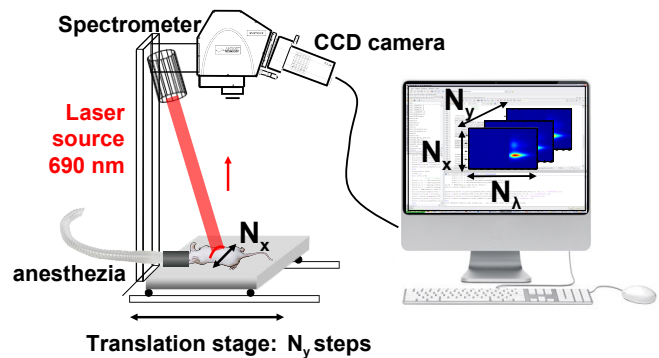


Fig. 14. Experimental set-up: acquisition on the animal and data processing.

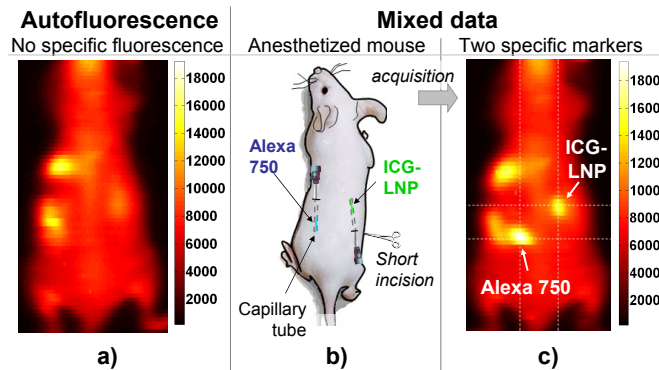


Fig. 15. a) autofluorescence acquisition b) Experiment: two capillary tubes filled with specific fluorescent markers are placed on the animal c) Acquisition with two specific markers tubes.

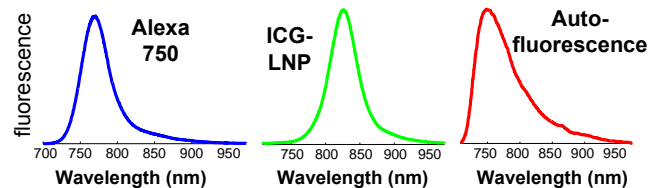


Fig. 16. Spectra of ICG-LNP and Alexa 750 acquired *ex vivo*, and average autofluorescence spectrum acquired on mice.

autofluorescence, ICG-LNP and Alexa 750 – have overlapping emission spectra.

Once the capillary tubes placed, the fluorescence acquisition of the animal is run. Obtained result is presented Figure 15-c.

## B. Results

We ran both algorithms we want to compare on our experimental data. For the sparse NMF algorithm, a sparsity value equal to 0.8 has been empirically chosen. The different unmixing results are presented Figure 17. Both algorithms succeeded to separate the three fluorescence sources: in the three fluorescence images obtained for each algorithm, the highest intensities correspond to true position of markers (see unmixed fluorescence contributions on Figure 17).

Nevertheless, sparse NMF algorithm gave more accurate results. This can first be noticed by comparison with the fluorescence contributions obtained with the classical algorithm,

particularly on ICG-LNP and autofluorescence ones: ICG-LNP contribution obtained with classical algorithm (see image 2 on 17 - a) contains background parts that were not accurately separated. Those background parts appear as missing on the autofluorescence contribution (see image 3 on Figure 17 - a, comparison can be made with the initial mixed data). Autofluorescence contribution can also directly be compared to the test acquisition without specific fluorescence (Figure 15-a). Finally, all the drawbacks underlined before, as wrong background remaining and not accurate separation, do not appear – or lesser – in the results obtained with the sparse NMF algorithm (see Figure 17-b, images 1,2 and 3).

Resulting  $A$  matrices translate all the observations previously made on images 1, 2 and 3. Indeed, the fluorescent markers weights (columns of  $A$ ) are expected to be spatially sparse to agree with the capillary tubes positions. Such result is reached with the sparse NMF algorithm (see weights of matrix  $A$  on Figure 17 - b)), while an incorrect background remains in columns of  $A$  for specific markers obtained with the classical NMF algorithm (Figure 17 - a)), as already noticed on images 1, 2 and 3 of Figure 17-a.

Moreover, we can look more in details at spectra obtained in matrix  $S$  with both algorithms: compared to emission spectra from our *ex vivo* and autofluorescence measurements (see Figure 16), specific fluorescence spectra obtained after classical NMF separation are moving further away from expected fluorescence spectra (see Figure 17-a, matrix  $S$ ) than those obtained with the sparse NMF algorithm (see Figure 17-b, matrix  $S$ ), that much closer to reality.

## VI. CONCLUSION

The specific fluorescence signal of specific markers used in optical imaging allows to detect targets, as tumors. Even if we use infra-red light, the autofluorescence of biological tissues is also detected in this wavelength range: it needs to be removed to get accurate detection results. We introduced the Non-negative Matrix Factorization as blind source separation method to unmix fluorescence spectra and eradicate autofluorescence. In order to improve detection of spatially sparse fluorescent markers, we suggested a new NMF algorithm with sparsity constrained. This algorithm asks for a wished sparsity value to reach: we propose a simplified way to approximate sparseness from the size of tumors we are looking for.

A study has been conducted on simulated data to compare sparse and non-sparse algorithms unmixing performances. Our sparse NMF algorithm appeared to be more robust to the random choice of initialization spectra, and returned results closer to the expected spectra than the non-sparse algorithm. NMF method without constraints offering infinity of solutions, we got interested in the range of admissible solutions for both algorithms: once more, sparse NMF algorithm gives better results than the algorithm without sparsity constraints.

Finally, both algorithms were run on *in vivo* acquisitions, and successfully unmixed up to three different fluorescent sources. Acquisitions had been obtained on a mouse with subcutaneous specific fluorescent markers. Once more, results are closer to reality with the sparse NMF algorithm than

with the non-constrained algorithm, which comforts simulation outcomes.

Nevertheless, even if fluorescent markers are expected to be specific enough to only accumulate around the tumors, reality of *in vivo* experiments is for now different. Indeed, once IV-injected, currently used markers travel in blood and lymphatic canals and spread everywhere in tissues: one may talk of non-specific signal. Only a few supplementary amount of markers accumulate around the tumor: the signal measured there is called specific signal. Even if more markers will accumulate around the tumor, a small amount of non-specific signal will always be detected everywhere in tissues. In such case, sparsity constraints may seem inappropriate for the unmixing problem. However, injected specific markers are developed to be more and more specific to tumors. We thus may expect in the near future that markers will be specific enough to avoid a non-specific signal detection, and use sparsity constraints for the unmixing step.

As optical imaging tries to detect deeper and deeper embedded tumors, NMF – as a preprocessing tool to remove autofluorescence signal and isolate specific fluorescence contributions, sharpened by several constraints – is an helpful tool. Sparse NMF algorithm we proposed greatly improves the preprocessing results in spatially sparse markers localization.

## REFERENCES

- [1] V. Ntziachristos, E.A. Schellenberger, J. Ripoll, D. Yessayan, E. Graves, A. Bogdanov, L. Josephson, and R. Weissleder, "Visualization of antitumor treatment by means of fluorescence molecular tomography with an annexin v-cy5. 5 conjugate," *Proceedings of the National Academy of Sciences*, vol. 101, no. 33, pp. 12294, 2004.
- [2] F. Leblond, S. C. Davis, P. A. Valds, and B. W. Pogue, "Pre-clinical whole-body fluorescence imaging: Review of instruments, methods and applications," *Journal of Photochemistry and Photobiology B: Biology*, vol. 98, no. 1, pp. 77 – 94, 2010.
- [3] C. Bremer, V. Ntziachristos, and R. Weissleder, "Optical-based molecular imaging: contrast agents and potential medical applications," *European Radiology*, vol. 13, no. 2, pp. 231 – 243, 2003.
- [4] B. Ballou, L. A. Ernst, and A. S. Waggoner, "Fluorescence imaging of tumors *in vivo*," *Curr. Med. Chem.*, vol. 12, pp. 795 – 805, 2005.
- [5] J. Hung, S. Lam, J. C. Leriche, and B. Palcic, "Autofluorescence of normal and malignant bronchial tissue," *Lasers in surgery and medicine*, vol. 11, no. 2, pp. 99–105, 1991.
- [6] D. Wood, G. Feke, D. Vizard, and R. Papineni, "Refining epifluorescence imaging and analysis with automated multiple-band flat-field correction," Tech. Rep., Nature methods, 2008.
- [7] R. Levenson and P. J. Cronin, "Spectral imaging of deep tissue," Jan. 2008, US Patent 7,321,791.
- [8] C. C. Hoyt, "Systems and methods for *in-vivo* optical imaging and measurement," Dec. 2006, US Patent App. 11/295,140.
- [9] P. Cronin and P. J. Miller, "Spectral imaging," Sept. 18 2005, US Patent App. 11/522,745.
- [10] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Blind source separation by higher-order singular value decomposition," *Proceedings of EURASIP European Signal Processing Conference (EUSIPCO)*, vol. 1, pp. 175–178, 1994.
- [11] R. C. Henry, "Current factor analysis receptor models are ill-posed," *Atmospheric Environment (1967)*, vol. 21, no. 8, pp. 1815 – 1820, 1987.
- [12] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Advances in neural information processing systems*, vol. 13, pp. 556–562, 2001.
- [13] C. Gobinet, E. Perrin, and R. Huez, "Application of nonnegative matrix factorization to fluorescence spectroscopy," in *Proceeding EUSIPCO*, 2004, pp. 6–10.

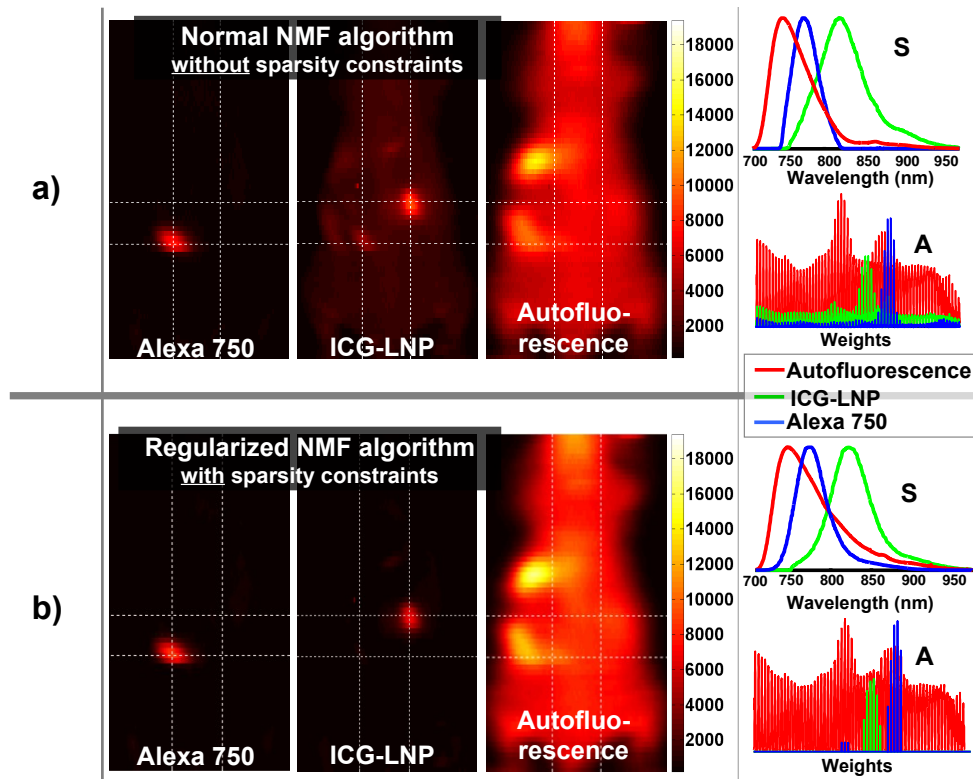


Fig. 17. Results of unmixing obtained with both algorithms: separated fluorescence contributions, and matrices  $A$  and  $S$  at convergence: a) with classical NMF algorithm, without sparsity constraints and b) with the regularized NMF algorithm, when taking sparsity constraints into account.

- [14] C. Gobinet, V. Vrabie, A. Tfayli, O. Piot, R. Huez, and M. Manfait, "Pre-processing and source separation methods for Raman spectra analysis of biomedical samples," in *Engineering in Medicine and Biology Society, 29th Annual International Conference of the IEEE*, 2007, pp. 6207–6210.
- [15] H. Xu, C. Kuo, and B. Rice, "Improved sensitivity by applying spectral unmixing prior to fluorescent tomography," *Proceeding to Biomedical Optics BIOMED*, p. paper BMC1, 2008.
- [16] H. Xu and B. W. Rice, "In-vivo fluorescence imaging with a multivariate curve resolution spectral unmixing technique," *Journal of Biomedical Optics*, vol. 14, no. 6, pp. 064011, 2009.
- [17] A. Cichocki, S. Amari, A.-H. Phan, and R. Zdunek, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley-Blackwell, Chichester, UK, 2009.
- [18] A. Cichocki, R. Zdunek, and S. Amari, "New algorithms for non-negative matrix factorization in applications to blind source separation," in *Proceeding to IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006, vol. 5, pp. V–V.
- [19] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics and Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.
- [20] H. Kim and H. Park, "Non-negative matrix factorization based on alternating non-negativity constrained least squares and active set method," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 713–730, 2008.
- [21] A. Cichocki, R. Zdunek, and S. Amari, *Independent Component Analysis and Blind Signal Separation*, vol. 3889, pp. 32–39, Springer Berlin / Heidelberg, 2006.
- [22] A. B. Hamza and D. J. Brady, "Reconstruction of reflectance spectra using robust nonnegative matrix factorization," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3637–3642, 2006.
- [23] I. Dhillon and S. Sra, "Generalized nonnegative matrix approximations with Bregman divergences," *Advances in neural information processing systems*, vol. 18, pp. 283–290, 2006.
- [24] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.
- [25] K. Stadlthanner, F.J. Theis, C.G. Puntonet, and E.W. Lang, "Extended sparse nonnegative matrix factorization," in *Computational Intelligence and Bioinspired Systems*. 2005, vol. 3512, pp. 249–256, Springer Berlin / Heidelberg.
- [26] Z. H. Jin, V. Josserand, S. Foillard, D. Boturyn, P. Dumy, M. C. Favrot, and J. L. Coll, "In vivo optical imaging of integrin  $\alpha$ v- $\beta$ 3 in mice using multivalent or monovalent crgd targeting vectors," *Molecular Cancer*, vol. 41, no. 6, pp. 1–9, 2007.
- [27] Y. Urano, D. Asanuma, Y. Hama, Y. Koyama, T. Barrett, M. Kamiya, T. Nagano, T. Watanabe, A. Hasegawa, P. L. Choyke, et al., "Selective molecular imaging of viable cancer cells with pH-activatable fluorescence probes," *Nature medicine*, vol. 15, no. 1, pp. 104–109, 2008.
- [28] R. Weissleder, C. H. Tung, U. Mahmood, and A. Bogdanov, "In vivo imaging of tumors with protease-activated near-infrared fluorescent probes," *Nature biotechnology*, vol. 17, pp. 375–378, 1999.
- [29] J. Razkin, V. Josserand, D. Boturyn, Z. Jin, P. Dumy, M. Favrot, J. L. Coll, and I. Texier, "Activatable fluorescent probes for tumour-targeting imaging in live mice," *ChemMedChem*, vol. 1, no. 10, pp. 1069–1072, 2006.
- [30] A.-S. Montcuquet, L. Hervé, F. Navarro, J.-M. Dinten, and J. I. Mars, "Nonnegative matrix factorization: a blind spectra separation method for in vivo fluorescent optical imaging," *J. Biomed. Opt.*, vol. 15, no. 5, pp. 056009–1 – 056009–14, 2010.
- [31] J.-C. Bariteaux, K. Hassler, and M. Unser, "An efficient numerical method for general regularization in fluorescence molecular tomography," *IEEE Trans Med Imaging*, vol. 29, no. 4, pp. 1075–1087, 2010.
- [32] S. Moussaoui, D. Brie, and J. Idier, "Non-negative source separation: range of admissible solutions and conditions for the uniqueness of the solution," in *Proceeding to IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2005, vol. 5, pp. 289–292.
- [33] F. P. Navarro, M. Berger, M. Goutayer, S. Guillermet, V. Josserand, P. Rizo, F. Vinet, and I. Texier, "A novel indocyanine green nanoparticle probe for non invasive fluorescence imaging in vivo," *Progress in biomedical optics and imaging*, vol. 10, no. 30, pp. 1–10, 2009.