This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

Constrained Tensor Decomposition for Longitudinal Analysis of Diffusion Imaging Data

Claudio Stamile, François Cotton, Dominique Sappey-Marinier, Sabine Van Huffel, *Fellow, IEEE*

***Abstract*—Analysis of complex data is still a challenge in medical image analysis. Due to the heterogeneous information that can be extracted from magnetic resonance imaging (MRI) it can be difficult to fuse such data in a proper way. One interesting case is given by the analysis of diffusion imaging (DI) data. DI techniques give an important variety of information about the status of microstructure in the brain. This is interesting information to use especially in longitudinal setting where the temporal evolution of the pathology is an important added value.**

In this paper, we propose a new tensor-based framework capable to detect longitudinal changes appearing in DI data in multiple sclerosis (MS) patients. We focus our attention to the analysis of longitudinal changes occurring along different white matter (WM) fiber-bundles. Our main goal is to detect which subset of fibers (within a bundle) and which sections of these fibers contain "pathological" longitudinal changes. The framework consists of three main parts: *i)* preprocessing of longitudinal diffusion acquisitions and WM fiber-bundles extraction, *ii)* data tensorization and rank selection, *iii)* application of a parallelized constrained tensor factorization algorithm to detect longitudinal "pathological" changes.

The proposed method was applied on simulated longitudinal variations and on real MS data. High level of accuracy and precision were obtained in the detection of small longitudinal changes along the WM fiber-bundles.

***Index Terms*—Multiple Sclerosis, Tensor Factorization, Diffusion Imaging, Longitudinal Analysis, Tractography.**

I. Introduction

Magnetic resonance imaging (MRI) is becoming the reference technique to assess brain related pathology. The evolution of fast magnetic acquisition techniques, allowed to use MRI to perform short term longitudinal follow-up of single patients in order to assess the pathological course.

This work is funded by the following projects: Henri Benedictus Fellowship of the Belgian American Educational Foundation. French National Research Agency (ANR) within the national program "Investissements d'Avenir" through the OFSEP project (ANR-10-COHO-002). EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (n° 339804); and in part by EU MC ITN TRANSACT 2012 (316679). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

Claudio Stamile, François Cotton and Dominique Sappey-Marinier are with CREATIS; CNRS UMR5220; INSERM U1206; Université de Lyon 1, INSA-Lyon, Villeurbanne, France (email: stamile@creatis.insa-lyon.fr; francois.cotton@chu-lyon.fr; dominique.sappey-marinier@univ-lyon1.fr)

Claudio Stamile and Sabine Van Huffel are with KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Leuven, Belgium (email: {Claudio.Stamile,Sabine.VanHuffel}@esat.kuleuven.be)

François Cotton is also with Service de Radiologie, Centre Hospitalier Lyon-Sud, Hospices Civils de Lyon, Pierre-Bénite, France

Dominique Sappey-Marinier is also with CERMEP - Imagerie du Vivant, Université de Lyon, Bron, France

Among all the available MRI biomarkers, diffusion tensor imaging (DTI) or, in general, diffusion imaging (DI), is one of the most interesting techniques to analyze structural changes in the brain. More in detail, in the last decade, different papers showed how image biomarkers derived from DI are sensible for the characterization of brain pathologies like brain tumors and multiple sclerosis (MS) [5], [11]. Moreover, DI was successfully applied in longitudinal follow-up of brain related pathology showing a good capability to enhance the presence of small longitudinal variations occurring in white matter (WM). These results are encouraging in exploiting DI in follow-up of relapsing remitting (RR) MS patients where small longitudinal changes occur in WM during the evolution of the pathology.

Main challenge in DI analysis is derived from the large number of information that is possible to derive. Indeed DI provides both structural and quantitative information. The former can be derived from the application of a tractography algorithm that allows to reconstruct the architecture of WM fibers. The latter is obtained by using a specific model to analyze diffusion like DTI. This information allows to obtain quantitative measures helpful to assess the properties of the tissue. If in this setting longitudinal information is added, it is quite easy to see how the corresponding data analysis problem starts to become quite complex.

Recently, different efforts were made in order to perform such complex analysis. In [6], the authors showed a method to analyze longitudinal changes in WM. Main limitation of this method is related to its capability to perform a longitudinal analysis restricted to two time-points. Moreover, it can not take into account multiparametric data at the same time. In our previous works, we overcome the limitation of this method [19], [17], [20]. In particular, in [20] we developed a method based on non-negative matrix factorization to perform longitudinal analysis in MS patients in cross-sections of the WM fiber-bundle. Since the major constraint in using a matrix factorization is given by the fixed number of the dimensions, this method was applied only to one specific cross-section of the bundle at time. As a result, only local information contained in a specific cross-section was taken into account during the analysis, ignoring the information available along different fibers within the bundle.

In the last years, tensor factorization techniques [16] emerged as a convincing tool to perform data fusion. More in detail, their capability to analyze data, that can be represented using more than two dimensions, made them a perfect tool easy to use in scenarios where more than two dimensions are needed to represent the data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

In this paper, we propose a new tensor-based framework to analyze longitudinal changes in MS patients occurring along WM fiber-bundles. Our main goal is to detect which subset of fibers within a WM fiber-bundle and which sections of those fibers contain "pathological" longitudinal changes and which are the time-points affected by those changes. We developed a complete fully automated pipeline to register and extract WM fiber-bundles from MRI data and analyze them using tensor factorization. Moreover, in order to optimize the computational time needed to perform the tensor factorization, we developed a parallel pipeline to perform such analysis.

This paper is structured as follows. In Section II, we provide a detailed description of the processing pipeline used in this work to extract WM fiber-bundles and register longitudinal data. In Section III, we describe how our problem to analyze longitudinal changes in WM fiber-bundle can be formalized using tensors. In Section IV we show how to speed-up the proposed framework using parallel programming. In Section V, we present our experimental campaign. In Section VI we show our results. In Section VII we discuss the performances, benefits and limitations of our method. Finally, in Section VIII, we draw our conclusions.

II. Our Data processing

Before describing or data processing pipeline, we need to state some assumptions made in this paper. More in detail, we assume that each patient has longitudinal MRI acquisition. In general, we denote with s the number of acquisitions for a single patient. Moreover, for each patient we assume that diffusion data are available at each time-point.

A. Data registration

As first step, each of the s time-points ($T_1 \dots T_s$) of DTI longitudinal acquisitions, are processed. Eddy current correction [9] was first applied on the diffusion volumes using the b0 volume ($b = 0s.mm^{-2}$) as reference. Tensor model was then computed using the FDT module of FSL [9]. In order to minimize any bias introduced by the registration process, the longitudinal data were co-registered using the diffusion tensor based method described in [10]. More in details, the following steps were performed: *i)* generation of a patient-specific template obtained from longitudinal diffusion tensor images, *ii)* co-registration of the resulting template to the Illinois Institute of Technology (IIT) atlas [22], and *iii)* co-registration of each time-point data into the IIT atlas space by applying the previously obtained transformations to the initial longitudinal data. After the co-registration the obtained tensor image was used to compute, for each time-point, 8 diffusion metric maps: fractional anisotropy (FA), mean diffusivity (MD), Compositional Kullback-Leibler Anisotropy (KLA), Angular Anisotropy (AA) [14], radial diffusivity (λ_r) and the three eigenvalues of the diffusion tensor (λ_1, λ_2, λ_3).

B. Fiber-bundle extraction

The fiber orientation density (FOD) information contained in the IIT Atlas were used in order to perform a probabilistic streamline tractography using MRTrix [21]. From the obtained tractography, the algorithm described in [18], coupled with the regions of interest (ROI) contained in the JHU atlas [7], was applied in order to extract the fiber-bundles. In order to analyze fiber-bundles an additional step is needed. Indeed, the output of the tractography could not be directly used for the analysis of the fiber-bundle since the number of points used to reconstruct the fibers varies. Moreover, start and end point of each fiber are not consistent within the same fiber-bundle. Fibers could start randomly from the two extremities of the bundle. In order to overcome those problems part of the pipeline described in [19] was applied to process the fiber-bundle. As first step we define common start/end points of each fiber within the bundle. A classical K-Means algorithm [13] is performed to generate two different clusters, R_1 for the starting points and R_2 for the ending points. Fiber points are then reordered from R_1 to R_2 and fibers that did not link the two clusters (broken fibers) are automatically removed. As final post-processing step each fiber is resampled with the same number $q = 100$ of points (also called nodes or cross-sections).

C. Fiber-bundle formalization

After the post-processing we can formalize the extracted fiber-bundle as set $F = \{f_1, f_2, \dots, f_v\}$ composed of v fibers $f_j = \{\boldsymbol{p_1}, \dots, \boldsymbol{p_q}\}$ where $\boldsymbol{p_i} = (x_i, y_i, z_i) \mid 1 \leq i \leq q$. The coordinate $\boldsymbol{p_i}$ is used to extract the voxel's value of one of the six diffusion maps ($FA(\boldsymbol{p_i})$ in case of FA) in the corresponding location of f_j. With this formalization it is possible to analyze the signal along each fiber within the bundle. For instance, the FA signal profile along a fiber $f_j \in F$ can be regrouped in the set $E_j = \{FA(\boldsymbol{p_1}), \dots, FA(\boldsymbol{p_q})\}$. This subdivision is graphically described in Figure 1. A common way to represent the signal profile along the whole fiber-bundle is to average all the signal profiles along each fiber $f_j \in F$. The global profile is then expressed using mean and standard deviation of the signal along the fiber bundle as presented in Figure 1.

Similarly, taking all the cross-sections $\boldsymbol{p}$ with the index i in each fiber $f \in F$ it is possible to analyze the global diffusion values in a particular cross-section of F. More in detail we can collect all the FA values belonging to a given cross-section of F defining the following set: $S_i = \{FA(\boldsymbol{p_i}) \mid \boldsymbol{p_i} \in f \; \forall \; f \in F\}$ where i is the fixed index representing the cross-section to analyze.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

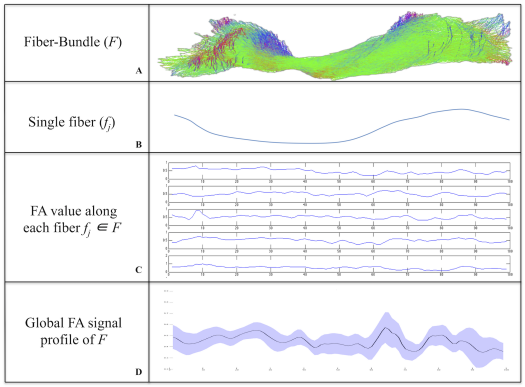

Figure 1. **A)** Original fiber-bundle F. **B)** A single fiber $f_j \in F$. **C)** FA signal along each fiber $f_j \in F$. **D)** Mean FA signal (black line) with standard deviation (coloured band) representing the global signal profile along F.

III. Fiber-Bundle as tensor

Before the introduction of our tensor-based formalism we need to define the notation used in this paper. More in detail, if not clearly stated, we will use the notation described in [12]. We denote with lower case letters e.g. a scalar values, with bold lower case letters e.g. $\boldsymbol{a}$ 1-dimensional vectors, with boldface uppercase letter e.g. $\boldsymbol{A}$ matrices and with boldface Euler script letters $\boldsymbol{\mathcal{A}}$ tensors.

In this work we describe a new method to analyze longitudinal changes visible along WM fiber-bundles using a tensor-based model. In Figure 2 we show the model of our tensor to describe the signal along a given fiber-bundle. The third-order tensor $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{v \times q \times m}$ is generated by concatenating all the longitudinal features extrapolated along all the fibers within a bundle.

More in detail, let z be the number of features we want to extrapolate along the fiber-bundle and let s be the number of acquired time-points. Since we have multiple features and multiple time-points, we define with M_{ij} $1 \leq i \leq z$, $1 \leq j \leq s$ the $i-th$ feature extrapolated at the $j-th$ time-point. Using this information the tensor $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{v \times q \times m}$, with $m = s * z$, is built. In Figure 3 we show a graphical representation describing how the z features of each time-point are stacked in the 3^{rd} mode of the tensor $\boldsymbol{\mathcal{T}}$.

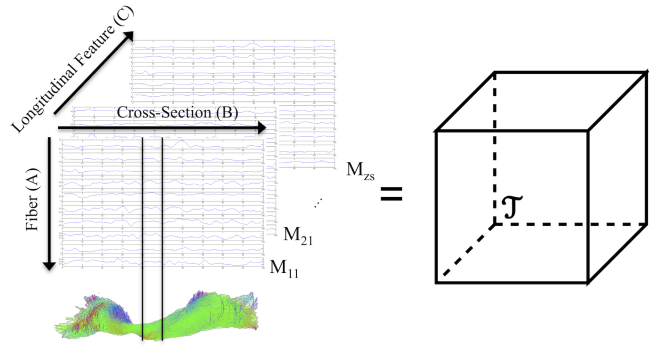

Figure 2. Tensorization of longitudinal diffusion features along a fiber-bundle (cortico-spinal tract in this case).

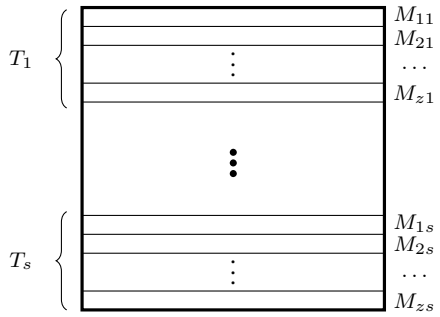

Figure 3. Representation of the 3^{rd} mode of the tensor $\boldsymbol{\mathcal{T}}$

This tensor-based representation gives us the possibility to analyze, in this one structure, local and global aspects of a fiber-bundle.

A. Tensor factorization using canonical polyadic decomposition

The canonical polyadic decomposition (CPD) decomposition factorizes a tensor into a sum of rank-one tensors. For example, given a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{v \times q \times m}$ we wish to decompose it as:

$$\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R} \boldsymbol{a_r} \circ \boldsymbol{b_r} \circ \boldsymbol{c_r} + \boldsymbol{\mathcal{E}} \tag{1}$$

where R is a positive integer, $\boldsymbol{a_r} \in \mathbb{R}^v, \boldsymbol{b_r} \in \mathbb{R}^q, \boldsymbol{c_r} \in \mathbb{R}^m \ \forall \ 1 \leq r \leq R$ are the mode-1, mode-2 and mode-3 vectors respectively and $\boldsymbol{\mathcal{E}} \in \mathbb{R}^{v \times q \times m}$ is the residual tensor. The symbol "$\circ$" represents the outer product.

The rank of a tensor $\boldsymbol{\mathcal{X}}$, denoted $rank(\boldsymbol{\mathcal{X}})$, is defined as the smallest number of rank-one tensors that generate $\boldsymbol{\mathcal{X}}$ as their sum. In other words, this is the smallest number of components in an exact CP decomposition, where "exact" means that there is equality in equation 1 with residual tensor $\boldsymbol{\mathcal{E}}$ set to zero. An exact CP decomposition with $R = rank(\boldsymbol{\mathcal{X}})$ components is called the canonical polyadic decomposition (CPD).

The CPD problem can be formalized as follows:

$$\min_{\hat{\boldsymbol{\mathcal{X}}}} \|\boldsymbol{\mathcal{X}} - \hat{\boldsymbol{\mathcal{X}}}\|_F \quad \text{with} \quad \hat{\boldsymbol{\mathcal{X}}} = \sum_{r=1}^{R} \boldsymbol{a_r} \circ \boldsymbol{b_r} \circ \boldsymbol{c_r} = [\![\boldsymbol{A}; \boldsymbol{B}; \boldsymbol{C}]\!] \tag{2}$$

In our specific case, since non-negative values are present in the analysis, we imposed a non-negativity constraint to the factorization. More in detail, in order to factorize our tensor, the following optimization problem is solved.

$$\min_{\boldsymbol{A} \geq 0, \boldsymbol{B} \geq 0, \boldsymbol{C} \geq 0} \|\boldsymbol{\mathcal{T}} - \boldsymbol{A} \circ \boldsymbol{B} \circ \boldsymbol{C}\| \tag{3}$$

The formulation of the tensor factorization using CPD is graphically described in Figure 4.

According to the tensorization process previously described, the factor matrix $\boldsymbol{A} \in \mathbb{R}^{v \times R}$ stacks in its columns all vectors a_r describing the contribution of each fiber during the factorization. Similarly, the factor matrix $\boldsymbol{B} \in \mathbb{R}^{q \times R}$ stacks all the b_r and contains the contribution of each fiber-bundle, and finally the factor matrix $\boldsymbol{C} \in \mathbb{R}^{m \times R}$, stacking

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

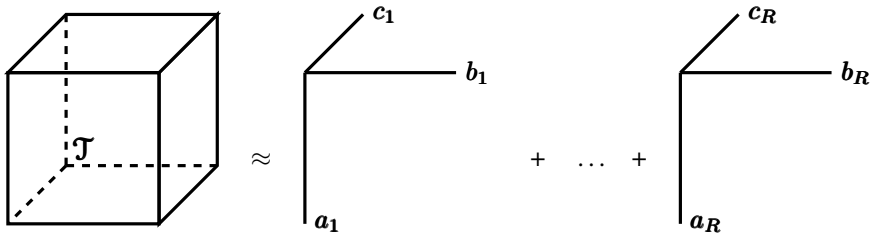

Figure 4. Canonical polyadic decomposition

all c_r, contains the contribution of each feature extracted in a specific time-point. We used the information contained in the $\boldsymbol{A}$ (fibers information) and $\boldsymbol{B}$ (cross-sectional information) factor matrices in order to discriminate fibers and cross-sections affected by longitudinal changes. The factor matrix $\boldsymbol{C}$ is used to identify which are, among the R sources, the ones containing "pathological" longitudinal changes. The idea behind the automatic selection of the sources containing information about the longitudinal "pathological" changes is described in Section III-C.

B. Rank estimation for factorization

An important issue in tensor factorization is the estimation of the rank for the CPD factorization of a tensor $\mathcal{T}$. The literature on estimation of decomposition rank from the tensor is limited. Tensorlab package [23] has a method *rankest*, which estimates the rank based on the L-curve of the number of rank-one terms in a CPD [3]. This approach allows to obtain the minimum number of component to get the best fit. In a real case scenario, this method can overestimate the real rank of the data. Indeed, the estimated rank tends to be larger than the real rank of components present in the image, due to noise or biological variations. In this work, since the proposed method is capable to automatically identify which are the sources of interest, (as we will describe in Section III-C), we use *rankest* function to estimate the rank of the factorization. We prefer to overestimate the real rank by using the *rankest* function, in order to avoid loss of information after the factorization.

C. Detection of longitudinal changes from tensor factorization

In order to detect the sources containing the changes generated by longitudinal variations, we generalize the method we previously proposed in [20]. In our previous work [20], we introduced a model to detect longitudinal changes in sources obtained by factorizing data using a non-negative matrix factorization (NMF). Here we generalize the previous model to the tensor case where each component to analyze is represented by a one dimensional vector.

After the factorization of the tensor $\mathcal{T}$ in R components, we obtained the vectors $\boldsymbol{a_r} \in \mathbb{R}^v, \boldsymbol{b_r} \in \mathbb{R}^q, \boldsymbol{c_r} \in \mathbb{R}^m \; \forall \, 1 \leq r \leq R$. In order to detect if a component $i \; 1 \leq i \leq R$ captures abnormal changes contained in certain features in specific time-points its vector $\boldsymbol{c_i}$ is used. We recall that the vector $\boldsymbol{c_i}$ contains the information of all the z features extracted in the s time-points as follows:

$$\boldsymbol{c_i} = [\underbrace{c_1^i, \cdots, c_z^i}_{T_1}, \cdots, \underbrace{c_{s*z-z}^i, \cdots, c_{s*z}^i}_{T_s}]$$

In this vector, each block of z consecutive elements represents the contribution of each diffusion feature, extracted in a specific time-point T_i. Starting from the vector $\boldsymbol{c_i}$ the goal is to detect if the $i{-}th$ components contains longitudinal variations. More in detail, we say that a longitudinal variation appears in the $i-th$ component if longitudinal variations are present in all the diffusion metrics belonging to certain time points of T_i. We define these "changed time points" as *outliers*. In order to clarify this concept, we report in Figure 5 a graphical example of outliers contained in component vector $\boldsymbol{c_i}$. In the rest of this section we will show how we can: *i)* detect the outliers in the component vector $\boldsymbol{c_i}$, *ii)* exploit the information derived from the outlier analysis in order to extract fibers and cross-sections affected by longitudinal pathological changes.

Since the tensor factorization generates R components, this check is performed for all $\boldsymbol{c_i}$ $1 \leq i \leq R$. Detection of outliers was performed using a density-based local outliers (LOF) algorithm [2].

The LOF algorithm allows to detect outliers by computing the LOF value for each element. The LOF value of each object represents the degree of the object to be an outlier compared to the other elements in the cluster. This value strongly depends on a single parameter $MinPts$, which represents the number of nearest neighbours used in defining the local neighborhood of the object [2]. The main problem related to the LOF is the difficulty to interpret resulting LOF scores since there are no clear rules that define when a point is an outlier. In order to properly detect the outliers, this value should be carefully selected for the specific dataset. A value is defined as "outlier" if and only if the LOF value is greater than a fixed threshold value ω.

In order to use the LOF algorithm to detect if $\boldsymbol{c_i}$ contains *outlying time-points*, we represent this vector as a matrix defined as follow:

$$\hat{\boldsymbol{C}}_{\boldsymbol{i}} = \begin{pmatrix} c_1^i & \cdots & c_z^i \\ \vdots & \ddots & \vdots \\ c_{s*z-z}^i & \cdots & c_{s*z}^i \end{pmatrix} \begin{matrix} T_1 \\ \vdots \\ T_s \end{matrix}$$

the matrix $\hat{\boldsymbol{C}}_{\boldsymbol{i}} \in \mathbb{R}^{s \times z}$ is the matrix representation of the original vector $\boldsymbol{c_i}$ where each of the s rows represents a time-point defined by its diffusion features.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

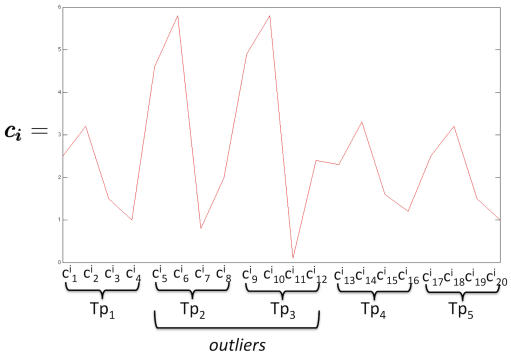

Figure 5. Graphical example showing a plot of a component vector $\boldsymbol{c_i}$ with $s = 5$ and $z = 4$. The vector contains outliers values from c^i_5 to c^i_{12} corresponding to time-point 2 and time-point 3. Detection of those outliers time-points allows to understand if the $i-th$ component "captures" longitudinal alterations. Moreover outlier detection in $\boldsymbol{c_i}$ allows to detect time-points containing longitudinal pathological changes.

The matrix $\hat{\boldsymbol{C}}_{\boldsymbol{i}}$ can now be used by the LOF algorithm in order to detect if *outlying time-points* are present in the corresponding $i-th$ component.

As a result of the application of the LOF algorithm on the matrix $\hat{\boldsymbol{C}}_{\boldsymbol{i}}$ a vector $\boldsymbol{l_i} \in \mathbb{R}^s$ is generated. This vector contains real values l^i_j $1 \leq j \leq s$ representing the LOF value corresponding to the $j-th$ row of the matrix $\hat{\boldsymbol{C}}_{\boldsymbol{i}}$. The LOF value of each row in $\hat{\boldsymbol{C}}_{\boldsymbol{i}}$ allows us to understand if a specific time-point in the $i-th$ component is an outlier. Or, in other words, this value allows to detect if the $i-th$ component of the tensor factorization captures "longitudinal pathological variations" appearing in the follow-up. We say that the $i-th$ component of the factorization contains pathological longitudinal changes if the following condition holds:

$$\exists\ l^i_j \in \boldsymbol{l_i} \mid l^i_j > \omega,\ 1 \leq j \leq s \tag{4}$$

where ω is the LOF threshold that allows to define whether an element is an outlier or not. By applying this procedure to each component it is possible to obtain two types of information: *i)* the components who "capture" the longitudinal pathological changes and *ii)* the time-points where the pathological longitudinal changes appear. This procedure is summarized in Algorithm 1. In detail, the algorithm take the component vectors $\boldsymbol{c_1}, \ldots, \boldsymbol{c_s}$ generated from the tensor factorization and the LOF threshold ω. The algorithm applies the LOF algorithm to each matrix representation of the component vector in order to detect which components and which time-points contain longitudinal changes. The algorithm stores all the time-points and the components containing the longitudinal changes in the sets U and Y respectively.

The information derived from the proposed algorithm can be used in order to detect which fibers (represented by the component matrix $\boldsymbol{A}$) and which cross-sections (represented by the component $\boldsymbol{B}$) are affected by longitudinal pathological changes.

More in detail, in order to verify whether a fiber f_i presents longitudinal changes the following condition is checked. Let $a_{y[i]}$ be the $i-th$ element of the component vector $\boldsymbol{a_y}$ obtained from the tensor factorization and let Y be the set containing

Algorithm 1: ALGORITHM FOR DETECTION OF COMPONENTS AND TIME-POINTS AFFECTED BY LONGITUDINAL CHANGES

CMPTPDET $(\boldsymbol{c_1}, \ldots, \boldsymbol{c_R}, \omega)$;
Input : the component vectors $\boldsymbol{c_1}, \ldots, \boldsymbol{c_R}$ obtained from the tensor factorization
a positive real value $\omega > 0$ representing the LOF threshold
Output: a set Y with the components detected as containing longitudinal changes
a set U containing the time-points affected by the longitudinal changes
begin
  $Y \leftarrow \emptyset$;
  $U \leftarrow \emptyset$;
  foreach $i \in \{1, \ldots, R\}$ **do**
    $\hat{\boldsymbol{C}}_{\boldsymbol{i}} =$ RESHAPE$(\boldsymbol{c_i})$;
    $\boldsymbol{l_i} =$ LOF$(\hat{\boldsymbol{C}}_{\boldsymbol{i}})$;
    foreach $j \in 1, \ldots, s$ **do**
      if $l^i_j > \omega$ **then**
        $U \leftarrow U \cup j$;
      end
    end
    $Y \leftarrow Y \cup i$;
  end
  return Y, U
end

the components detected as "outliers" by the LOF algorithm. We say that f_i is affected by longitudinal pathological changes if the following condition holds:

$$\exists y \in Y \mid a_{y[i]} > a_{r[i]} \forall 1 \leq r \leq R, r \neq y \tag{5}$$

Roughly speaking, this condition checks whether the fiber f_i has its maximal contribution in one of the components detected as "outliers" by the LOF algorithm.

Similarly, in order to check if a specific cross-section j $(1 \leq j \leq q)$ of the fiber-bundle is affected by longitudinal changes, the following condition is also checked. Let $b_{y[j]}$ be the $j-th$ element of the component vector $\boldsymbol{b_y}$ obtained from the tensor factorization and let Y be the set containing the components detected as "outliers" by the LOF algorithm. We say that the cross-section j is affected by longitudinal changes if the following condition holds:

$$\exists y \in Y \mid b_{y[j]} > b_{r[j]} \forall 1 \leq r \leq R, r \neq y \tag{6}$$

Like for the fibers, this condition checks whether the cross-section j has its maximal contribution in one of the components detected as "outliers" by the LOF algorithm.

The procedures used for the detection of all the fibers and cross-sections affected by the longitudinal changes are described in Algorithm 2 and Algorithm 3 respectively. Both the algorithms check, for each fiber and cross-section, whether the conditions defined in equation 5 and 6 are satisfied. The algorithms store all the fibers and the cross-section detected as containing the pathological longitudinal changes in the sets P and H respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

Algorithm 2: ALGORITHM FOR DETECTION OF FIBERS AFFECTED BY LONGITUDINAL CHANGES

FIBDET $(\boldsymbol{a_1}, \ldots, \boldsymbol{a_R}, Y, v)$;
Input : the component vectors $\boldsymbol{a_1}, \ldots, \boldsymbol{a_R}$ obtained from the tensor factorization
the set Y with the components detected as containing longitudinal changes by Algorithm 1
number of fibers v
Output: A set P with the fibers containing longitudinal pathological changes
begin
  $H \leftarrow \emptyset$;
  foreach $i \in \{1, \ldots, q\}$ **do**
    if $\exists y \in Y \mid a_{y[i]} > a_{k[i]} \forall 1 \leq k \leq R, k \neq y$ **then**
      $P \leftarrow P \cup i$
    end
  end
  return P
end

Algorithm 3: ALGORITHM FOR DETECTION CROSS-SECTION AFFECTED BY LONGITUDINAL CHANGES

CSDET $(\boldsymbol{b_1}, \ldots, \boldsymbol{b_R}, Y, q)$;
Input : the component vectors $\boldsymbol{b_1}, \ldots, \boldsymbol{b_R}$ obtained from the tensor factorization
the set Y with the components detected as containing longitudinal changes by Algorithm 1
number of cross-sections q
Output: A set H with the cross-sections containing longitudinal pathological changes
begin
  $H \leftarrow \emptyset$;
  foreach $i \in \{1, \ldots, q\}$ **do**
    if $\exists y \in Y \mid b_{y[i]} > b_{k[i]} \forall 1 \leq k \leq R, k \neq y$ **then**
      $H \leftarrow H \cup i$
    end
  end
  return H
end

IV. PARALLEL IMPLEMENTATION OF THE PROPOSED METHOD

In order to reduce the computation time needed to perform the factorization of the fiber-bundle, we parallelized the algorithm according to the "divide et impera" programming paradigm. Instead of computing the CPD and LOF on the whole fiber-bundle, we split the fiber-bundle in small subsets of fibers. In order to split the whole fiber-bundle in K different sub fiber-bundles, for each pair of fibers $f_a, f_b \in F$ we compute the Minimum Average Direct Flip (MDF) metric [4]. The MDF metric computes the distance between two fibers composed by q points according to the following equations:

$$d_{direct}(f_a, f_b) = d(f_a, f_b) = \frac{1}{q}\sum_{i=1}^{q} \|\boldsymbol{p_i^a} - \boldsymbol{p_i^b}\| \quad (7)$$

$$d_{flipped}(f_a, f_b) = d(f_a, f_b^{\#}) = d(f_a^{\#}, f_b) \quad (8)$$

$$MDF(f_a, f_b) = min(d_{direct}(f_a, f_b), d_{flipped}(f_a, f_b)) \quad (9)$$

where $\|\boldsymbol{p_i^a} - \boldsymbol{p_i^b}\|$ represents the Euclidean distance between the $i-th$ point of f_a and f_b. In equation 8, $f_a^{\#}$ denotes the flipped version of f_a, e.g. $f_a = \{\boldsymbol{p_1}, \boldsymbol{p_2}, \ldots, \boldsymbol{p_q}\}$, $f_a^{\#} = \{\boldsymbol{p_q}, \boldsymbol{p_{q-1}}, \ldots, \boldsymbol{p_1}\}$. The MDF distance is a metric on the space of fibers, it respects the triangle inequality and it is fast to compute [4].

Let v the total number of fibers, we can compute the MDF distance for each pair of fibers building the positive symmetric matrix $\boldsymbol{M} \in \mathbb{R}^{v \times v}$ where $m_{ij} \in \boldsymbol{M}$ represents the MDF distance computed between the fibers f_i and f_j. The matrix $\boldsymbol{M}$ is then used as distance matrix for the K-medoids clustering algorithm in order to generate a set $Cl = \{F_1, F_2, \ldots, F_K\}$ of K clusters (Algorithm 4). In our prospective, a cluster $F_i \in Cl$ represents a sub fiber-bundle generated from the original fiber-bundle F such that $\bigcup_{i=1}^{K} F_i = F$.

Algorithm 4: ALGORITHM FOR CLUSTER GENERATION

WMCL (F, K);
Input : a set $F = \{f_1, f_2, \ldots, f_v\}$ of fibers
a positive integer $K > 1$ representing the number of clusters to generate
Output: a set $Cl = \{F_1, F_2, \ldots, F_K\}$ of clusters
begin
  $M \leftarrow$ `zeros(v, v)`;
  foreach ***pair*** (f_i, f_j) *s.t.* $f_i \in F,\ f_j \in F$ **do**
    $M[i, j] = MDF(f_i, f_j)$;
  end
  $Cl =$ `KMedoids`(F, M, K);
  return Cl
end

After their computation, each sub fiber-bundle is assigned to a different processor in order to perform the sub fiber-bundle analysis. In this analysis, each processor, independently, computes the CPD for the assigned sub-bundle. The results of this factorization are then used to detect the time points, the fibers and the cross-sections affected by the longitudinal pathological changes using Algorithm 1, Algorithm 2 and Algorithm 3 respectively. The results obtained by each processor are finally merged into different *synchronized* shared variables (Algorithm 5). At the end of this process the algorithm will generate three sets U, P and H representing the time points, the fibers and the cross-sections affected by pathological longitudinal changes respectively.

The parallel execution of the algorithm could be graphically represented as a diagram (Figure 6). Each branch represents the computation flow independently executed by each of the K processors. At the end of the parallel execution, each processor adds its local solution to the complete solution set.

Algorithm 5: PARALLEL ALGORITHM FOR WM ANALYSIS USING TENSOR FACTORIZATION

WMTF (F, ω, K, v, q);
Input : F set containing all the fibers of a bundle, ω LOF threshold parameter,
K number of parallel processors, v total number of fibers, q total number of cross-sections
Output: U set containing the time-points affected by the longitudinal changes
P set of fibers containing longitudinal changes
H set of cross-section containing longitudinal changes.
$U \leftarrow \emptyset$;
$P \leftarrow \emptyset$;
$H \leftarrow \emptyset$;
$F_1, \ldots, F_K \leftarrow$ WMCL(F, K);
foreach $F^{'} \in \{F_1, \ldots, F_K\}$ **do**
  StartNewProcess;
  $\boldsymbol{\mathcal{T}} \leftarrow GenTensor(F^{'})$;
  $R \leftarrow rankest(\boldsymbol{\mathcal{T}})$;
  $\boldsymbol{a_1}, \ldots, \boldsymbol{a_R}, \boldsymbol{b_1}, \ldots, \boldsymbol{b_R}, \boldsymbol{c_1}, \ldots, \boldsymbol{c_R} \leftarrow \min_{\substack{\boldsymbol{a_r} \geq \boldsymbol{0}, \boldsymbol{b_r} \geq \boldsymbol{0}, \boldsymbol{c_r} \geq \boldsymbol{0}, \\ 1 \leq r \leq R}} \|\boldsymbol{\mathcal{T}} - \sum_{r=1}^{R} \boldsymbol{a_r} \circ \boldsymbol{b_r} \circ \boldsymbol{c_r}\|$;
  $Y, U^{'} \leftarrow$ CMPTPDET$(\boldsymbol{c_1}, \ldots, \boldsymbol{c_R}, \omega)$;
  synchronized;
  $U \leftarrow U \cup U^{'}$;
  $P \leftarrow P \cup$ FIBDET$(\boldsymbol{a_1}, \ldots, \boldsymbol{a_R}, Y, v)$;
  $H \leftarrow H \cup$ CSDET$(\boldsymbol{b_1}, \ldots, \boldsymbol{b_R}, Y, q)$;
  EndOfProcess;
end

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

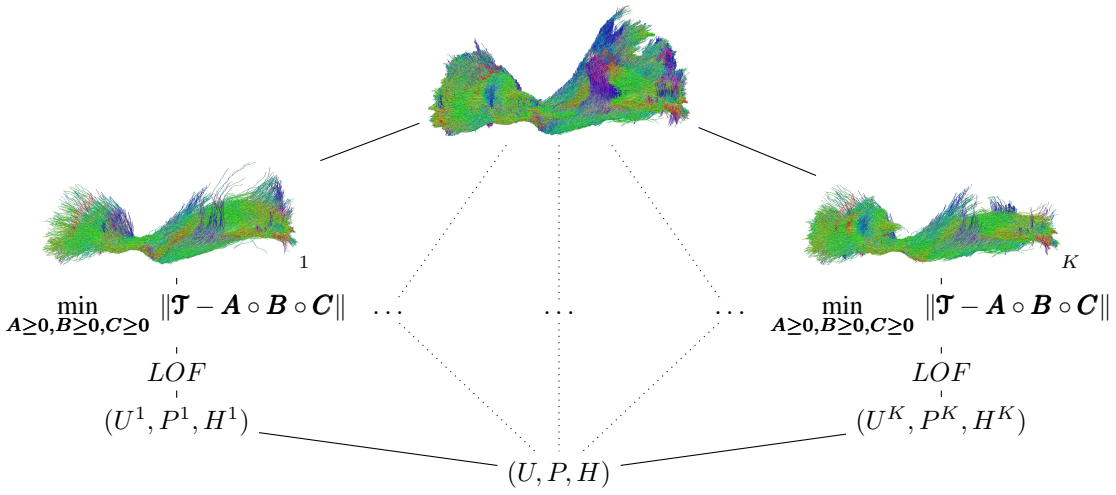

Figure 6. Graphical representation of parallel execution of the proposed algorithm

V. Experiments

A. Subjects

Four relapsing-remitting (RR) MS patients (3 women and 1 man, mean ($\pm$ SD) age: 36.8 $\pm$ 9.5 years; median disease duration: 4.24y; max 16.5 y) (median Expanded Disability Status Scale (EDSS)=2.5, range=[0; 4]) and one healthy control (HC) subject (age: 24 years) were included in this study. Inclusion criteria specified that studied patients were diagnosed as RR MS if they present at least one new Gadolinium-enhancing lesion during the six months preceding study enrolment. All patients had stopped their treatment for at least one year and have not started any during the study period. In order to limit the nephrogenic damage risks associated to Gadolinium injection, creatinine clearance was checked every 2 weeks after inclusion. A clearance higher than 60ml/min was an exclusion criterion. This study was approved by the local ethics committee (CPP Sud-Est IV) and the French national agency for medicine and health products safety (ANSM). Written informed consent was obtained from all patients and the control subject prior to study initiation.

B. MRI protocol

All subjects underwent a weekly examination for a period of two months (8 time-points from T_1 to T_8). MRI protocol included a DTI and a FLAIR acquisition, that were performed on a 3T Philips Achieva system (Philips Healthcare, Best, The Netherlands) with a 16-channels head-coil. The DTI image set consisted of the acquisition of 60 contiguous 2mm-thick slices parallel to the bi-commissural plane (AC-PC), and were acquired using a 2D Echo-Planar Imaging (EPI) sequence (TE/TR = 60/8210 ms, FOV = 224x224x120 mm) with 32 gradient directions ($b = 1000 s.mm^{-2}$). The nominal voxel size at acquisition (2x2x2 mm) was interpolated to 0.875x0.875x2 mm after reconstruction. The FLAIR Vista 3D sequence (TE/TR/TI = 356/8000/2400 ms, FOV=180x250x250 mm) consisted of the acquisition of 576 slices of 0.43 mm thickness.

C. Experiments on simulated longitudinal Variations

As described in [17], a simulation study was performed in order to quantify the performances of the proposed method. More in details, 100 different longitudinal variations were simulated on the diffusion map of a control subject. All the variations were generated along 10 different fiber-bundles, namely, left and right cortico-spinal tract, inferior-fronto occipital fasciculi, cingulum, and forceps major and minor of corpus callosum. In order to simulate the longitudinal variations on diffuse data, the method we proposed in [20], and described in the Appendix, was used. More in detail, we randomly selected several cross-sections and fibers within a specific bundle. For each voxel belonging to these regions, a generalized Gaussian probability density function (GGPDF) with parameters $\mu_\eta, \alpha_\eta, \beta_\eta, \mu_\rho, \alpha_\rho, \beta_\rho$ (see the Appendix for a complete explanation of those parameters) randomly chosen were used to simulate the longitudinal changes. Three different tests were performed: *i)* detection of fibers containing longitudinal changes, *ii)* detection of cross-sections affected by longitudinal alterations and *iii)* identification of abnormal time-points.

Performance measurements used in this work are based on the analysis of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) instances classified during the testing phase. Performances of longitudinal variations were evaluated using accuracy (Ac), precision (Pr) and sensitivity (Se) respectively defined as: $Ac = \frac{TP+TN}{P+N}$, $Pr = \frac{TP}{TP+FP}$ and $Se = \frac{TP}{TP+FN}$. Since multiple tests were performed, mean and standard deviation of accuracy, precision and sensitivity ($\overline{Ac}\pm\sigma_{Ac}$, $\overline{Pr}\pm\sigma_{Pr}$, $\overline{Se}\pm\sigma_{Se}$) were computed.

Best parameter values for *MinPts* and ω were obtained trough grid-search. More in details, we selected a range of values for both parameters and we took the couple with the highest performance. *MinPts* values range from 1 to 8 (the total number of time-points) and the ω interval was given by $\{2, 4, 6, 8, 10, 12\}$.

Generalization of the performances was obtained using a two parts validation method. In the first part, the grid-search, previously described, was performed on simulated variations.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

In the second part, the method was applied on a new set of simulated variations using the best parameters discovered in the first step.

Moreover, in order to show the improvement obtained using tensor factorization, we compared the proposed method with our previous NMF-based method [20]. Since the algorithm proposed in [20] allows just a local cross-sectional analysis of the fiber-bundle, we compared the capability of the two methods to identify only cross-sections and time-points affected by longitudinal pathological changes.

D. Experiments on real MS follow-up data

Four RR MS patients (see section V-A) were selected due to the presence of visible longitudinal alterations assessed by our neuroradiologist (FC). DTI data of each patient were processed using our proposed pipeline described in section II and III. Among the 20 fiber-bundles (contained in the JHU atlas), the Cortico-Spinal Tract (CST), fronto-occipital fasciculus (IFOF) and superior longitudinal fasciculus (SLF) were selected because of the presence of longitudinal MS alterations.

VI. Results

A. Detection of affected fibers, cross-sections and time-points on simulated data

Performances obtained for the detection, in simulated data, of affected fibers within a bundle are reported in table I. The table reports the value of accuracy, precision and sensitivity obtained in the task of the extraction of fibers affected by longitudinal changes. The tests were performed using different diffusion features. Best results were obtained with using as diffusion features $\langle \lambda_2, \lambda_3 \rangle$ with an accuracy, precision and sensitivity respectively of 0.97, 0.79 and 0.40. Satisfying results were also obtained using $\langle \lambda_2, \lambda_3, FA, MD \rangle$ and $\langle \lambda_2, \lambda_3, FA \rangle$ having respectively an accuracy of 0.77 and 0.79, a precision of 0.73 and 0.76 and a sensitivity of 0.20 and 0.23. Worst results were obtained using $\langle KLA, AA \rangle$ and $\langle KLA, FA, AA \rangle$ having respectively an accuracy of 0.71 and 0.77, a precision of 0.69 and 0.69 and a sensitivity of 0.16 and 0.15.

Performances obtained in the detection of affected cross-sections within a bundle are reported in table II. Like the previous case, globally, the best results were achieved using $\langle \lambda_2, \lambda_3 \rangle$ as feature set with an accuracy, precision and sensitivity respectively of 0.63, 0.98 and 0.95. The features set composed by $\langle \lambda_2, \lambda_3, FA \rangle$ show low values in terms of accuracy (0.40) and sensitivity (0.56) but the highest level of precision (1.00). Like for the tests performed for the detection of fibers affected by longitudinal changes, worst results were achieved by the features set $\langle KLA, AA \rangle$ and $\langle KLA, FA, AA \rangle$. Moreover, in table II, we show the performances obtained with the tensor-based and NMF methods to detect cross-sections, of a fiber-bundle, affected by pathological longitudinal changes. From the table it is possible to see how the NMF method method always outperforms tensor factorization in accuracy. More in detail, NMF reaches the best accuracy (0.74) with $\langle \lambda_2, \lambda_3, FA \rangle$ and $\langle \lambda_2, \lambda_3, FA, MD \rangle$. Similar results are visible in sensitivity, indeed NMF always outperforms tensor factorization except for the feature set $\langle \lambda_2, \lambda_3 \rangle$ where tensor factorization reaches the best performance. Contrarily, for the accuracy, the tensor factorization outperforms, in all the different diffusion features used, the NMF method.

Performances obtained in the detection of time-points affected by longitudinal changes are reported in table III. Best results were achieved using $\langle \lambda_2, \lambda_3 \rangle$ as features with an accuracy, precision and sensitivity respectively of 0.84, 0.93 and 0.96. Contrary to the previous case, the second best features set is $\langle \lambda_2, \lambda_3, FA, MD \rangle$ having accuracy, precision and sensitivity respectively of 0.73, 0.87 and 0.81. In this case $\langle \lambda_2, \lambda_3, FA \rangle$ globally shows results comparable to $\langle KLA, AA \rangle$ and $\langle KLA, FA, AA \rangle$. The largest difference is in the accuracy who reaches 0.60 for $\langle \lambda_2, \lambda_3, FA \rangle$ and 0.50 for both $\langle KLA, AA \rangle$ and $\langle KLA, FA, AA \rangle$. Moreover, as additional experiment, in table III we compared the performances in detection of time-points affected by longitudinal changes obtained using the proposed tensor factorization algorithm with the NMF method. From the table it is possible to see how the proposed tensor factorization algorithm always outperforms the NMF method in all the performance measurements and in all the combination of features used in the experiment. Regarding the parameters used to perform the analysis used by the tensor factorization algorithm, the best results were achieved with $MinPts = 3$ and $\omega = 8$. Accuracy performances with different values of $MinPts$ and ω are reported in Table IV.

Table I
Results of mean and standard deviation (in parenthesis) of accuracy, precision and sensitivity for detection of longitudinal affected fibers using different combination of diffusion parameters.

	Accuracy	Precision	Sensitivity
λ_2,λ_3	**0.97 (0.06)**	**0.79 (0.08)**	**0.40 (0.11)**
λ_2,λ_3, FA	0.77 (0.29)	0.73 (0.11)	0.20 (0.02)
λ_2,λ_3, FA, MD	0.79 (0.32)	0.76 (0.15)	0.23 (0.05)
KLA,AA	0.71 (0.23)	0.69 (0.13)	0.16 (0.01)
KLA,FA,AA	0.77 (0.20)	0.66 (0.15)	0.15 (0.04)

Table II
Results obtained using tensor and non-negative matrix factorization. Results are reported with mean and standard deviation (in parenthesis) of accuracy, precision and sensitivity for detection of longitudinal affected cross-sections using different combination of diffusion parameters.

	Tensor Factorization			Non-Negative Matrix Factorization		
	Accuracy	Precision	Sensitivity	Accuracy	Precision	Sensitivity
λ_2,λ_3	0.63 (0.10)	0.98 (0.02)	**0.95 (0.15)**	0.65 (0.12)	0.60 (0.08)	0.92 (0.10)
λ_2,λ_3, FA	0.40 (0.12)	**1.00 (0.01)**	0.56 (0.12)	**0.74 (0.11)**	0.68 (0.04)	0.90 (0.09)
λ_2,λ_3, FA, MD	0.47 (0.13)	0.97 (0.03)	0.74 (0.23)	**0.74 (0.08)**	0.68 (0.12)	0.89 (0.11)
KLA,AA	0.30 (0.10)	0.96 (0.03)	0.64 (0.30)	0.59 (0.07)	0.55 (0.12)	0.91 (0.10)
KLA,FA,AA	0.32 (0.12)	0.96 (0.03)	0.63 (0.28)	0.56 (0.11)	0.53 (0.09)	0.85 (0.05)

Table III
Results obtained using tensor and non-negative matrix factorization. Results are reported with mean and standard deviation (in parenthesis) of accuracy, precision and sensitivity for detection of longitudinal affected time-points using different combination of diffusion parameters.

	Tensor Factorization			Non-Negative Matrix Factorization		
	Accuracy	Precision	Sensitivity	Accuracy	Precision	Sensitivity
λ_2,λ_3	**0.84 (0.12)**	**0.93 (0.09)**	**0.96 (0.13)**	0.72 (0.30)	0.86 (0.18)	0.78 (0.31)
λ_2,λ_3, FA	0.60 (0.20)	0.76 (0.10)	0.54 (0.16)	0.57 (0.36)	0.75 (0.17)	0.53 (0.35)
λ_2,λ_3, FA, MD	0.73 (0.30)	0.87 (0.18)	0.81 (0.32)	0.67 (0.36)	0.84 (0.18)	0.78 (0.33)
KLA,AA	0.50 (0.30)	0.71 (0.21)	0.82 (0.30)	0.50 (0.36)	0.67 (0.19)	0.56 (0.37)
KLA,FA,AA	0.50 (0.32)	0.74 (0.19)	0.75 (0.20)	0.50 (0.36)	0.66 (0.19)	0.57 (0.34)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

B. Evaluation of parallel implementation

In order to underline the difference between serial and parallel implementation, we show the computational time obtained by using the two implementations. More in detail, the implementations were run 10 times on the same task and the computational time (in terms of mean and standard deviation) were measured. For the parallel implementation the same tasks were executed with different parallel processors ($2 \leq K \leq 7$). Results in terms of execution time are reported in Figure 7. Since the serial execution time does not depend on the number of parallel process it is always stable. An important decrease in computational time is already clearly visible starting with $K = 2$ where the serial implementation takes 899.8 ± 14.0 seconds while the parallel one takes 550.3 ± 10.3 seconds. We define with speed-up factor the ratio between the mean computational time of the serial implementation and the parallel implementation. Already with $K = 2$, the speed-up factor, is 1.63. Moreover, it is possible to see how by increasing the number of processors the computational time decreases. The minimum and maximum speed-up factors are 1.63 and 2.6 reached for $K = 2$ and $K = 4$ respectively. Results show how the improvement in computational performance is clear using the parallel implementation.

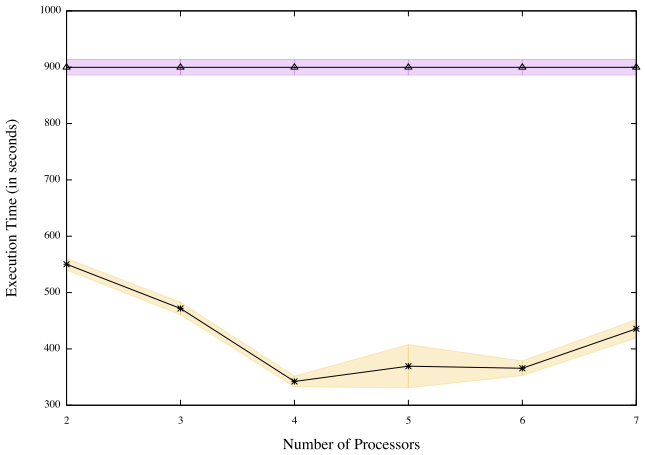

Figure 7. Mean and standard deviation, computed from the 10 different runs, of the computation time for Serial ($\triangle \pm$) and Parallel ($* \pm$) implementations.

C. Detection of affected fibers, cross-sections and time-points on real data

The proposed method was also tested on real data. Based on the results obtained in the previous sections, $MinPts = 3$, $\omega = 8$ and $K = 4$ were selected as parameters and a feature vector composed by $\langle \lambda_2, \lambda_3 \rangle$ ($M = 2$) was used.

Results of application on MS real data are illustrated in Figure 8. In all the figures, the mean FA signal profile along the fibers, within the bundle, detected by our method as affected by longitudinal variations are reported. Time-points detected as containing longitudinal changes are reported with the symbol "*" in the name. Cross-sections detected as affected by those changes are underlined with black lines.

In Figure 8 A and Figure 8 B we reported the FA signal profile of two different MS subjects. The signal profile refers to the CST in the subset of fibers detected as containing longitudinal pathological changes. It is possible to see how in those subsets of fibers, in the time-points and in the cross-sections detected as longitudinal changed, the FA signal goes down due to the presence of small longitudinal MS alterations.

Same behaviour is also visible in the other two MS subjects reported in Figure 8 C and Figure 8 D respectively. The former represents the IFOF In the cross-sections and in the time-points identified by our method, longitudinal changes in FA are visible. More in detail, the FA signal goes down in the detected cross-sections due to the presence of MS lesions. In Figure 8 D, we reported the SLF. Like for the other fiber-bundles, in the subset of fibers and cross-sections identified as affected by longitudinal changes, alterations in FA signal profile are visible.

VII. Discussion

In this work we described a new tensor-based method to automatically analyze longitudinal changes in WM fiber-bundles of MS patients. More in detail, we provided a complete pipeline capable to automatically extract fiber-bundles and register longitudinal data to a common template. As major message, we formalized the problem using a tensor-based formalism to detect local scale longitudinal variations caused by rapid inflammatory processes in MS patients.

Moreover, we improved the computational speed of our method by providing a parallel implementation of the algorithm. We used the "divide-et-impera" paradigm to subdivide the main problem in sub-problems. We then merged the sub-solutions of each sub-problem into an unique final solution. In order to perform a better parallelization, we split the fiber-bundle in small sub-bundles using a modified version of K-medoids. As first step, a global distance matrix, based on the MDF metric, was computed between each pair of fibers within a bundle. The obtained matrix was then used as distance matrix for the K-medoids algorithm. The tensor decomposition pipeline was then applied independently to the data extracted in each sub-bundle. The results obtained by the factorization in each sub-problem were then merged in an unique solution. The decrease in computation time, between serial and parallel implementation, is well described in the test we made (Section VI-B). In particular using 4 parallel processors ($K = 4$) we got a speed-up factor of 2.6 (from 899.8 seconds to 342.2 seconds). It should be noted that the number K of parallel processors mainly depends on two factors: *i)* the number of parallel processors available on the machine used to run the algorithm, *ii)* the number of fibers within the bundle. Indeed, if K is large (thus the value $\lceil \frac{v}{K} \rceil$ is small) the number of fibers in each process could not be enough to compute an accurate tensor decomposition.

Results on simulated data, generated using GGPDF [20], show the capability of our method to detect and delineate the presence of longitudinal changes in both fibers and cross-sections of WM fiber-bundles. The experiment section (Section V) was also enriched by the analysis of the features used to compute the factorization of the longitudinal data. The diffusion features used to perform the analysis play also an

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

important role on the final results. Best results were obtained using $\langle \lambda_2, \lambda_3 \rangle$ as features. The introduction of other diffusion parameters does not drastically improve the performances.

As final experimental step, the proposed method was applied to real MS patients showing its capability to detect and delineate time-points, fibers and cross-sections identified, by the neuroradiologist, as affected by longitudinal changes. With these experiments, we showed that also on real data it is possible to identify regions containing longitudinal pathological changes. This gave us the possibility to extract two types of information: *i)* segmentation of the regions (fibers and cross-sections) affected by the longitudinal variations and *ii)* the delineation of the time-points affected by the longitudinal variations. Detection of longitudinal changes were validated by our neuroradiologist (FC).

An important limitation is derived by the low performances, in terms of sensitivity, obtained in the extraction of the fibers affected by longitudinal changes. The low performances could be explained by the CPD model we applied to perform the factorization. Indeed, CPD uses only rank-1 terms to factorize the tensor. With this restriction, representation of biological variations typical of MRI data could not be perfectly estimated. It is important to observe that, in our reference context, Precision should be privileged over Sensitivity because the number of fibers generated by tractography algorithms usually does not reflect the number of real fibers of a human brain.

Another crucial point of investigation is related to the rank selection. Indeed, in this work we used a "brute force" way to compute the rank using the *rankest* function provided in Tensorlab [23]. This process usually tends to overestimate the real number of components and, as results, noise or data variability can be modelled by the factorization. The main consequence related to this overestimation is a decrease in performances obtained in detection of "pathological" components. Other approaches, based on hyper-parameter optimization, such as the Tree-of-Parzen-Estimators (TPE) algorithm [1], can also be used to find the optimal rank thereby reducing the computational time. This problem points out that further improvements are expected by developing a problem-specific rank estimator. This is part of future work.

Moreover, it should be noted that the proposed method was applied only on a few real-life MS patients recording. A more extensive study including more large-scale experiments on larger patient databases is need to obtain more solid conclusions. As future work, we plan to test our method also on a larger dataset with more patients and with a longer time interval between two scans.

Since the proposed method is completely unsupervised and model-free, it is capable to perform the analysis also including more time points or modalities. Indeed, since the values acquired in a new time-point are simply stacked in the 3^{rd} mode of the tensor, no modification is needed in the algorithm. Furthermore, this framework is easy to extend, features derived from other modalities (like T1, T2, etc...) can be concatenated in the 3^{rd} mode of the tensor (as described in Figure 3) without requiring additional work.

The proposed method can be seen as a general framework capable to extend our previous method based on the NMF [20]. Indeed, due to the capability of the tensor to represent high dimensional data, fiber-bundles can be analyzed globally without generating a local view (like in cross-sectional analysis). In the comparison we performed with the NMF method, we saw that the better performances, in cross-section analysis, were obtained by the NMF method compared to the tensor factorization. These results are due to the local analysis performed by NMF. Indeed, this method analyzes each cross-section independently performing a specific local analysis focused on cross-section. In the comparison of the performances obtained in time-point detection, the proposed tensor factorization algorithm outperformed the NMF algorithm. This shows how analysis of multiple information at the same time, really helps to improve the detection of small longitudinal changes in WM fiber-bundles.

Moreover, compared to [19], [17], [6], the proposed method allows to *i)* analyze multiple DTI features taking into account more than 2 time-points *ii)* delineate regions, fibers and detect time-points affected by longitudinal changes.

VIII. Conclusion

In this work, we proposed a new tensor-based framework for the analysis of longitudinal changes occurring along WM fiber-bundles. We described how constrained tensor factorization is a potential tool to analyze multi-dimensional data that can not be fully described using a simple matrix representation.

In order to reach our goal, two main challenges related to tensor factorization were analyzed and solved, namely rank selection and computational time for the factorization. The former was solved by estimating L-curve error of the CPD with increasing number of rank-one terms. The latter was improved by splitting the problem in sub-tasks using the "divide et impera" paradigm.

The method was tested and validated on simulated data and real data. Moreover the performances obtained with the proposed method were also compared with the NMF-based method proposed in [20]. We generated simulated data using our previously proposed simulation paradigm [20]. For the real data, we used a dataset containing MS subjects affected by small longitudinal changes visible in a short (weekly) interval. According to the results obtained in this work, the method we proposed will offer a valuable tool for longitudinal data analysis in neurodegenerative diseases.

Multiple future improvements of the proposed method are planned. First of all, we will improve the validation set by adding multimodal data. Indeed, we plan to apply the NTF algorithm on datasets containing different MRI modalities like MRS, T2, T1, etc.. We also plan to improve the detection of cross-sections affected by longitudinal changes by introducing regularization for the component matrix representing cross-sections. Moreover, as we reported in Section VII, CPD could be too restrictive for some applications as it does not model all variability in the data [8]. We plan to increase the performance, especially for the cross-section component, using block term decomposition (BTD) instead of CPD. Indeed, using a BTD model, it is possible to model the variability on the data by performing a so-called rank $(L_r, L_r, 1)$ BTD [8].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

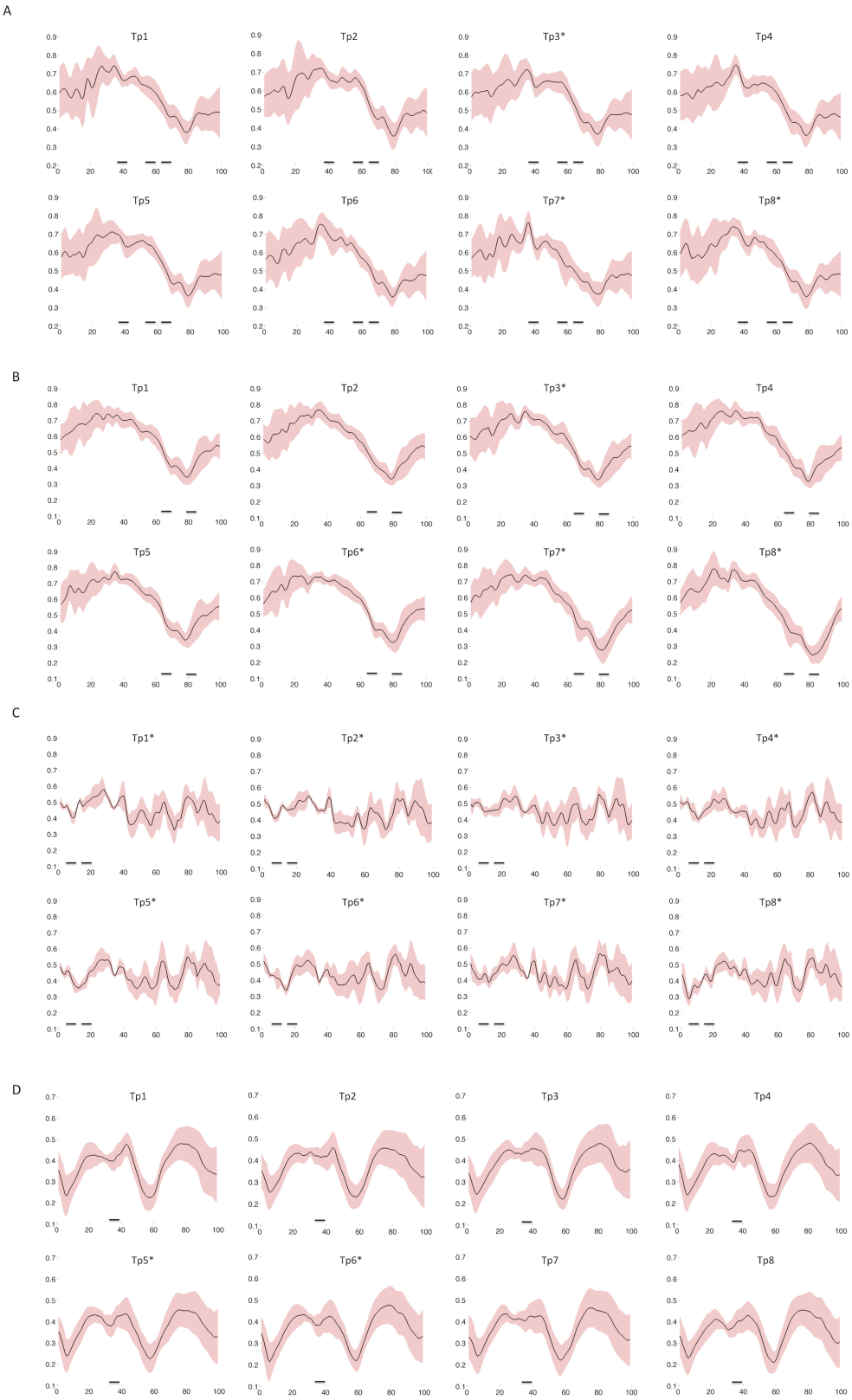

Figure 8. Mean FA signal profile of **A)** cortico-spinal tract Subject 1, **B)** cortico-spinal tract Subject 2, **C)** fronto-occipital fasciculus Subject 3 and **D)** superior longitudinal fasciculus of Subject 4 along the subset of fibers identified by our method as affected by longitudinal changes. Time-points detected as containing longitudinal changes, detected by our method, contain the symbol "*" in the name. Cross-sections detected as affected by longitudinal pathological changes are underlined with black lines.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2019.2933138, IEEE Journal of Biomedical and Health Informatics

APPENDIX

A. Longitudinal Variations Simulator (LVS)

In this section we briefly describe the model we used to generated longitudinal variations introduced in [20].

The model assumes that in different time-points new variations appear in the image. Each of those variations can change both shape and voxel values.

We consider that the variation takes a spherical shape of radius η that could change on time. Moreover, we assume that diffusion values of voxels belonging to this region could change during the longitudinal evolution of a factor ρ, called reduction coefficient. This coefficient ρ is used to change the voxel's diffusion values inside the spherical region in a given time-point according to the following equations:

$\lambda_2^* = \lambda_2 + \rho * (\lambda_1 - \lambda_2) \quad \lambda_3^* = \lambda_3 + \rho * (\lambda_1 - \lambda_3)$

with $0 \leq \rho \leq 1$. FA, MD, and other tensor metrics are then recomputed in the given time-point using λ_1 and the new λ_2^*, λ_3^* values. The proposed LVS model could be summarized using the parametric function $\mathcal{S}$ defined as:

$$\mathcal{S} : \begin{cases} \eta(t) = \mathcal{G}(t, \mu_\eta, \alpha_\eta, \beta_\eta) \\ \rho(t) = \mathcal{G}(t, \mu_\rho, \alpha_\rho, \beta_\rho) \end{cases}$$

the function $\mathcal{G}$ is the generalized Gaussian probability density function (GGPDF) defined as: $\mathcal{G}(x, \mu, \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-(\frac{|x-\mu|}{\alpha})^\beta}$

where $x, \mu, \alpha, \beta \in \mathbb{R}$ with $\alpha, \beta \geq 0$ and Γ denotes the gamma function. This particular distribution includes the normal Gaussian distribution ($\mathcal{N}$) when $\beta = 2$ (with mean μ and variance $\frac{\alpha^2}{2}$) and it includes also the Laplace distribution when $\beta = 1$.

B. Parameter selection for $MinPts$ and ω

In order to describe how ω and $MinPts$ parameters can affect the performances of the proposed method, we reported in table IV how accuracy performances change for detection of longitudinal affected fibers using $\langle \lambda_2, \lambda_3 \rangle$ as diffusion parameters.

The best results are obtained when both $MinPts$ and ω parameters have values in range $1 \leq MinPts \leq 3 \wedge 6 \leq \omega \leq 10$. Values outside this range show low level of accuracy. More in details, large values of $MinPts$ degrade the accuracy too much and are therefore not recommended.

Table IV
MEAN ACCURACY FOR DETECTION OF LONGITUDINAL AFFECTED FIBERS USING λ_2, λ_3 AS DIFFUSION PARAMETERS.

ω \ *MinPts*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
2	0.59	0.62	0.70	0.80	0.65	0.60	0.50	0.33
4	0.66	0.72	0.75	0.81	0.72	0.66	0.60	0.50
6	0.70	0.87	0.82	0.90	0.77	0.70	0.60	0.55
8	0.80	0.90	0.97	0.94	0.91	0.80	0.68	0.60
10	0.77	0.80	0.91	0.85	0.75	0.70	0.61	0.52
12	0.76	0.77	0.84	0.86	0.76	0.65	0.58	0.40

REFERENCES

[1] J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, "Algorithms for Hyper-Parameter Optimization", NIPS'2011, 2011.

[2] M.M. Breunig, H.-P. Kriegel, R.T. Ng, J. Sander, "LOF: Identifying density-based local outliers," ACM SIGMOD, vol. 29, pp. 93-104, 2000.

[3] J. L. Castellanos, S. Gomez, V. Guerra, "The triangle method for finding the corner of the L-curve," Applied Numerical Mathematics, vol. 43(4), pp. 359-373, 2002.

[4] E. Garyfallidis, M. Brett, M. M. Correia, G. B. Williams, I. Nimmo-Smith, "Quickbundles, a method for tractography simplification," Front Neurosci, vol. 6: 175, 2012.

[5] E.R. Gerstner, A.G. Sorensen, "Diffusion and diffusion tensor imaging in brain cancer," Semin Radiat Oncol., vol. 21, pp. 141-146, 2011.

[6] A. Grigis, V. Noblet, F. Blanc, F. Heitz, J. de Seze, S. Kremer, J. P. Armspach, "Longitudinal change detection: inference on the diffusion tensor along white matter pathways," Med Image Anal, vol. 17, pp. 375-386, 2013.

[7] K. Hua, J. Zhang, S. Wakana, H. Jiang, X. Li, D. S. Reich, "Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification," Neuroimage, vol. 39, pp. 336-347, 2008.

[8] B. Hunyadi, D. Camps, L. Sorber, W. Paesschen, M. De Vos, S. Huffel, L. De Lathauwer, "Block term decomposition for modelling epileptic seizures," EURASIP J. Adv. Signal Process. vol. 139, 2014.

[9] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith, "FSL," NeuroImage, vol. 62, pp. 782-790, 2012.

[10] S. Keihaninejad, H. Zhang, N. S. Ryan NS, I. B. Malone, M. Modat, M. J. Cardoso, D. M. Cash, N. C. Fox, S. Ourselin, "An unbiased longitudinal analysis framework for tracking white matter changes using diffusion tensor imaging with application to Alzheimer's disease," Neuroimage, vol. 72, pp. 153-163, 2013.

[11] G. Kocevar, C. Stamile, S. Hannoun, F. Cotton, S. Vukusic, F. Durand-Dubief, D. Sappey-Marinier, "Graph Theory-Based Brain Connectivity for Automatic Classification of Multiple Sclerosis Clinical Courses," Front. Neurosci., doi: 10.3389/fnins.2016.00478, 2016.

[12] T. G. Kolda, B. W. Bader, "Tensor Decompositions and Applications," SIAM Review, vol. 51(3), pp. 455-500, 2009.

[13] J. MacQueen, "Some methods for classification and analysis of multivariate observations," Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1(14), 1967.

[14] F. Prados, I. Boada, A. Prats-Galino, J. A. Martín-Fernández, M. Feixas, G. Blasco, J. Puig, S. Pedraza, "Analysis of New Diffusion Tensor Imaging Anisotropy Measures in the Three-Phase Plot," JMRI, vol. 31, pp. 1435-1444, 2010.

[15] M. Reuter, N. J. Schmansky, H. D. Rosas, B. Fischl, "Within-subject template estimation for unbiased longitudinal image analysis," Neuroimage, vol. 61(4), pp. 1402-1418, 2012.

[16] L. Sorber, M. VanBarel, L. De Lathauwer, "Structured data fusion," Selected Topics in Signal Processing, vol. 9(4), pp. 586-600, 2015.

[17] C. Stamile, G. Kocevar, F. Cotton, D. Sappey-Marinier, "A genetic algorithm-based model for longitudinal changes detection in white matter fiber-bundles of patient with multiple sclerosis," Computers in Biology and Medicine, vol. 84, pp. 182-188, 2017.

[18] C. Stamile, F. Cauteruccio, G. Terracina, D. Ursino, G. Kocevar, D. Sappey-Marinier, "A Model-Guided String-Based Approach to White Matter Fiber-Bundles Extraction," BIH 2015: Brain Informatics and Health, LNCS 9250, pp. 135-144, 2015.

[19] C. Stamile, G. Kocevar, F. Cotton, F. Durand-Dubief, S. Hannoun, C. Frindel, D. Rousseau, D. Sappey-Marinier, "A Sensitive and Automatic White Matter Fiber Tracts Model for Longitudinal Analysis of Diffusion Tensor Images in Multiple Sclerosis," PLoS ONE 11(5): e0156405. doi:10.1371/journal.pone.0156405, 2016.

[20] C. Stamile, G. Kocevar, F. Cotton, F. Maes, D. Sappey-Marinier, S. Van Huffel, "Multiparametric Nonnegative Matrix Factorization for Longitudinal Variations Detection in White-Matter Fiber Bundles," IEEE Journal of Biomedical and Health Informatics, vol. 21(5), pp. 1393-1402, 2017.

[21] J. D. Tournier, F. Calamante, A. Connelly, "MRtrix: Diffusion tractography in crossing fiber regions," International Journal of Imaging Systems and Technology, vol. 22, pp. 53-66, 2012.

[22] A. Varentsova, S. Zhang, K. Arfanakis, "Development of a high angular resolution diffusion imaging human brain template," Neuroimage, vol. 91, pp. 177-186, 2014.

[23] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, L. De Lathauwer, "Tensorlab 3.0," Available online, Mar. 2016. URL: http://www.tensorlab.net/