Data-driven separation of MRI signal components for tissue characterization

Rahbek, Sofie; Madsen, Kristoffer H.; Lundell, Henrik; Mahmood, Faisal; Hanson, Lars G.

Published in:
Journal of Magnetic Resonance

Link to article, DOI:
10.1016/j.jmr.2021.107103

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Journal of Magnetic Resonance 333 (2021) 107103

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

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

Data-driven separation of MRI signal components for tissue characterization ☆

Sofie Rahbek [a], Kristoffer H. Madsen [b,c], Henrik Lundell [b], Faisal Mahmood [d,e], Lars G. Hanson [a,b,*]

[a] Department of Health Technology, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
[b] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, 2650, Denmark
[c] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
[d] Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital, Odense C 5000, Denmark
[e] Department of Clinical Research, University of Southern Denmark, Odense 5000, Denmark

ARTICLE INFO

Article history:
Received 8 March 2021
Revised 14 October 2021
Accepted 2 November 2021
Available online 5 November 2021

Keywords:
Magnetic resonance imaging
non-negative matrix factorization
Data-driven decomposition
Monotonous slope
Tissue characterization
Diffusion
Relaxometry

ABSTRACT

Purpose: MRI can be utilized for quantitative characterization of tissue. To assess e.g. water fractions or diffusion coefficients for compartments in the brain, a decomposition of the signal is necessary. Imposing standard models carries the risk of estimating biased parameters if model assumptions are violated. This work introduces a data-driven multicomponent analysis, the monotonous slope non-negative matrix factorization (msNMF), tailored to extract data features expected in MR signals.
Methods: The msNMF was implemented by extending the standard NMF with monotonicity constraints on the signal profiles and their first derivatives. The method was validated using simulated data, and subsequently applied to both ex vivo DWI data and in vivo relaxometry data. Reproducibility of the method was tested using the latter.
Results: The msNMF recovered the multi-exponential signals in the simulated data and showed superiority to standard NMF (based on the *explained variance*, *area under the* ROC *curve*, and *coefficient of variation*). Diffusion components extracted from the DWI data reflected the cell density of the underlying tissue. The relaxometry analysis resulted in estimates of edema water fractions (EWF) highly correlated with published results, and demonstrated acceptable reproducibility.
Conclusion: The msNMF can robustly separate MR signals into components with relation to the underlying tissue composition, and may potentially be useful for e.g. tumor tissue characterization.

1. Introduction

Quantitative assessment of tissue properties has become possible with specific MRI sequences that enable extraction of parameters such as T_1 and edema water fraction. With a quantitative characterization of the tissue, it is possible not only to compare data across scans and subjects, but also to capture changes at sub-voxel scale [1,2]. This can potentially improve our understanding of pathophysiology and generate biomarkers for e.g. abnormality detection, disease staging or treatment response assessment. However, probing tissue parameters requires acquisition of high-dimensional datasets and appropriate parameterization. A well-known example is the calculation of the apparent diffusion coefficient (ADC) from a set of pulsed gradient spin-echo scans with varying diffusion weighting (b-value). The ADC describes the mean diffusivity of water in a voxel and is obtained by fitting the diffusion-weighted signal to a mono-exponential model, as expressed in the Stejskal-Tanner equation [3]. Important factors modulating the diffusivity are the volume of extracellular water with a less restricted mobility [4], the size of restricting domains such as cells, and further, the orientation of elongated restrictions which makes the diffusivity directionally dependent. However, a voxel can contain multiple tissue compartments (partial volume effect) or orientational dispersion of anisotropic domains, and the diffusion can be affected by water exchange between compartments [5,6]. Furthermore, biophysical processes such as flow and perfusion can cause extra signal decay [7,8]. Hence, tissues exhibit non-exponential signal decay if a broad range of b-values is used. Several more complex models have been suggested and used to improve fitting, e.g. the

☆ **Funding information** Danish Cancer Society, Grant Number: R167-A10637-17-S2; European Union's Horizon 2020 research and innovation programme, Grant agreement No 804746.
* Corresponding author at: Technical University of Denmark, Ørsteds Plads, Building 349, Kgs. Lyngby 2800, Denmark.
E-mail address: lghan@dtu.dk (L.G. Hanson).

https://doi.org/10.1016/j.jmr.2021.107103
1090-7807/

intravoxel incoherent motion model or water exchange models [6]. Imposing models, however, always carries the risk of estimating biased or uninformative parameters, because the model may be misleading, or too restricted or flexible considering the underlying signal variation [8,9]. Nevertheless, decomposition of the signal into multiple compartment-related components is desired for tissue characterization, which is an essential challenge that is repeatedly encountered in quantitative analysis of MRI data. Mapping parameters for relaxometry or magnetic resonance fingerprinting, for example, faces the ill-posed problem of decomposing a signal into a weighted sum of exponentials or other latent factors [1,7]. A class of methods designed to solve this ill-posed inversion, is regularized voxel-wise spectral fitting [10–12], which in the case of multi-exponential decay is also referred to as regularized inverse Laplace transforms [13]. A frequently used version is the non-negative least squares (NNLS) technique [14,15] which has proven useful, e.g. for estimation of myelin water fractions (MWF) [14,10,16]. The regularization is necessary to handle the inevitable large impact of noise in a voxel-wise fitting regime [17], but requires the choice of a penalty parameter. Though a cross-validated or data-driven estimation of this is possible [16], it may introduce a bias in quantitative estimates due to the well known bias/variance trade-off. Additionally, the method relies on an identification of spectral clusters or averaging across regions of interest (ROIs) requiring prior knowledge of the tissue homogeneity [17,18].

Another group of strategies avoiding these limitations and presumptive physiological models is blind source separation (BSS) techniques, including widely used unsupervised linear dimensionality reduction (LDR). Here, all measurements are analyzed simultaneously and represented as a product of two factors; a matrix of fundamental basis functions and a matrix of the associated spatial distributions [9,19,20]. An attractive feature of LDR techniques is their insensitivity to partial volume effects often hampering traditional signal modeling. What sets the different LDR techniques apart is the assumptions on the structure of the factors. The well-known principal components analysis yields orthogonality between the principal components [21–23], while the independent component analysis, often used for separation of functional MRI signals, ensures independence between them [19,24]. While these methods may result in low-error reconstructions of the data since the underlying singular value decomposition is guaranteed to produce the minimum error for a given dimensionality reduction [25], the components may be physically unrealistic and difficult to interpret. For medical image intensities, a natural assumption is component-wise non-negativity leading to the LDR technique, non-negative matrix factorization (NMF). This method is popular within a broad range of applications ranging from text mining to image processing because the non-negativity assumption promotes an attractive part-based representation with more realistic factors, although the general non-uniqueness of NMF solutions can render the interpretation of the estimated components difficult [19,26–28]. In general, NMF is non-convex and the estimated solution is often sensitive to initialization and small perturbations of the input data. In practice, a common strategy to tackle these issues is to use priors on the factors or regularization during optimization [19,27,29].

This work proposes a novel extension of the NMF by introducing additional constraints on the structure of the factors, which tailors the method for extraction of multi-exponential or similar signals, expected for many phenomena both within and beyond the field of MRI. This loosely defined class of functions is called *approximately multi-exponential* in the following. The extension is inspired by Bhatt and Ayyar who introduced the monotonous NMF [30], mNMF. They enforced monotony of NMF components, which in practise give results similar to NMF for a range of MR signals and is insufficient to enforce realistic signal behaviour (see Supplementary Material). Hence, we introduce a method, which constrains the problem further. It imposes two conditions on the non-negative basis functions: they should be monotonous and have a monotonous first derivative (slope). Thus, the proposed method is referred to as *monotonous slope non-negative matrix factorization* (msNMF). It is expected that the new constraint on the first derivative makes msNMF superior to the mNMF in regards to separation of mixed MRI signal components. Yet, the method leaves more freedom to the signal components compared to the spectral fitting techniques where the decomposition is within a pre-determined multi-exponential basis constraining also higher order derivatives. In more general terms, the additional constraints enforce prior expectations of MR signal decay and represents a balance between model-based analysis that may rely on inadequate model assumptions, and model-free analysis that may be too flexible to give interpretable results.

We introduce and demonstrate the msNMF for multi-component analysis of MRI data, and show examples where the estimated components are related to the underlying tissue structure, thus demonstrating the use of the method as a robust alternative to advanced modeling.

2. Methods

2.1. Non-negative matrix factorization

The NMF approximates a given data matrix, $\mathbf{X} \in \mathbb{R}^{m\times n}$, with the low rank matrix, $\mathbf{WH}$:

$$\mathbf{X} = \mathbf{WH} + \mathbf{E} \tag{1}$$

$\mathbf{W} \in \mathbb{R}_{+}^{m\times k}$ is a set of non-negative basis functions, $\mathbf{H} \in \mathbb{R}_{+}^{k\times n}$is the associated non-negative weights for the linear combination of basic functions to reconstruct the data, and $\mathbf{E} \in \mathbb{R}^{m\times n}$is an error matrix accounting for noise and non-factorizable signals. With the assumption of Gaussian i.i.d. residuals in $\mathbf{E}$, the Frobenius norm is used as objective function, $\|\mathbf{X} - \mathbf{WH}\|_F^2$. This is a non-convex optimization problem. However, using the two-block coordinate descent framework known as the alternating non-negative least squares (ANLS) algorithm, it can be formulated as two convex sub-problems:

$$\min_{\mathbf{W}\geqslant 0} \|\mathbf{X} - \mathbf{WH}\|_F^2 \quad \text{and} \quad \min_{\mathbf{H}\geqslant 0} \|\mathbf{X} - \mathbf{WH}\|_F^2 \tag{2}$$

The problem is solved by alternately re-estimating one of the two factors, $\mathbf{W}$ or $\mathbf{H}$, under a non-negativity constraint, while keeping the other fixed.

In this study demonstrating analysis of volumetric 4D MRI datasets with m measurements and n signal-carrying voxels, the k basis functions in $\mathbf{W}$ are referred to as *signal components*. The weights in $\mathbf{H}$, specifying how these are mixed in each voxel, are referred to as *mixture maps*.

2.2. Monotonous slope non-negative matrix factorization

The idea of introducing constraints that enforce monotonicity in the NMF emerges from expectations for noise-free MR signal behaviour. It can be shown, that the group of mono- and multi-exponential functions (described by: $b_0 + \sum_{n=1}^{N} b_n \cdot e^{a_n x}$) are monotonous, so the continuous derivative has no zeros, provided the exponents a_n have a common sign, and the arbitrary number of non-zero coefficients b_n also share sign. The offset b_0 has no impact. The derivatives fall within the same class of functions, and are therefore also monotonous. In unusual specific cases involving pronounced intermediate water exchange, mostly in

combination with inversion recovery, the signal no longer complies with the assumption of shared sign coefficients b_n, but normal multi-exponential signal decay and longitudinal magnetization recovery fulfills this requirement adopted in the following [31–33].

The two constrained least squares problems in (2) can easily be restated in dual form to allow optimization via standard quadratic programming. Eq. 3 shows the general case of such reformulation:

$$\min_{\mathbf{x}} \tfrac{1}{2}\|\mathbf{Rx}-\mathbf{d}\|^2 \longrightarrow \min_{\mathbf{x}} \tfrac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Px}+\mathbf{f}^{\mathrm{T}}\mathbf{x} \quad \text{subject to}: \ \mathbf{Gx} \leqslant \mathbf{a} \qquad \text{where } \mathbf{P}=\mathbf{R}^{\mathrm{T}}\mathbf{R} \text{ and } \mathbf{f}=-\mathbf{R}^{\mathrm{T}}\mathbf{d} \tag{3}$$

In the optimization problem at hand, $\mathbf{x}$ is a vectorization of either $\mathbf{W}$ or $\mathbf{H}$. The inequality constraint can be revised to include both the non-negativity and monotonicity constraints. This approach was inspired by the work of Bhatt and Ayyar [30] and results in the following reformulated optimization with the appropriate constraints added:

$$\text{Given } \mathbf{W}: \quad \min_{\mathbf{H}} \ \mathrm{vec}(\mathbf{H})^{\mathrm{T}}(\mathbf{I}_n \otimes \mathbf{W}^{\mathrm{T}}\mathbf{W})\mathrm{vec}(\mathbf{H}) - 2\mathrm{vec}(\mathbf{X}^{\mathrm{T}}\mathbf{W})^{\mathrm{T}}\mathrm{vec}(\mathbf{H})$$

$$\text{subject to}: \quad -\mathbf{I}_{nk}\mathrm{vec}(\mathbf{H}) \leqslant \mathbf{0}_{nk}$$

$$\text{Given } \mathbf{H}: \min_{\mathbf{W}} \ \mathrm{vec}(\mathbf{W})^{\mathrm{T}}(\mathbf{H}\mathbf{H}^{\mathrm{T}} \otimes \mathbf{I}_m)\mathrm{vec}(\mathbf{W}) - 2\mathrm{vec}(\mathbf{X}\mathbf{H}^{\mathrm{T}})^{\mathrm{T}}\mathrm{vec}(\mathbf{W})$$

$$\text{subject to}: \quad \begin{bmatrix} -\mathbf{I}_{mk} \\ \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathrm{vec}(\mathbf{W}) \leqslant \mathbf{0}_{mk}$$

vec() is a column vectorization of the matrix coefficients and $\otimes$ is the Kronecker tensor product. For optimization of $\mathbf{H}$ (Eq. 4), only the non-negativity constraint is included in the inequality. For optimization of $\mathbf{W}$ (Eq. 5), the new constraints are added in $\mathbf{A} \in \mathbb{R}^{(m-1)k\times mk}$ ensuring a monotonous behavior of the basis functions, and in $\mathbf{B} \in \mathbb{R}^{(m-2)k\times mk}$ ensuring monotonous behavior of their first derivatives. $\mathbf{A}$ and $\mathbf{B}$ can be designed such that the basis functions are either increasing or decreasing (Supplementary Material). The implementation enables mixing of basis functions to involve both decaying and growing signal components if desired. This may become relevant, e.g. for saturation recovery or if a Rician noise bias arises for small signal values.

The programming language MATLAB 2018b (MathWorks, Inc., Natick, Massachusetts, United States) was chosen for implementation, and the trust-region-reflective algorithm was used for the quadratic programming.

To facilitate straightforward adoption into existing processing pipelines the software is made publicly available, https://github.com/sofierahbek/msNMF.

2.3. Practical implementation

Initialization and stopping criterion A basic random initialization of $\mathbf{H}$ with coefficients drawn independently from a uniform distribution between 0 and 1 was used. For small datasets, all voxels were given as input to the iterative ANLS which continued until a stopping criterion was met, being either a maximum of 500 iterations or a relative change in factors below 10^{-8}. To reduce the unavoidable risk of ending in a local minimum due to the non-convex nature of the optimization problem, the process was carried out 10 times (each time with a different initialization), and the lowest cost solution was chosen. This is also known as a "multi-start" strategy [27,19]. For large datasets ($> 10^5$ observations), a stochastic gradient descent framework was utilized due to memory limitations which also served to reduce computation time. Data was divided into randomly selected batches, and the input batch was replaced during optimization for every T'th iteration. This continued until all data had been included (one full epoch) or until fulfilling a convergence criterion based on relative reduction in cost function (for the full data) averaged across the latest five results. Here, the threshold for the relative reduction of the cost function was set to 0.1%, the input batch size was chosen to $\approx 1\%$ of all observations and T was set to 30. We found that the optimization is robust to these choices and that they will primarily affect computation time rather than the obtained result [34].

Rank determination The resulting factorization of data is dependent on the number of components, i.e. the rank, k, which was chosen using prior knowledge or a trial-and-error approach, where the factorization and its residual are used as feedback for rank adjustments [19,27]. If anatomical structures were visible in the residual image, the rank was increased. Conversely, if two components were very similar, the rank was decreased.

Preprocessing steps A few preprocessing steps were implemented to reduce the impact of noise on the factorization. First, the data used as input to the msNMF was required to carry the highest signal in the first measurement (m = 1) to exclude irrelevant signal behaviour from vascular components, for example. Secondly, data was normalized to have initial value 1 before factorization. This removes the influence of the initial signal amplitudes during estimation of the signal components in $\mathbf{W}$. A subsequent projection of the original (unscaled) data onto these ($\mathbf{H} = (\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{X}$) leaves $\mathbf{H}$ with the amount of total signal distributed over the identified components. The introduced data normalization simultaneously amplifies the noise, which was compensated by weighting the ANLS with the initial signal amplitudes. An updated objective function is given in Supplementary Material.

2.4. Validation

Simulated data was used to validate the implementation (visualized in Supplementary Material, Figure S1). To imitate a set of realistic MR measurements, eight non-equidistant samples of three decaying signals, two mono-exponential and one bi-exponential function ($y_1 = e^{-0.01x}, y_2 = 0.6e^{-0.08x} + 0.4e^{-0.03x}, y_3 = e^{-0.15x}$), were generated and mixed across 64x64 voxels. The bi-exponential component was included to demonstrate the method's ability to detect realistic signals beyond the mono-exponential domain. Normally distributed complex noise ($\sigma = 0.06$) was added to the data. This left the final magnitude data with a relatively low SNR and a Rician noise distribution, which is more realistic than the assumption of Gaussian independent and identically distributed (i.i.d) residuals implied by the formulation of the cost function. Knowing the underlying data structure it was possible to quantitatively evaluate the result of the method. The *explained variance* (EV), with 1 indicating a perfect reconstruction, was used to compare the reconstructed curves in $\mathbf{W}$ with the true signals ($[y_1, y_2, y_3]$). The *area under the* ROC *curve* (AUC) was used to compare the spatial distributions in $\mathbf{H}$ with the true mixture maps. This measure explains to what extent the spatial pattern of the components is recognized, and is not affected by the absolute values of the loadings in $\mathbf{H}$. The performance of the msNMF was compared to that of the standard NMF. For a fair comparison, the exact same conditions for both factorizations were used except for the monotonicity constraints, i.e. data normalization and weights were also applied for the NMF which makes it more constrained than standard NMF. Additionally, a stability test was carried out to investigate how much the added constraints stabilize the solution across 50 random initializations. The *coefficient of variation* (CV) was used as

metric with a low percentage indicating high stability. The specified value is the average CV across all elements in **H** (leaving out background voxels).

2.5. Data

The application of the msNMF was demonstrated using two previously published MRI datasets: ex vivo DWI of a monkey brain acquired with 12 b-values at two different temporal scales [35], and in vivo multiple spin-echo imaging with 40 echo times for multi-exponential T_2 (MET_2) analysis of rat spinal cord [11]. The relatively large number of measurements in both datasets yielded well-characterized signal curves, which in both cases were expected to exhibit an approximately multi-exponential signal decay.

Ex vivo DWI data Data recorded and published by Lundell et al. [35] with multidimensional diffusion encoding revealed microstructural details in a post mortem brain of a Vervet monkey. The sample was perfusion fixed and placed in a buffer solution using procedures optimised for post mortem DWI [36], and data was recorded with a 4.7 T pre-clinical MRI system. We refer to the original publication for more details regarding the experimental setup and the data, which is available on reasonable request [35]. In short, the experiment probed the effect of anisotropy as well as time-dependent diffusion, depending on the size of restrictions. The latter was explored by comparing two diffusion acquisitions recorded with the same echo time TE = 68 ms and so-called "tuned" and "detuned" directional encoding gradient waveforms with an average diffusion time of $\sim$ 10 ms ("short diffusion time") and of $\sim$ 20 ms ("long diffusion time") respectively. Both acquisitions included the same 12 b-values in the range 242–4832 s/mm^2 in five slices of central brain. The difference in diffusion time is expected to influence the signal response to restricted diffusion on length scales $\sim$ 3–10 μm, whereas free diffusion will provide similar signal signatures [5]. All brain voxels from the two data sets were used as input for the msNMF and the stochastic gradient descent framework as described in Section 2.3 was used. This resulted in a computation time of 266 s using a workstation equipped with a 28-core system having 16 GB RAM. Signals were decomposed using an initial rank $k = 3$ due to expectations of at least a bi-component signal decay in the brain tissue and an additional component from the surrounding buffer solution. The resulting mixture maps (**H**) for the two data sets were normalized, leaving the coefficients as signal *fractions*, and compared through voxel-wise subtraction. The decomposition was evaluated using knowledge of cell densities in primate brain and findings by Lundell et al. [35].

In vivo MET_2 data Relaxometry data originally presented in [11] by Harkins et al. demonstrated a model for intra-myelinic edema (IME) quantification in rat spinal white matter evaluated by comparisons with histology. In this study, 24 rats were divided into three groups of eight, each receiving a different amount of the IME-inducing toxin hexachlorophene (HCP) in their diet: 0 ppm, 300 ppm and 600 ppm respectively. Forty T_2-weighted spin-echo images were recorded at 9.4 T using a fluid-attenuated inversion-recovery-prepared multiple spin-echo imaging sequence [37] (TI = 2 s, TR = 6 s). The first 32 echoes were collected with an echo spacing of 9 ms starting from TE = 7.4 ms. The last 8 were collected with an echo spacing of 50 ms. Recordings were obtained for a 1.5 mm thick axial slice (0.2x0.2 mm^2 in-plane resolution) of the cervical spinal cord, and histology data were obtained using light microscopy of 1 μm thick slices dissected from the same location for three rats from each group. Both MR images and histology data were analyzed for the water content of edema, myelin and intra/extra axonal space in four different white matter tracts: dorsal cortical spinal tract, funicilus gracilis, rubrospinal tract, and vestibulospinal tract. These regions-of-interest (ROIs) defined the voxels to include in the msNMF, and the input signals were decomposed into three decaying components, one for each of the mentioned compartments, i.e. the rank best suited for comparison was known in advance. **H** was normalized to provide the fraction of each compartment and to facilitate comparison to both histology values and imaging metrics (water fractions) estimated by Harkins et al. The water fractions were defined as in Mackay et al. [14], assuming proportionality between total signal amplitude and total water content. Signal from myelin lipids was considered to be negligible given its rapid T_2 decay (<1 ms) [38–40].

Reproducibility test As a novel method based on a non-convex problem and random initializations, it is sensible to test the reproducibility of the full framework, which was done using the in vivo data [11] and a split-half (NPAIRS) approach [41]: The full data was divided in two sets, each consisting of every second echo, i.e. half of the variables but all observations. The resulting **H** for each dataset was compared using Pearson correlation coefficient to calculate a *reproducibility* measure. The two factorizations were evaluated by projection of **H** onto the other unseen dataset and calculating the EV for each data reconstruction. The average defined a *prediction* measure. A *prediction*/*reproducibility* curve (pr-curve) showing the trade-off between prediction accuracy and pattern reproducibility for a range of ranks k was produced.

3. Results

3.1. Validation data

Fig. 1 shows the final factorization of the simulated data. The true exponential signals and true mixture maps are included for visual comparison, and the quantitative comparison revealed $EV_{\mathbf{W}} = 0.984$ and $AUC_{\mathbf{H}} = 0.999$. The three components are well separated but the spatial distributions seen in **H** are not flawlessly presenting homogeneous circles, especially the purple and yellow component contain traces of other signal components and a higher level of noise (confirmed by the indicated summary statistics). The lower mean value for the purple component illustrates that part of the signal has been incorrectly distributed to the other components, consistent with the "shadow" seen in the yellow mixture map. Please note that even for a hypothetical method perfectly identifying the true signals (dotted lines in Fig. 1a), the residual noise amplitude will be spatially inhomogeneous due to the structure and noise level of the data.

In comparison, the standard NMF extracted erroneous signal components and slightly worse spatial distributions (Supplementary Material, Figure S3). The evaluation metrics were correspondingly lower than those obtained with the msNMF ($EV_{\mathbf{W}} = 0.922$ and $AUC_{\mathbf{H}} = 0.989$). The mean CV for 50 initializations was 5.07 % for the NMF and 1.13 % for the msNMF, i.e. a higher stability of the solution was observed for the msNMF. A table in Supplementary Material shows the msNMF result for simulated data with coil arrays (8, 16 and 32 elements) and other SNR (σ) levels, included to demonstrate the dependence on noise characteristics.

3.2. Ex vivo DWI data

Example raw data are shown in Fig. 2a in the shape of a somewhat T_2-weighted image (TE = 68 ms, b = 242 s/mm^2). Trial-and-error testing of the rank led to a final choice of k = 3. A decrease to k = 2 resulted in a residual map with clear cerebral structures. Conversely, an increment to k = 4 resulted in two of the four components to be very similar for both factors, and **H** was distinctly contaminated with noise compared to the decomposition presented here using k = 3. The resulting three diffusion-related

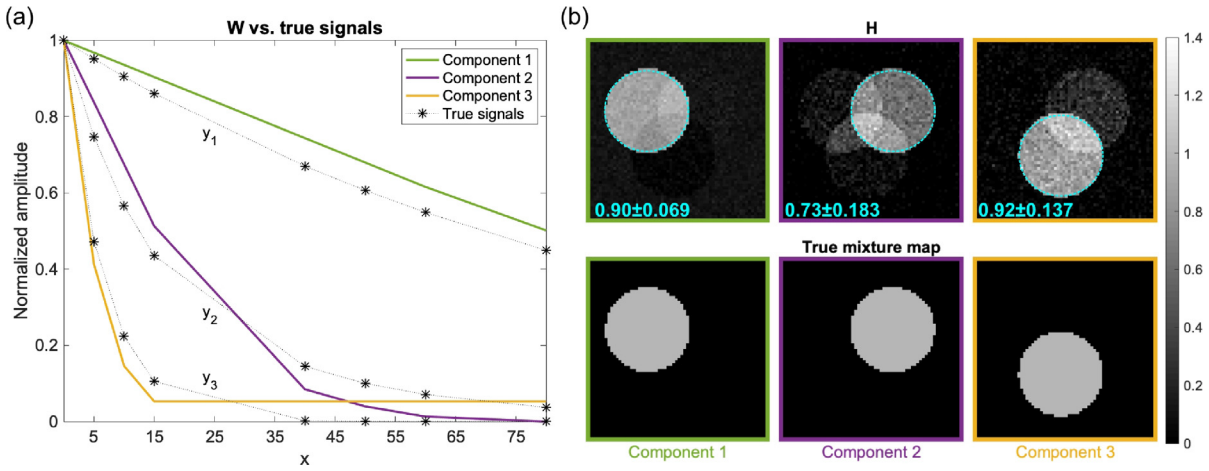

Fig. 1. True data compared to the msNMF result. (a) Estimated signal components, **W**, and true signals ($y_1 = e^{-0.01x}, y_2 = 0.6e^{-0.08x} + 0.4e^{-0.03x}, y_3 = e^{-0.15x}$). (b): Associated mixture maps, **H**, indicated by frame colors. The cyan numbers are mean±SD across the encircled voxel values. Similar illustrations obtained with standard NMF and the mNMF are provided in Supplementary Material.

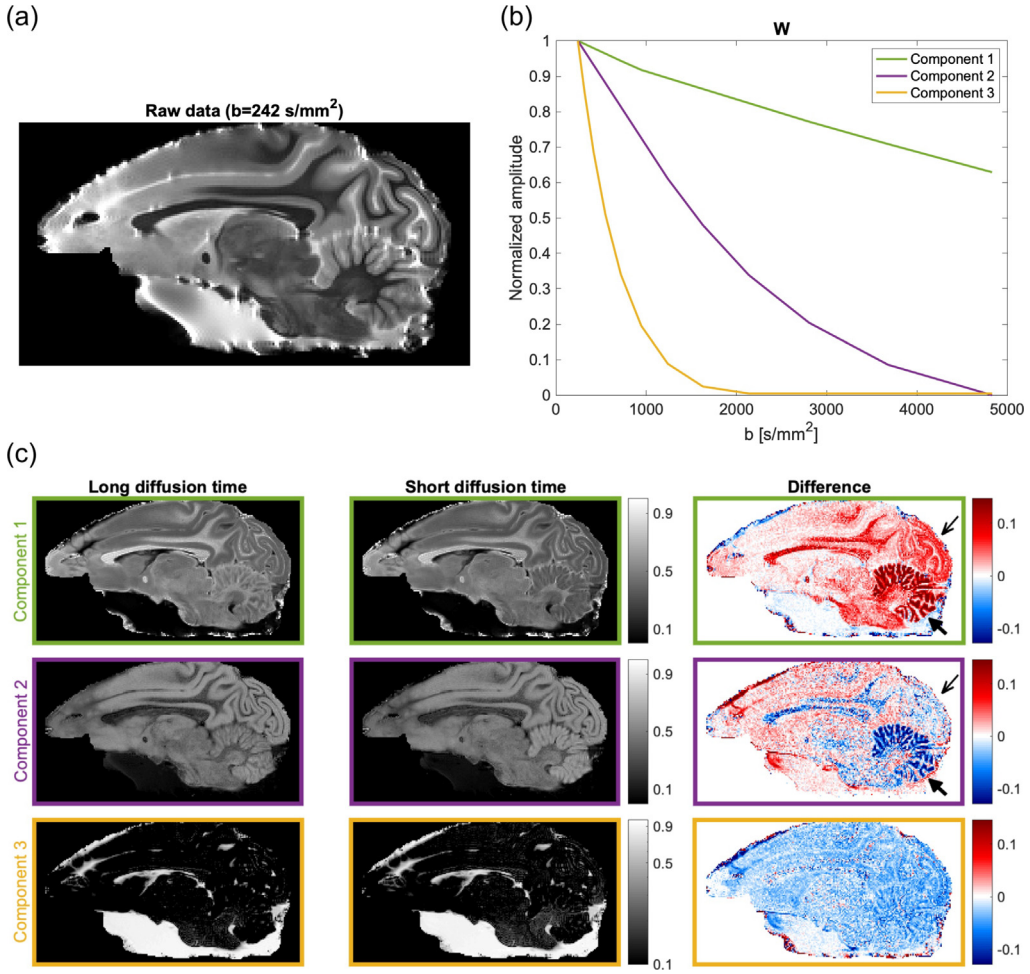

Fig. 2. (a): Measured image of the monkey brain for the lowest acquired b-value (242 s/mm^2) (b): The signal components, **W**, from the msNMF. (c): The associated normalized mixture maps, **H**, labelled by the frame colors. A logarithmic colorscale is used for the yellow component. Maps are given for both data sets and the right column shows the difference between the two ("Short diffusion time" subtracted from "Long diffusion time"). The thin and thick black arrows mark the conspicuous visual cortex and cerebellum, respectively.

components seen in Fig. 2b reveal long, short and intermediate signal lifetime (green, yellow and purple curve, respectively).

The associated normalized mixture maps in Fig. 2c reveal a relatively high fraction of the purple signal component in the cortex, creating a clear separation of grey matter from white matter, which mostly contains the green signal component. The signal from aqueous compartments (buffer solution surrounding the brain sample) is almost exclusively described by the yellow component. The difference between the two sets of maps, provided to capture the effect of time-dependent diffusion, is presented in the third column of images. Positive values (red colors) indicate that an increase in diffusion time results in a higher fraction of the respective component. This effect is observed for the long-lived signal component (green), especially in white matter including corpus callosum, visual cortex (thin arrow), and the cerebellar cortex (thick arrow), which exhibit high contrast. Simultaneously, the opposite effect (negative values) is seen for the intermediate decaying signal component (purple) in these areas. The cerebellum is particularly prominent (thick arrow), and superior to this, the visual cortex (thin arrow) clearly differs from the rest of the cerebral cortex, which otherwise exhibits positive changes for this signal component. The small uniform change across the brain seen for the short-lived signal (yellow) is in agreement with the expectation that a fast decaying liquid component is time independent.

3.3. In vivo MET_2 data

Fig. 3a and 3b show the result of the msNMF: the signal components and mixture maps, respectively. The mixture maps illustrate the distribution of the signals across the spinal cord for a rat from each group. Since only voxels from the white matter tracts have been included in the factorization, the gray matter signals may not be well represented. For all three components, the contrast between the butterfly-shaped gray matter and the surrounding white matter increases with increasing amount of HCP in the diet. Especially the abundance of the long-lived component (green) is increased in specific areas of the white matter for rats on HCP diet. The three signal components are thought to reflect relaxation of water within different compartments in the white matter. Fitting a mono-exponential decay to the curves resulted in relaxation times of 12 ms, 42 ms and 181 ms for the yellow, purple and green curves respectively. With expectations of IME to cause T_2 hyperintensity [11,42], the long-living signal component is consistent with IME. Thus, the fraction of the green component (given by the normalized **H**) is considered an estimate of the EWF.

Fig. 4b presents the mean EWF for the four ROIs visualized in Fig. 4a averaged across the eight rats in each group. The error bars indicate the inter-animal standard deviation. For comparison, Fig. 4c shows the equivalent result presented in [11]. The two bar plots are very similar considering the relative increase of EWF in the different ROIs for the HCP groups. Quantitatively, the paper reports slightly higher values than the method proposed here.

Fig. 5 compares the EWF estimates with the histology-based edema water content fraction (W_{ed}) for three rats from each group. The correlation metric for the linear regression together with the associated p-value reveal a high agreement with histology data (r^2 = 0.88, $P<0.001$). The trendline is closer to unity, the intra-ROI standard errors are lower, and the correlation metric is slightly higher in comparison to the corresponding analysis in the original paper [11]. For both approaches, non-zero intercepts remain in the comparisons with histology estimates.

Fig. 6 shows the result of the split-half test, the pr-curve. The plot reveals a high Pearson correlation (>0.8) for a rank below 5, and a high predictability for a rank above 1. This indicates an acceptable reproducibility of the framework for the chosen rank k = 3 with predictability and correlation both being high.

4. Discussion

4.1. Simulated data

The simulated data was included to validate the method before applying the msNMF to measured data where the true underlying structure is unknown. Three fundamental signals, only eight temporal samples, and a Rician noise distribution were selected to generate a clinically realistic MR dataset, and a non-trivial decomposition problem.

It was confirmed both by the visual comparisons (Fig. 1) and evaluation metrics that the msNMF provided a decomposition in high accordance with the true data, thus validating the implementation. The example, however, demonstrates both the ability and limitation of the method. The imperfect mixture maps observed for the purple and yellow component, in particular, are consistent with a lower integrated signal intensity for these components and a larger similarity between them. Both properties can make them more difficult to distinguish in the factorization. Comparisons with results from the standard NMF and mNMF [30] confirmed that adding constraints tailored to this type of data improves the result and increases the stability of the solution.

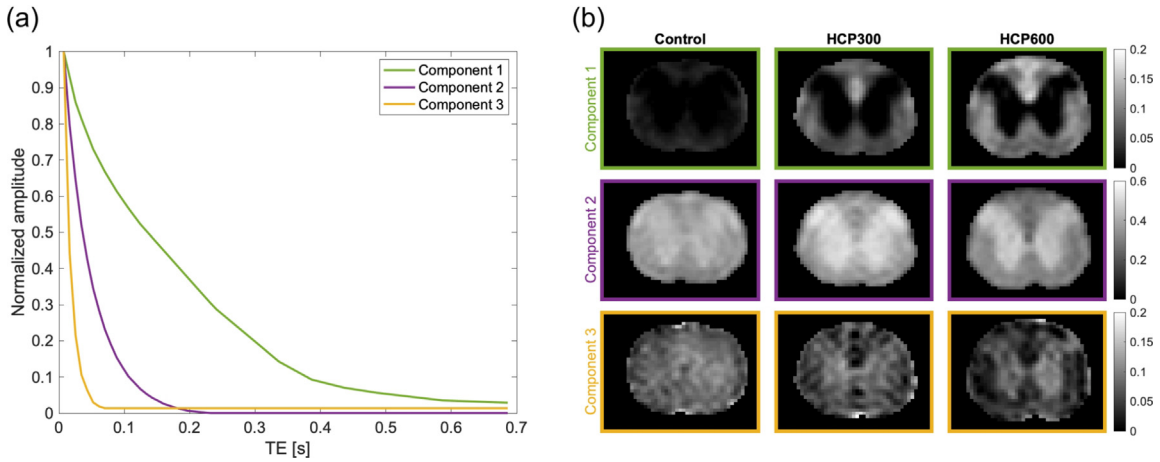

Fig. 3. Result of the msNMF for MET_2 analysis of rat spinal cord. (a): The signal components, **W**. (b): The associated mixture maps, **H**, indicated by the frame and label colors. Each column of images shows the maps for a specific rat (one from each group).

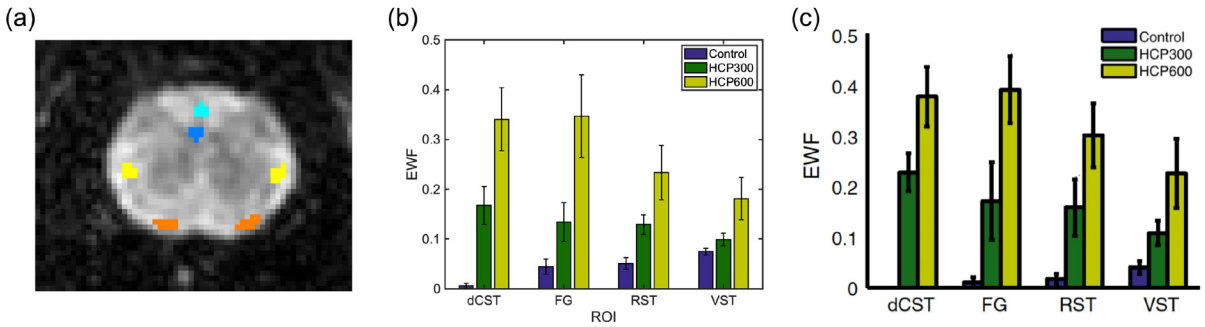

Fig. 4. EWF estimates for each of the spinal cord tracts (ROIs). (a): The ROI locations for an example rat; dorsal cortical spinal tract (blue), funicilus gracilis (cyan), rubrospinal tract (yellow), vestibulospinal tract (orange). (b): Estimates provided by the msNMF: the mean fraction of component 1 (green in Fig. 3) within a ROI averaged across rats. (c): Estimates reported in [11]. Error-bars indicate inter-animal sta.ndard deviation.

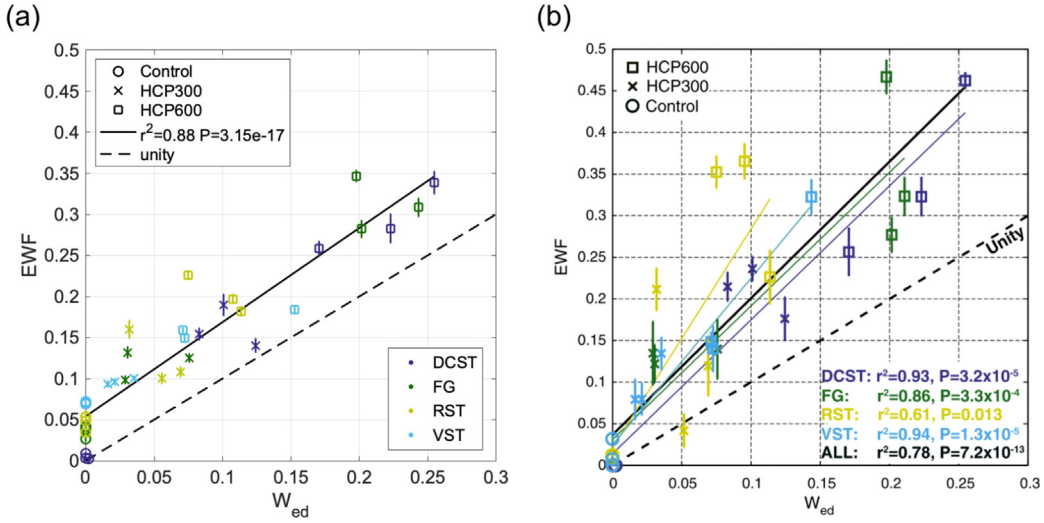

Fig. 5. (a): Scatter plot comparing EWF estimates from the msNMF with W_{ed} from histology data. Colors indicate the ROI, symbols indicate the rat group, and error bars indicate intra-ROI standard error. The black solid line shows a linear regression for all data merged together ($r^2 = 0.88$). (b): The equivalent scatter plot presented in the source paper [11].

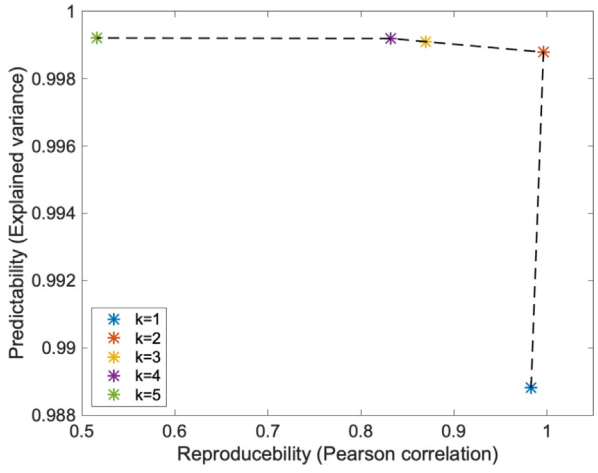

Fig. 6. Prediction/reproducibility curve for the msNMF of the MET_2 data set. The curve is constructed by varying the rank of the decomposition and calculating the Pearson correlation and explained variance using the split-half test.

The superiority to mNMF is particularly important, as this method, among several BSS techniques extended with priors or regularization in order to solve related problems [13,15,43–46], is most similar to our proposed method. We emphasize that the msNMF does not rely on specification of regularization parameters, nor requires information from several contrast dimensions, and we therefore consider it out of scope to compare msNMF to most alternative BSS methods imposing particular signal behaviour. The simple prior assumption of monotonicity of the individual signals and their first-order derivatives have to our knowledge not been utilized in a decomposition method before. These constraints intuitively fit many types of MR data.

A common problem of the Rician or noncentral chi distributed noise present in MRI magnitude data, is a positive bias (noise floor) for low SNR, and the effect is exacerbated using large coil arrays and conventional sum-of-squares coil combination [47,48]. In the validation example, this effect became evident for the yellow signal component (Fig. 1) for the added noise amplitude ($\sigma = 0.06$). A pragmatic strategy to reduce such signal-dependent noise bias is to estimate a single increasing component with monotonous decreasing slope in addition to the decreasing components. This may absorb a considerable fraction of the noise floor. To demonstrate such flexibility of the msNMF, the strategy was tested for coil array data in addition to the single element coil. The addition of a spatially uniform growing component was found to improve the estimates (Supplementary Material, Figure S4 and Table S1). The uniformity condition on the mixture map was used to avoid excessive "cancellation" of the decaying signal components, which may

occur if the noise floor is not reached for all signals and thus is difficult to identify. However, the presence of a positive bias depends on the level of signal in the given voxel, so the assumption of a constant bias across all voxels is not accurate in general. Even so, the proposed strategy can be advantageous to directly account for noise or other sources of bias as demonstrated with the simulated data. Also, it is possible to enforce more involved constraints for the mixture map such as a low frequency basis for each tissue class. For the animal data, the noise bias was not an issue so the described correction strategy was not applied.

The msNMF has proven effective for decomposition of the simulated data generated as a weighted mixture of three main signal components. However, signals of biological origin is often expected to have a continuous distribution of decay rates, e.g. a range of diffusivities from variation in cell sizes, centered around a mean decay rate. In the Supplementary Material (Figure S2), an extended example of the simulated data is presented to show how the method handles such data. For two of the main signal components, the decay rate was drawn from a normal distribution. As the result reveals, the mean signals were recognized, i.e. the decomposition outcome did not change significantly. This shows that msNMF provides an informative and compact representation of the data without necessarily describing all aspects of the underlying data.

As already mentioned, another typical quantitative analysis strategy is the NNLS approach. In spite of the necessity for regularization and averaging across ROIs, the NNLS approach is a relevant reference for the proposed strategy, and results of the method for two sets of simulated data are provided in the Supplementary Material (Figure S5). The presented results rely on post-processing involving k-means partitioning of the raw spectra into three clusters. Final results are comparable to those for the msNMF considering single-coil data, but are more affected by the Rician noise for the coil array data reconstructed using conventional sum-of-squares. Notably, our method decomposes mixed signals in the domain of the acquisition, e.g. the echo time domain, and is not relying on critical or equidistant sampling of the distribution of T_2 values, for example.

An alternative group of strategies worth mentioning is the CORE (component-resolved NMR) processing family, typically used for analysis of diffusion-ordered NMR spectroscopy (DOSY) data sets. Examples are CORE [49], GRECORD [50], OUTSCORE [51], SILT-DOSY [52], and one of the newest presented in the literature, the InSpect algorithm [18]. As for the BSS methods, these techniques have replaced the univariate processing (voxel-wise fitting) with a multivariate decomposition. However, they still require choosing an explicit MR signal model for the reconstruction [18]. As described in the Introduction, complex biological conditions make it difficult to formulate an accurate signal model a priori. Our introduction of msNMF is motivated by this and a direct comparison to methods from the CORE family is therefore not possible.

4.2. Ex vivo DWI data

The DWI dataset was used to investigate whether the extracted components were in fact related to the local diffusion and thus a "fingerprint" of the underlying tissue micro-structure. Being an ex vivo dataset with a broad range of b-values, measurements for two different temporal scales, and an accompanying thorough analysis published in [35], it was well-suited for testing the proposed method.

The comparison of the mixture maps differing by diffusion time (Fig. 2c, right column), uncovered areas of high cell density, especially visual cortex and cerebellum [53,54]. In more detail, the enhancement of the long-lived diffusion-weighted signal (green component) for increased diffusion time is associated with the degree of restrictions in the local tissue environment [35,55]. Therefore, the identified areas can be further categorized into very restricted environments (cerebellum) and intermediate restricted cell environments (visual cortex). This strong contrast across different areas of the cerebral and cerebellar cortex, revealed by the maps, have to our knowledge not been presented based on other MR contrast mechanisms before.

Thus, using two diffusion times and the msNMF for decomposition, the effect of time-dependent diffusion was convincingly demonstrated, exactly as Lundell et al. aimed at when setting up their experiment. Their analysis consisted of a voxel-wise calculation of the powder averaged signal difference at b = 4800 s/mm^2 normalized to the b = 0 signal. The analysis revealed similar time-dependent diffusion contrast, but was dependent on estimating the b = 0 signal from a 3rd order expansion of the signal in b and thus delivered a more noisy map than the one presented in Fig. 2c.

Through the use of non-conventional diffusion encodings with general gradient waveforms, different effects related to the signal attenuation in DWI (e.g. anisotropy, restrictions, exchange) can independently be probed and detected as direct features of the data or contrasts between measurements exploring single encoding dimensions [35,56]. Time-dependent diffusion is a promising marker of cell morphology and density in tumors [57]. This makes the combination of msNMF applied to novel DWI data a potentially powerful tool for characterization of tumor tissue that is generally difficult to model, and thus relevant for e.g. therapy response evaluation. However, to adapt the acquisition for in vivo measurements, fewer b-values (i.e. shorter scan time) and a more simple gradient waveform are desired. For this data, it may thus be interesting to investigate whether the separation into the three diffusion-related components can be obtained with fewer measurements, for example. Such analysis is beyond the scope of this paper but possible now that the components are characterized.

Consideration should also be given to what tissue information can be derived if only one data set was available, e.g if only the left column of mixture maps is used (long diffusion time). The three macrostructural brain tissues are easily distinguished. Gray matter is separated from white matter, and firm tissue from aqueous solutions. If the three decaying signal components each are fitted to a mono-exponential function, the resulting decay constants would be 0.1×10^{-3}, 0.6×10^{-3} and 2.3×10^{-3} mm^2/s respectively for the green, purple and yellow signals. These values are within a range of ADC values observed in the literature for white matter, gray matter and CSF respectively, though the ADC is dependent on acquisition method, handling of ex vivo material, temperature etc. [58–60]. The fitted values are thus only included here to relate the found components to a familiar characterization of diffusion. It is evident that especially the purple signal component deviates from an exponential decay as it reaches 0 for b = 4842 s/mm^2. This is an unrealistic behaviour indicating that all noise/left-over signal is assigned to the green component for the last measurement, possibly an effect of the implicit sparsity of the NMF [19]. The effect can potentially be reduced with signal sampling at even higher b-values to allow all individual signals to reach the noise floor, so they are easier to resolve.

A noticeable feature of the factorization is the alternative representation of the signal variation between the two tissue types, gray and white matter. A difference in the mixture of (mainly) two signal components discriminates the tissue types rather than a difference in a mono-exponential decay constant. This exemplifies that the analysis provides a compact representation of the data that may reveal relevant signal features without necessarily reflecting the underlying time-dependent diffusion phenomena [61], especially when data points are limited. The

decomposition-based representation may robustly reveal useful information about the local diffusion compared to a simplified signal model vulnerable to noise and partial volume effects. Here, it proved possible to capture cell density variation by means of the factorization, thereby demonstrating a connection between components and microstructure.

4.3. In vivo MET_2 data

To further demonstrate the versatility of the method, in vivo data was analyzed quantitatively to extract estimates of edema water content similar to those obtained by Harkins et al. in [11]. Simultaneously, the method was tested as an alternative solution for MET_2 analysis, which in general is challenged due to the ill-posed problem of inverting a sum of exponentials [31], but also highly desired in order to obtain stronger quantitative relationships between MR contrast and tissue microstructure [37,11,10].

The three extracted signal components (Fig. 3) are consistent with T_2-relaxation in the different compartments of white matter, i.e. IME (green component), intra/extra cellular space (purple component), and myelin water (yellow component). The visual contrast change observed in the mixture maps for HCP-affected rats affirmatively indicates that the green component can be interpreted as the "IME signal", and literature confirms a decay constant within the range 5–50 ms for water trapped in myelin sheaths [39].

The strategy used in Harkins et al. is similar to the NNLS approach tested for the simulated data, but assumes a Gaussian prior for the distribution of decay rates, and was found superior to the standard NNLS [62]. The T_2-distribution was fitted with up to three Log-Gaussian shaped peaks (components), and this was followed by a k-means algorithm to cluster the T_2-values into groups of short, intermediate and long decay. The resulting mean decay constants of the distributions were roughly 9 ms, 40 ms, and 150 ms, respectively, thus not far from the values obtained by a mono-exponential fit to the three signal components of the msNMF (12 ms, 44 ms, and 182 ms). The fitted decay rates are presented only to include a well-known characterization of the signal profiles and show that results are consistent with the decomposition reported by the source paper [11] (also confirmed by the bar-plots presented in Fig. 4).

Comparison to histology (Fig. 5a) is important for evaluation of the proposed factorization. The high agreement between EWF estimates and histology, revealed by an r^2 of 0.88 ($P< 0.001$), confirms that the decomposition is related to the underlying compartmentalization of the tissue, and demonstrates the method's ability to detect IME severity from T_2-weighted relaxometry measurements. However, similar to the results reported by Harkins et al., a small overestimation of EWF compared to W_{ed} was observed. In general, one must be critical to the absolute values, from both factorization (or any model) and histology. Explanations for the discrepancies were discussed in [11].

Although the split-half experiment revealed a high reproducibility for a decomposition of rank $k = 3$, and the CV metric for the simulation data revealed a relatively high stability of the method, a stable quantitative measure is in general not guaranteed. A variability of the absolute EWF estimates across multiple runs of the algorithm (as part of the initialization strategy) was indeed observed, however at a lower level than the subject-variability. This means that the relative difference of EWF between rat groups and ROIs was stable, and the correlation of r^2=0.88 with histology was robustly observed (SD = 0.0019 for 50 runs).

The rank of the factorization was predetermined to $k = 3$ to match the decomposition presented by Harkins et al. However, the pr-curve (Fig. 6) supports that two components may be sufficient for a representative reconstruction of the data. Reasons could be that the water exchange effectively merges several physical compartments and diminish the MWF [11], or simply that compartments co-vary so a linear combination of two components is sufficient to describe the relaxation within three different compartments. This also explains why estimation of the MWF may not be trustworthy and thus is not considered here, although it is sometimes a target for relaxometry measurements. Dependence on the choice of dimensionality of the sub-system is a common challenge for decomposition methods as there is no established procedure for rank determination, and metrics from heuristic approaches suggested in the literature are highly dependent on the data properties, making it difficult to know when a metric is suitable [27,63]. The pr-curve in Fig. 6 illustrates the problem, as for a range of rank values ($k = [2, 4]$) an acceptable compromise is seen between reproducibility and degree of data explanation. The suggested "trial-and-error" rank-adjustment guided by left-over anatomy in the residual image relies on a subjective inspection, which is a limitation of the method present when the underlying sub-system dimensionality is unknown. That being said, the rank is the only free parameter of the analysis (with or without the inclusion of a bias-compensating component). In the spectrum of techniques ranging from restrictive modeling to completely data-driven signal processing, the msNMF is in the unrestrictive end while relying on the existence of compartments and using this to enforce realistic signal behaviour, though possibly approximate, e.g. due to pronounced intermediate exchange or parameter variability within physical compartments. The technique used by Harkins et al. is also flexible, but still assumes a distribution of one to three Log-Gaussian shaped peaks of exponential time-constants.

Although assumptions of the signal behaviour are less specific than in model-based analysis, the monotonocity requirement must essentially be fulfilled. For clinical multi-echo spin-echo sequences at fields higher than 1.5 T, for example, transients will typically be present causing initial signal oscillations, which makes the msNMF inappropriate as analysis strategy. Also, RF inhomogeneity correction may be necessary to make signal decay rates independent of position in the coil. No such transients or inhomogeneity were observed in the MET_2 data, recorded with a preclinical scanner delivering a train of near-accurate 180° refocusing flip angles, why it was suited for demonstration of the msNMF.

Overall, the pursuit of micro-structural specificity, pioneered with the myelin water determination for multiple sclerosis patients, have throughout the last decade been in rapid development, and several strategies exist for analysis of relaxometry data [14,10,39]. The msNMF is not promoted as a replacement or advocated as a general new tool for MET_2 analysis. The two clinical data examples included in this paper are used to exemplify the characteristics and potential of the proposed analysis method, but not to promote the method for these applications in particular or in general. Many MR signals fall within the class of *approximately multi-exponential* signals and when prerequisites are met, the method offers a data-driven decomposition which may be useful either for direct analysis or to inspire modeling, e.g. let the detected data structure inform priors for a more specific parameterization. When a monotonous slope signal decay or increase is expected, this relatively simple and rapid analysis can give a compact, informative representation of data.

5. Conclusion

The msNMF robustly separated MR signals into identifiable components specific to the underlying structure, and produced quantitative results consistent with those published in the source papers using tailored analysis. The method is insensitive to partial-volume effects and requires only a choice of rank, and

optionally correction of noise biases. The msNMF is potentially applicable to a broad range of multi-dimensional MR data. The demonstrated sensitivity to cell density, for example, makes it a candidate for tumor tissue characterization needed for prognosis and treatment planning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Mark D. Does and Kevin D. Harkins for their cooperativeness and sharing of MET_2 data, and Nirav Bhatt and Arun Ayyar for providing their software for monotonous NMF. The study was funded by the Danish Cancer Society (grant R167-A10637-17-S2) and Henrik Lundell additionally received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 804746).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jmr.2021.107103.

References

[1] M.D. Does, J.L. Olesen, K.D. Harkins, T. Serradas-Duarte, D.F. Gochberg, S.N. Jespersen, N. Shemesh, Evaluation of principal component analysis image denoising on multi-exponential MRI relaxometry, Magnetic Resonance in Medicine 81 (2019) 3503–3514, https://doi.org/10.1002/mrm.27658.

[2] S.M. Meyers, I.M. Vavasour, B. Mädler, T. Harris, E. Fu, D.K. Li, A.L. Traboulsee, A. L. Mackay, C. Laule, Multicenter measurements of myelin water fraction and geometric mean T 2: Intra- and intersite reproducibility, Journal of Magnetic Resonance Imaging 38 (2013) 1445–1453, https://doi.org/10.1002/jmri.24106.

[3] G. Soujanya Chilla, C. Heng Tan, C. Xu, C. Loo Poh, Diffusion weighted magnetic resonance imaging and its recent trend-a survey, Quant Imaging Med Surg 5 (2015) 407–422, https://doi.org/10.3978/j.issn.2223-4292.2015.03.01, URL www.amepc.org/qims..

[4] F. Fornasa, Diffusion-weighted magnetic resonance imaging: What makes water run fast or slow?, Journal of Clinical Imaging Science 1 (2011) 1–7, https://doiorg/10.4103/2156-7514.81294.

[5] J.S. Nielsen, T.B. Dyrby, H. Lundell, Magnetic resonance temporal diffusion tensor spectroscopy of disordered anisotropic tissue, Scientific Reports 8 (2018), https://doi.org/10.1038/s41598-018-19475-y.

[6] A.R. Padhani, G. Liu, D. Mu-Koh, T.L. Chenevert, H.C. Thoeny, T. Takahara, A. Dzik-Jurasz, B.D. Ross, M. Van Cauteren, D. Collins, D.A. Hammoud, G.J. Rustin, B. Taouli, P.L. Choyke, Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations, Neoplasia 11 (2009) 102–125, https://doi.org/10.1593/neo.81328, http://linkinghub.elsevier.com/retrieve/pii/S1476558609800249.

[7] M. Nagtegaal, P. Koken, T. Amthor, M. Doneva, Fast multi-component analysis using a joint sparsity constraint for MR fingerprinting, Magnetic Resonance in Medicine 83 (2020) 521–534, https://doi.org/10.1002/mrm.27947.

[8] D.K. Jones, M. Cercignani, Twenty-five pitfalls in the analysis of diffusion MRI data, NMR in Biomedicine 23 (2010) 803–820, https://doi.org/10.1002/nbm.1543.

[9] E. Eyal, H. Degani, Model-based and model-free parametric analysis of Breast dynamic-contrast-enhanced MRI, NMR in Biomedicine 22 (2009) 40–53, https://doi.org/10.1002/nbm.1221.

[10] A.L. Mackay, C. Laule, Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin, Brain Plasticity 2 (2016) 71–91, https://doi.org/10.3233/bpl-160033.

[11] K.D. Harkins, W.M. Valentine, D.F. Gochberg, M.D. Does, In-vivo multi-exponential T 2, magnetization transfer and quantitative histology in a rat model of intramyelinic edema, NeuroImage: Clinical 2 (2013) 810–817, https://doi.org/10.1016/j.nicl.2013.06.007.

[12] M. Nagtegaal, P. Koken, T. Amthor, J. de Bresser, B. Mädler, F. Vos, M. Doneva, Myelin water imaging from multi-echo T2 MR relaxometry data using a joint sparsity constraint, NeuroImage 219 (2020), https://doi.org/10.1016/j.neuroimage.2020.117014.

[13] M. Molina-Romero, P.A. Gómez, J.I. Sperl, M. Czisch, P.G. Sämann, D.K. Jones, M. I. Menzel, B.H. Menze, A diffusion model-free framework with echo time dependence for free-water elimination and brain tissue microstructure characterization, Magnetic Resonance in Medicine 80 (2018) 2155–2172, https://doi.org/10.1002/mrm.27181.

[14] A. Mackay, K. Whittall, J. Adler, D. Li, D. Paty, D. Graeb, In vivo visualization of myelin water in brain by magnetic resonance, Magnetic Resonance in Medicine 31 (1994) 673–677, https://doi.org/10.1002/mrm.1910310614.

[15] D. Kim, J.P. Haldar, Nonnegative matrix factorization for tissue mixture modeling with noisy MR magnitude image sequences, Proceedings - International Symposium on Biomedical Imaging 2015-July (2015) 1028–1031, https://doi.org/10.1109/ISBI.2015.7164046.

[16] T.A. Bjarnason, C.R. McCreary, J.F. Dunn, J.R. Mitchell, Quantitative T2 analysis: The effects of noise, regularization, and multivoxel approaches, Magnetic Resonance in Medicine 63 (2010) 212–217, https://doi.org/10.1002/mrm.22173.

[17] V. Wiggermann, I.M. Vavasour, S.H. Kolind, A.L. MacKay, G. Helms, A. Rauscher, Non-negative least squares computation for in vivo myelin mapping using simulated multi-echo spin-echo T2 decay data, NMR in Biomedicine XXX (2020) 1–14, https://doi.org/10.1002/nbm.4277.

[18] P.J. Slator, J. Hutter, R.V. Marinescu, M. Palombo, L.H. Jackson, A. Ho, L.C. Chappell, M. Rutherford, J.V. Hajnal, D.C. Alexander, Data-Driven multi-Contrast spectral microstructure imaging with InSpect: INtegrated SPECTral component estimation and mapping, Medical Image Analysis 71 (2021) 102045, https://doi.org/10.1016/j.media.2021.102045.

[19] N. Gillis, The Why and How of Nonnegative Matrix Factorization, arXiv 7 Mar (2014) 1–25. http://arxiv.org/abs/1401.5226. arXiv:1401.5226.

[20] H. Benali, I. Buvat, F. Frouin, J.P. Bazin, J. Chabriais, R. Di Paola, Factor Analysis of Medical Image Sequences (FAMIS): Fundamental principles and applications, in: E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, B. Burtschy (Eds.), New Approaches in Classification and Data Analysis, Springer Berlin Heidelberg, Berlin, Heidelberg, 1994, pp. 619–627.

[21] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006, URL http://www.library.wisc.edu/selectedtocs/bg0137.pdf. doi:10.1641/B580519. arXiv:0-387-31073-8..

[22] D.J. Bartholomew, International Encyclopedia of Education, 3rd Edition, Chapter 18: Principal Components Analysis, 3rd ed., Elsevier Ltd, 2010. doi:10.1016/B978-0-08-044894-7.01358-0. arXiv:arXiv:1011.1669v3.

[23] M. Mudrova, A. Prochazka, Principal component analysis in image processing, in: Proceedings of the MATLAB Technical Computing Conference, 3, 2005, p. 4. http://dsp.vscht.cz/konference_matlab/matlab05/prispevky/mudrova/mudrova.pdf.

[24] H. Huang, J. Lu, J. Wu, Z. Ding, S. Chen, L. Duan, J. Cui, F. Chen, D. Kang, L. Qi, W. Qiu, S.W. Lee, S.J. Qiu, D. Shen, Y.F. Zang, H. Zhang, Tumor tissue detection using blood-oxygen-level-dependent functional MRI based on independent component analysis, Scientific Reports 8 (2018) 1–16, https://doi.org/10.1038/s41598-017-18453-0.

[25] P.M. Kim, B. Tidor, Subsystem identification through dimensionality reduction of large-scale gene expression data, Genome Research 13 (2003) 1706–1718, https://doi.org/10.1101/gr.903503.

[26] K. Huang, N.D. Sidiropoulos, A. Swami, Non-Negative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition, IEEE Transactions on Signal Processing 62 (2014) 211–224, https://doi.org/10.1109/TSP.2013.2285514.

[27] Y.X. Wang, Y.J. Zhang, Nonnegative matrix factorization: A comprehensive review, IEEE Transactions on Knowledge and Data Engineering 25 (2013) 1336–1353, https://doi.org/10.1109/TKDE.2012.51.

[28] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (1999) 788–791.

[29] H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, S.H. Jensen, Theorems on positive data: On the uniqueness of NMF, Computational Intelligence and Neuroscience 2008 (2008), https://doi.org/10.1155/2008/764206.

[30] N. Bhatt, A. Ayyar, Monotonous (Semi-) Nonnegative Matrix Factorization, in, in: Proceedings of the Second ACM IKDD Conference on Data Sciences, 2015, pp. 92–97, https://doi.org/10.1145/2732587.2732600.

[31] M.D. Does, Inferring brain tissue composition and microstructure via MR relaxometry, NeuroImage 182 (2018) 136–148, https://doi.org/10.1016/j.neuroimage.2017.12.087.

[32] J. Pfeuffer, U. Flo, W. Dreher, D. Leibfritz, Restricted diffusion and exchange of intracellular water: theoretical modelling and diffusion time dependence of 1H NMR measurements on perfused glial cells, NMR in Biomedicine 11 (1998) 19–31.

[33] S.C. Grant, D.L. Buckley, S. Gibbs, A.G. Webb, S.J. Blackband, MR microscopy of multicomponent diffusion in single neurons, Magnetic Resonance in Medicine 46 (2001) 1107–1112, https://doi.org/10.1002/mrm.1306.

[34] Y. Bengio, Practical recommendations for gradient-based training of deep architectures, in: Practical Recommendations for Gradient-Based Training of Deep Architectures, in: G. Montavon, G.B. Orr, K.R. Müller (Eds.), Neural Networks: Tricks of the Trade, Lecture Notes in Computer Science, vol 7700. Springer, volume 7700, Springer, Berlin Heidelberg, 2012, pp. 437–478, https://doi.org/10.1007/978-3-642-35289-8-26, arXiv:1206.5533.

[35] H. Lundell, M. Nilsson, T.B. Dyrby, G.J. Parker, P.L. Cristinacce, F.L. Zhou, D. Topgaard, S. Lasič, Multidimensional diffusion MRI with spectrally modulated gradients reveals unprecedented microstructural detail, Scientific Reports 9 (2019) 1–12, https://doi.org/10.1038/s41598-019-45235-7.

[36] T.B. Dyrby, W.F. Baaré, D.C. Alexander, J. Jelsing, E. Garde, L.V. Søgaard, An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets, Human Brain Mapping 32 (2011) 544–563, https://doi.org/10.1002/hbm.21043.

[37] K.D. Harkins, A.N. Dula, M.D. Does, Effect of intercompartmental water exchange on the apparent myelin water fraction in multiexponential T 2 measurements of rat spinal cord, Magnetic Resonance in Medicine 67 (2012) 793–800, https://doi.org/10.1002/mrm.23053.
[38] K.P. Whittall, A.L. MacKay, D.A. Graeb, R.A. Nugent, D.K. Li, D.W. Paty, In vivo measurement of T2 distributions and water contents in normal human brain, Magnetic Resonance in Medicine 37 (1997) 34–43, https://doi.org/10.1002/mrm.1910370107.
[39] P.S. Tofts, Quantitative MRI of the Brain: Measuring Changes Caused by Disease, John Wiley and Sons Inc., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2003.
[40] H.W. Fischer, P.A. Rinck, Y. van Haverbeke, R.N. Muller, Nuclear relaxation of human brain gray and white matter: Analysis of field dependence and implications for MRI, Magnetic Resonance in Medicine 16 (1990) 317–334, https://doi.org/10.1002/mrm.1910160212.
[41] P.M. Rasmussen, L.K. Hansen, K.H. Madsen, N.W. Churchill, S.C. Strother, Model sparsity and brain pattern interpretation of classification models in neuroimaging, Pattern Recognition 45 (2012) 2085–2100, https://doi.org/10.1016/j.patcog.2011.09.011.
[42] J.A. Cohen, R.S. Fisher, M.G. Brigell, R.G. Peyster, S. Gordon, The potential for Vigabatrin-induced intramyelinic edema in humans, Epilepsia 41 (2000) 148–157, https://doi.org/10.1111/j.1528-1157.2000.tb00134.x.
[43] J. Zhong, N. Didonato, P.G. Hatcher, Independent component analysis applied to diffusion-ordered spectroscopy: Separating nuclear magnetic resonance spectra of analytes in mixtures, Journal of Chemometrics 26 (2012) 150–157, https://doi.org/10.1002/cem.2423.
[44] B. Jeurissen, J.-D. Tournier, J. Sijbers, Tissue-type segmentation using non-negative matrix factorization of multi-shell diffusion-weighted MRI images, Proc. Intl. Soc. Mag. Reson. Med. 23 (2015) 0349.
[45] D. Christiaens, S. Sunaert, P. Suetens, F. Maes, Convexity-constrained and nonnegativity-constrained spherical factorization in diffusion-weighted imaging, NeuroImage 146 (2017) 507–517, https://doi.org/10.1016/j.neuroimage.2016.10.040.
[46] I. Toumi, B. Torrésani, S. Caldarelli, Effective processing of pulse field gradient NMR of mixtures by blind source separation, Analytical Chemistry 85 (2013) 11344–11351, https://doi.org/10.1021/ac402085x.
[47] N. Dikaios, S. Punwani, V. Hamy, P. Purpura, S. Rice, M. Forster, R. Mendes, S. Taylor, D. Atkinson, Noise estimation from averaged diffusion weighted images: Can unbiased quantitative decay parameters assist cancer evaluation?, Magnetic Resonance in Medicine 71 (2014) 2105–2117, https://doiorg/10.1002/mrm.24877.
[48] O. Dietrich, J.G. Raya, S.B. Reeder, M. Ingrisch, M.F. Reiser, S.O. Schoenberg, Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics, Magnetic Resonance Imaging 26 (2008) 754–762, https://doi.org/10.1016/j.mri.2008.02.001.
[49] P. Stilbs, K. Paulsen, Global least-squares analysis of large, correlated spectral data sets and application to chemical kinetics and time-resolved fluorescence, Review of Scientific Instruments 67 (1996) 4380–4386, https://doi.org/10.1063/1.1147539.
[50] P. Stilbs, Automated CORE, RECORD, and GRECORD processing of multi-component PGSE NMR diffusometry data, European Biophysics Journal 42 (2013) 25–32, https://doi.org/10.1007/s00249-012-0794-8.
[51] A.A. Colbourne, S. Meier, G.A. Morris, M. Nilsson, Unmixing the NMR spectra of similar species - Vive la différence, Chemical Communications 49 (2013) 10510–10512, https://doi.org/10.1039/c3cc46228e.
[52] B. Yuan, Y. Ding, G.M. Kamal, L. Shao, Z. Zhou, B. Jiang, P. Sun, X. Zhang, M. Liu, Reconstructing diffusion ordered NMR spectroscopy by simultaneous inversion of Laplace transform, Journal of Magnetic Resonance 278 (2017) 1–7, https://doi.org/10.1016/j.jmr.2017.03.004.
[53] C.E. Collins, D.C. Airey, N.A. Young, D.B. Leitch, J.H. Kaas, Neuron densities vary across and within cortical areas in primates, Proceedings of the National Academy of Sciences of the United States of America 107 (2010) 15927–15932. doi:10.1073/pnas.1010356107.
[54] J. Wen, M.S. Goyal, S.V. Astafiev, M.E. Raichle, D.A. Yablonskiy, Genetically defined cellular correlates of the baseline brain MRI signal, Proceedings of the National Academy of Sciences of the United States of America 115 (2018) E9727–E9736. doi:10.1073/pnas.1808121115.
[55] C. Najac, F. Branzoli, I. Ronen, J. Valette, Brain intracellular metabolites are freely diffusing along cell fibers in grey and white matter, as measured by diffusion-weighted MR spectroscopy in the human brain at 7 T, Brain Structure and Function 221 (2016) 1245–1254, https://doi.org/10.1007/s00429-014-0968-5.
[56] D. Topgaard, Multidimensional diffusion MRI, Journal of Magnetic Resonance 275 (2017) 98–113, https://doi.org/10.1016/j.jmr.2016.12.007.
[57] O. Reynaud, Time-dependent diffusion MRI in cancer: Tissue modeling and applications, Frontiers in Physics 5 (2017) 1–16, https://doi.org/10.3389/fphy.2017.00058.
[58] D. Xing, N.G. Papadakis, C.L.H. Huang, V.M. Lee, T.A. Carpenter, L.D. Hall, Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain, Magnetic Resonance Imaging 15 (1997) 771–784, https://doi.org/10.1016/S0730-725X(97)00037-4.
[59] S. Rane, T.Q. Duong, Comparison of In Vivo and Ex Vivo Diffusion Tensor Imaging in Rhesus Macaques at Short and Long Diffusion Times, The Open Neuroimag Journal 5 (2011) 172–178, https://doi.org/10.2174/1874440001105010172.
[60] A. Roebroeck, K.L. Miller, M. Aggarwal, Ex vivo diffusion MRI of the human brain: Technical challenges and recent advances, NMR in Biomedicine 32 (2019) 1–14, https://doi.org/10.1002/nbm.3941.
[61] D.S. Novikov, V.G. Kiselev, S.N. Jespersen, On modeling, Magnetic Resonance in Medicine 79 (2018) 3172–3193, https://doi.org/10.1002/mrm.27101.
[62] G.J. Stanisz, R.M. Henkelman, Diffusional anisotropy of T2 components in bovine optic nerve, Magnetic Resonance in Medicine 40 (1998) 405–410, https://doi.org/10.1002/mrm.1910400310.
[63] L. Muzzarelli, S. Weis, S.B. Eickhoff, K.R. Patil, Rank Selection in Non-negative Matrix Factorization: Systematic comparison and a new MAD metric, International Joint Conference on Neural Networks (IJCNN), volume July, IEEE, Budapest, Hungary, 2019, pp. 1–8, https://doi.org/10.1109/IJCNN.2019.8852146.