





RESEARCH ARTICLE

A comparison of non-negative matrix underapproximation methods for the decomposition of magnetic resonance spectroscopy data from human brain tumors

Gulnur Ugan^{1,2}  | Carles Arús^{1,2}  | Alfredo Vellido^{1,3}  |
Margarida Julià-Sapé^{1,2} 

¹Centro de Investigación Biomédica en Red (CIBER), Madrid, Spain

²Departament de Bioquímica i Biologia Molecular and Institut de Biotecnologia i Biomedicina (IBB), Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

³IDEAI-UPC Intelligent Data Science and Artificial Intelligence Research Center, Universitat Politècnica de Catalunya (UPC) BarcelonaTech, Barcelona, Spain

Correspondence

Margarida Julià-Sapé, Departament de Bioquímica i Biologia Molecular and Institut de Biotecnologia i Biomedicina (IBB), Universitat Autònoma de Barcelona (UAB), Barcelona, Spain.

Email: margarita.julia@uab.cat

Funding information

H2020-EU.1.3. - EXCELLENT SCIENCE - Marie Skłodowska-Curie Actions, grant number H2020-MSCA-ITN-2018-813120. *Proyectos de investigación en salud 2020*, grant numbers PI20/00064 and PI20/00360. Spanish *Ministerio de Economía y Competitividad* SAF2014-52332-R. AEI Research Project PID2019-104551RB-I00. *Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina* (CIBER-BBN [<http://www.ciber-bbn.es/en>, accessed on September 12, 2022], CB06/01/0010), an initiative of the *Instituto de Salud Carlos III* (Spain) co-funded by EU *Fondo Europeo de Desarrollo Regional* (FEDER). Spanish AEI PID2019-104551RB-I00 grant. *XartecSalut*, 2018 XARDI 00016 and 2021 XARDI 00021.

Abstract

Magnetic resonance spectroscopy (MRS) is an MR technique that provides information about the biochemistry of tissues in a noninvasive way. MRS has been widely used for the study of brain tumors, both preoperatively and during follow-up. In this study, we investigated the performance of a range of variants of unsupervised matrix factorization methods of the non-negative matrix underapproximation (NMU) family, namely, sparse NMU, global NMU, and recursive NMU, and compared them with convex non-negative matrix factorization (C-NMF), which has previously shown a good performance on brain tumor diagnostic support problems using MRS data. The purpose of the investigation was 2-fold: first, to ascertain the differences among the sources extracted by these methods; and second, to compare the influence of each method in the diagnostic accuracy of the classification of brain tumors, using them as feature extractors. We discovered that, first, NMU variants found meaningful sources in terms of biological interpretability, but representing parts of the spectrum, in contrast to C-NMF; and second, that NMU methods achieved better classification accuracy than C-NMF for the classification tasks when one class was not meningioma.

KEYWORDS

brain tumor, glioblastoma, low grade glioma, magnetic resonance spectroscopy, meningioma, metastasis, non-negative matrix factorization, non-negative matrix underapproximation

Abbreviations: A2, astrocytoma grade II; ACC, accuracy; AGG, aggressive; a.u., arbitrary units; AUC, area under the ROC curve; BER, balanced error rate; C-NMF, convex non-negative matrix factorization; D-C-NMF, discriminant convex non-negative matrix factorization; DR, dimensionality reduction; GL, glioblastoma; G-NMU, global NMU; LDA, linear discriminant analysis; LGG, low grade glioma; L-NMU, Lagrangian NMU; ME, metastasis; MRS, magnetic resonance spectroscopy; NMF, non-negative matrix factorization; NMU, non-negative matrix underapproximation; NO, normal; R-NMU, recursive NMU; S-NMU, sparse NMU; WHO, World Health Organization.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *NMR in Biomedicine* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Over the past decades, several noninvasive methods have been introduced for the preoperative characterization of human brain tumors. In clinical practice, radiologists provide information about the presence, anatomical localization, boundaries, and the possible type and grade of an abnormal brain mass using magnetic resonance (MR) imaging modalities. One of the MR modalities available to radiologists is MR spectroscopy (MRS), which provides information about the metabolites in solution in the volume of tissue under study, in the millimolar range of concentration. MRS is regarded as a promising technique but has not yet reached its full clinical potential. Reasons for this lack of generalized use in clinical practice include the difficulty in its processing and quantification and the variability in the metabolic patterns found in patients, which require assembling large datasets. In the past, several studies have demonstrated the added value of MRS in the preoperative clinical evaluation of brain tumors.¹⁻³ However, there is still room for improvement, specifically to overcome the bottleneck of brain tumor MRS data sparsity: there is an important degree of patient-to-patient heterogeneity in the MRS-recorded metabolic patterns of some tumor types and also partially overlapping patterns among tumor classes. For example, glioblastoma (GL) is a tumor of glial origin, is the most malignant primary brain tumor (grade IV), and its MRS normally shows a necrotic-like pattern dominated by lipids. However, on the one hand, GL MRS patterns have traditionally been difficult to distinguish from those of brain metastasis (ME), as they present with similar MRS patterns⁴; on the other hand, GL data acquired with the single voxel (SV) variant of MRS show high variability, with different degrees of necrosis.⁵ When the necrotic pattern is not dominating the spectrum, other peaks can be observed, such as those corresponding to metabolites that are typical of proliferation (choline-containing compounds, among others), which make the pattern more similar to lower grades of glial tumors.¹ Therefore, it would be desirable to find ways to deal with this sparsity and high variability to improve on the preoperative prediction of type and grade of brain tumors using MRS. The most obvious way would be to assemble larger datasets to improve on the MRS-based classification, but there are two problems regarding this: one is that brain tumors can be considered rare diseases (1.3% of all cancers⁶), and the second is that, too often, patient MRS data are not shared for a variety of reasons, ranging from local practice and regulations to difficulties in handling data formats and establishing consensus acquisition conditions.^{7,8}

Unsupervised feature extraction methods have been used to help overcome the sparsity problem in MRS data; independent component analysis (ICA),⁹⁻¹⁵ non-negative matrix factorization (NMF),^{16,17} convex non-negative matrix factorization (C-NMF),¹⁸ and discriminant NMF (D-C-NMF),¹⁹ among others, have been applied to MRS data of brain tumors.

ICA has been used with MRS datasets in several studies¹⁰⁻¹⁴ for different purposes, including the extraction of individual features that could match individual metabolites¹²; as a feature extraction methodology to further classify the type and grade of brain tumors^{11,12}; or to distinguish white matter from gray matter.¹⁴ The robustness of ICA features has also been studied in brain tumors for finding interpretable components of MRS.¹⁰

The use of C-NMF and D-C-NMF showed that whole interpretable spectral patterns could be identified that were highly correlated with class label.¹⁹⁻²² Furthermore, a comparative study²¹ showed that C-NMF with k-means initialization performed better for decomposition of human brain MRS than principal components analysis (PCA) and ICA of human brain MRS.

The objective of the current study is to test novel, unsupervised approaches designed to deal with data sparsity in the context of MRS data feature extraction. For that purpose, we consider the available family of non-negative matrix underapproximation (NMU)²³ methods, whose application has already shown^{24,25} that sparsity constraints allow the algorithm to detect local important patterns. We chose three different NMU methods to compare with C-NMF, namely, sparse non-negative matrix underapproximation (S-NMU), global non-negative matrix underapproximation (G-NMU), and recursive non-negative matrix underapproximation (R-NMU), as source extraction techniques to apply to our data, to evaluate how locally detected patterns on MRS could be used in brain tumor classification. That is, we use these NMU variants as dimensionality reduction (DR) methods prior to classification.

The rest of this paper is organized as follows: Section 2 describes the MRS dataset under analysis and the methodology applied, Section 3 compiles the results of the experiments for different tasks, and Section 4 discusses them.

2 | MATERIALS AND METHODS

2.1 | Dataset

The MRS data analyzed in this work are SV. That is, for each patient there is a single spectrum corresponding to a 4–8 cm³ cuboidal volume located within the tumor core. The multicenter data analyzed in this study were taken from the European research project INTERPRET (2000–2002).^{1,2,26} MRS data were acquired according to medical ethics regulations, with the Helsinki Declaration and the Spanish “Ley Orgánica de Protección de Datos de Carácter Personal (LOPD), Ley Orgánica 15/1999” and the “95/46/EU directive on data protection, December 13th, 1999”. All patients or their legal representatives signed informed consent forms, agreeing to the study and to the use of their deidentified (anonymized) data for research. Class labeling was performed according to the 2000 World Health Organization (WHO) classification of brain tumors after histopathological analysis of a biopsy sample taken from the solid tumor region.^{1,2,26} Data were acquired at 1.5 T and at short echo time (20–32 ms) (STE) from brain tumor patients and healthy controls. The analyzed dataset includes 22 astrocytoma grade II (A2; low grade glial

tumors, which are infiltrative and have the potential to evolve to GL), 86 GL (high grade malignant tumors of glial origin), 38 ME (high grade malignant tumors whose origin is outside of the brain), 58 low-grade meningiomas (MM; low grade, from meningeal origin), 124 aggressive (AGG; ME + GL), and 22 normal brain parenchyma spectra from normal healthy controls (NO). Although the data size may seem small, the dataset is the largest available multicenter MRS database.^{1,2,26} The MRS signal is described in the frequency domain in parts per million units (ppm), where specific frequency ranges correspond to different MRS-detectable metabolites present in the tissue in the millimolar range of concentration. Details of the ppm range and corresponding metabolites are described in several studies.^{1,2,26} Data were processed as in Tate et al.¹ Briefly, 195 frequency intensity values were sampled between the 0.5 and 4.2 ppm interval, in which the most relevant metabolite signals can be found. Figure 1 shows a summary of the spectra used in the analysis. Note that for some of the comparisons we made, the GL and ME are also merged in a superclass (AGG) because of their high grade of malignancy and previous literature also merging them as superclass for machine learning studies.^{19,21,26,27}

2.2 | Methods

2.2.1 | Classification tasks

To evaluate the methods, we chose several diagnostic problems frequently investigated in the literature of in vivo MRS of brain tumors (for a review, see Julià-Sapé et al.²). Some of the problems achieve high accuracy, while others are difficult, and involve dealing with heterogeneous and/or overlapping patterns. In this sense, and to compare the methods with C-NMF, the choice of specific problems was the same as in Ortega-Martorell et al.,²¹ where C-NMF was in turn compared with PCA. The six chosen problems were as follows:

1. A2 versus NO: Can normal brain be distinguished from infiltrative tumors?
2. A2 versus ME versus NO: Are grades II and IV well differentiated and distinct from normal brain tissue?
3. A2 versus GL versus NO: Are grades II and IV still well differentiated and distinct from normal brain tissue when one of the classes (GL) is (spectroscopically) heterogeneous?
4. A2 versus MM versus NO: Can low grades of distinct nature (A2, infiltrative/aggressive, MM, noninfiltrative/benign) be differentiated?
5. A2 versus AGG (GL + ME) versus NO: Can we distinguish between grades including one heterogeneous class against normal brain?
6. A2 versus AGG (GL + ME) versus MM: Can we distinguish between grades including one high-grade heterogeneous class and two different low-grade types?

The fifth and sixth problems can be considered as classical discrimination problems using in vivo MRS of brain tumors in patients.²⁷⁻³⁰

2.2.2 | Methods overview

NMF

The main idea for NMF¹⁷ is to provide a mathematical model that shows the perception of objects (originally images) as weighted combinations of their constituent parts. In NMF, the data matrix or observed data $M_{d \times n}$, where d is the data dimensionality and n is the number of observations, is approximately factorized into two non-negative matrices: the source matrix $W_{d \times k}$, and the mixing matrix $V_{k \times n}$, where k is the number of sources and k is less than n . The product of these two matrices is a good approximation of the original data matrix:

$$M \approx WV, \tag{1}$$

where WV is a low-rank approximation of M . Then the NMF problem can be written as follows:

$$\min_{W \in \mathbb{R}^{d \times k}, V \in \mathbb{R}^{k \times n}} \|M - WV\|_F^2 \text{ such that } V \geq 0 \text{ and } W \geq 0. \tag{2}$$

C-NMF

C-NMF¹⁸ was defined to relax the non-negativity constraint of NMF and allow both W and V to have negative entries. This is important in the application area of in vivo MRS, because spectra often contain negative values (in particular, for MRS, at long echo times), so that their true sources would be expected to contain negative values as well. In C-NMF, an additional constraint requires that $W = MA$, where $A_{n \times k}$ is an auxiliary adaptive weight matrix that fully determines W . Then Equation (1) can be written as

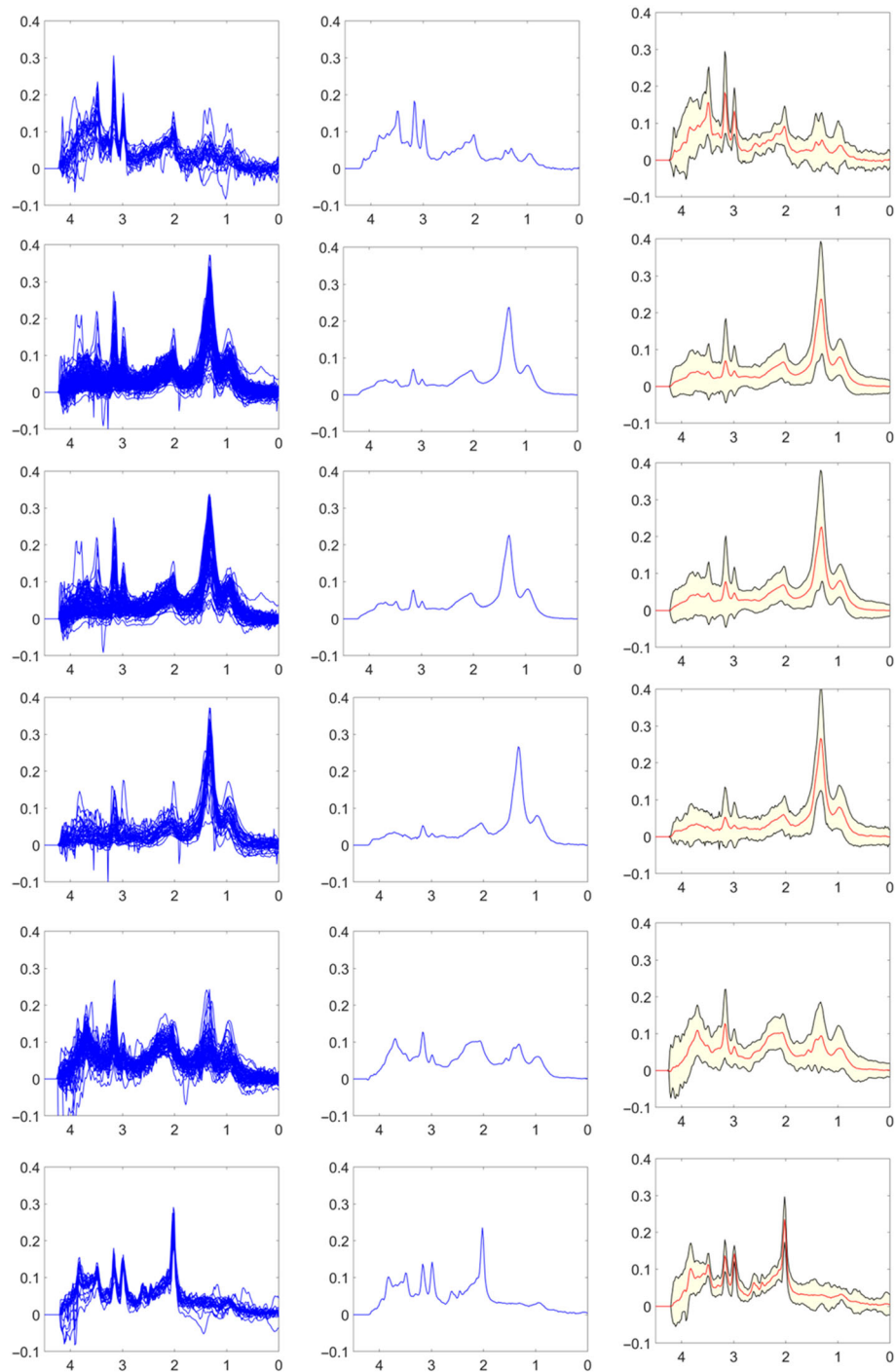


FIGURE 1 Overlaid spectra, mean spectra, and standard deviation (std) of each class. First row to last row: astrocytoma grade II (A2), aggressive (AGG), glioblastoma (GL), metastasis (ME), meningioma (MM), and normal (NO). Left column to right column: overlaid spectra, mean spectrum, and mean \pm std of overlaid spectra. See Data S3 (in the supporting information) for an enlarged version of the figure.

$$M \approx MAV. \quad (3)$$

NMU

This method is a rank-one NMF with $r = 1$, where r is the number of extracted sources, which is easier than higher factorization ranks. When rank-one NMF is applied, the factor matrices are vectors v and w . However, the general problem of higher rank NMF is nondeterministic polynomial time-hard (NP-hard), but, if the rank-one factor solution is applied iteratively (finding subsequent optimal rank-one NMF solutions), then

subtracting wv^T (the product of v and w) from M and applying rank-one factor NMF to this new matrix, if the rank-one approximation does not contain negative values, the strategy becomes computationally feasible.

In order to overcome the negativity, if a constraint is added ensuring that $wv^T \leq M$, an upper-bound constraint is introduced to the general NMF problem, and this new problem is known as NMU.²⁴ It is written as

$$\min_{\substack{w \in \mathbb{R}^{m \times r} \\ V \in \mathbb{R}^{r \times n}}} \|M - WV\|_F^2 \text{ such that } V \geq 0 \text{ and } W \geq 0 \text{ and } WV \leq M, \tag{4}$$

where V and W are built as the sum of all rank-solutions $w_1v_1^T + w_2v_2^T + w_3v_3^T + \dots + w_nv_n^T$.

Importantly, this recursive procedure enables a more localized part-based composition, in the sense that different basis elements tend to describe disjoint parts of the input data.

L-NMU

In this variant of NMU, the Lagrangian relaxation-based nonlinear optimization scheme is combined with Equation (2).

G-NMU

This method merges L-NMU with the rank- r problem.

R-NMU

In R-NMU,³¹ the L-NMU is run successively r times to compute r rank-one approximations, subtracting each approximation from the input matrix M before computing the next one.

S-NMU

This method has an additional sparsity constraint because NMU is typically not able to detect all small objects. In image analysis, it refers to object detection or segmentation. In MRS, it would refer to patterns or individual peaks, although this is a localized part-based composition technique. The sparsity constraint forces NMU to extract more localized features.

Among the NMU methods, only G-NMU requires initialization. To obtain comparable results, *k-Means* initialization was used for C-NMF and G-NMU. The number of iterations was set to 5000 and, to achieve robust results, each experiment was repeated 50 times. The NMU methods were implemented on Matlab and the code is available from the URL <https://sites.google.com/site/nicolagillis/code>. The computational time of each method is reported in Data S4. The fastest computational time was for two-source extraction with R-NMU for Problem 1 with 0.89 s, and the highest computational time was recorded for Problem 4 with eight S-NMU source extraction as 9.81 s. The average computational times for all of the problems were calculated as 4.95, 3.11, 3.88, and 4.78 s for C-NMF, G-NMU, R-NMU, and S-NMU, respectively.

2.2.3 | Number of extracted sources

For each classification task, the number of extracted sources starts from the number of classes and goes up to eight sources. That is two sources for binary classification tasks (e.g., A2 vs. NO), aiming to evaluate whether a two-source extraction yields interpretable results. The rationale for eight sources is that the maximum number of classes is four (A2 vs. GL + ME vs. MM or A2 vs. GL + ME vs. NO) and we aimed to explore whether the methods could discern the heterogeneity inside of a class (two sources per class). In particular, the GL class has long been reported to display heterogeneous spectral patterns in SV MRS and in this dataset, too.^{1,5} For each classification task, the number of extracted sources started from the number of classes and went up to eight sources; that is, from two sources for A2 versus NO and from three for the other tasks.

2.3 | NMF/NMU methods for source extraction as a DR procedure prior to classification

The selected NMF/NMU methods were used as a DR technique to compare the performance of different methods over each classification task at STE with those obtained by Ortega-Martorell et al.²¹

2.3.1 | Classification settings

Standard linear discriminant analysis (LDA with 5-fold cross validation) was used. The dataset was randomly divided into five folds. In each fold, 80% of the dataset was kept as training and 20% was kept as testing. After five folds, the mean performance of five iteration training and testing

were calculated. The main motivation for choosing LDA was that this technique can be seen as a baseline classifier in machine learning, that is, if a classification problem performs well with LDA, then more complex classifiers such as logistic regression, random forests, or artificial neural networks are likely to perform well.

Classifier performance was evaluated using the following measures: balanced error rate (BER), class-based accuracies (ACC), and class-based area under the ROC curve (AUC) with 95% confidence interval (CI) for both training and testing phases. The best performance was defined as the smallest BER in the testing phase.

3 | RESULTS

For the sake of brevity, only the best performing methods for each classification task in the testing phase are presented in this section. All classification results for each task, split by source extraction method, are shown in Data S1. All results of AUC, class-based ACCs, BER with 95% CIs for both training and testing phases, can be found in Data S2. For each classification problem, the best result obtained by the source extraction method for each different number of extracted sources is shown next. Additionally, testing accuracies of each class as bar plot and BER for number of sources against each classification task are presented as line graphs. In the following sections, we organized the results according to each problem, in order.

3.1 | Problem 1 (A2 vs. NO): distinguishing between normal brain and infiltrative tumors

In principle, this was the simplest problem. Unexpectedly, as can be seen in Figure 2, C-NMF was the worst performing method of all, mostly due to failure in the A2 class recognition. Even so, BER in the test set for C-NMF was in the 0.05–0.1 interval for up to seven sources. The other methods had perfect performance in the test set (Figure 2), even at the lowest number of sources. We attribute this to the spectral patterns extracted by C-NMF, with the two sources closely resembling the mean of each of the classes (Figures 3 and 4). By contrast, the other methods extracted two patterns that do not resemble the mean spectra. In all of them, the second source is basically the N-acetylaspartate (NAA) peak, a biomarker for normal brain, as main signal. For C-NMF, as the number of sources increased, more intermediate patterns appeared that could be interpreted through qualitative spectroscopist evaluation (see the red asterisks in Figure 4). For example, for eight sources, numbers one, five, and six seem to emphasize low signal-to-noise ratio (SNR) macromolecular/mobile lipids components, as well as minor contributions from the main metabolite-derived peaks.

3.2 | Problem 2 (A2 vs. ME vs. NO): Can grades II and IV be separated and are they distinct from normal brain?

In this problem, we had three homogeneous classes, very different from each other. In terms of performance (Figure 5), on the one hand, S-NMU and C-NMF behaved best, in a similar way and quite stably across the number of sources extracted. On the other hand, from five sources

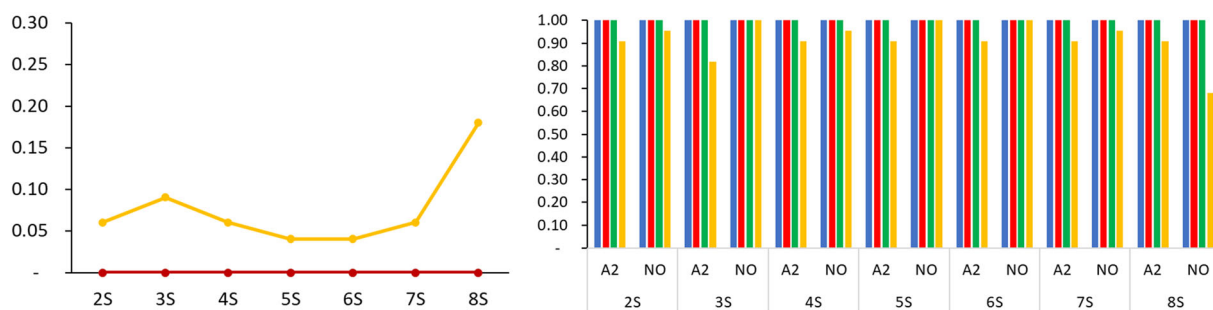
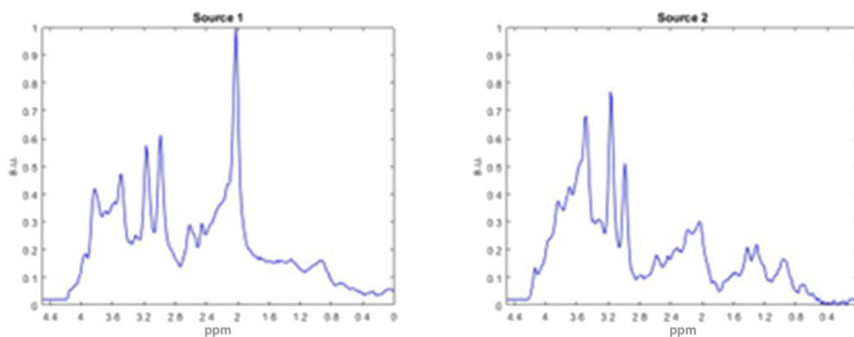
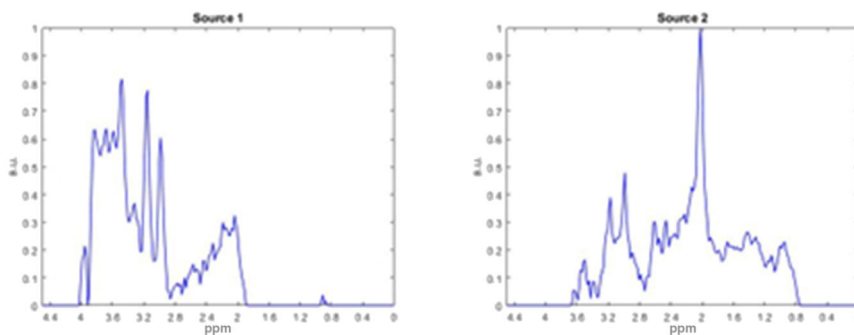


FIGURE 2 Testing performances for Problem 1 (astrocytoma grade II [A2] vs. normal [NO]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). G-NMU, R-NMU, and S-NMU overlapped on the $y = 0$ axis and they are shown in red. C-NMF (yellow). Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = class accuracy. See Data S3 for an enlarged version of the figure.

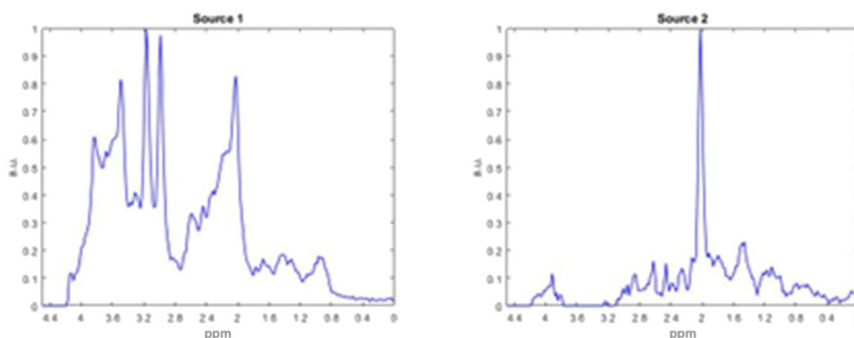
C-NMF



G-NMU



R-NMU



S-NMU

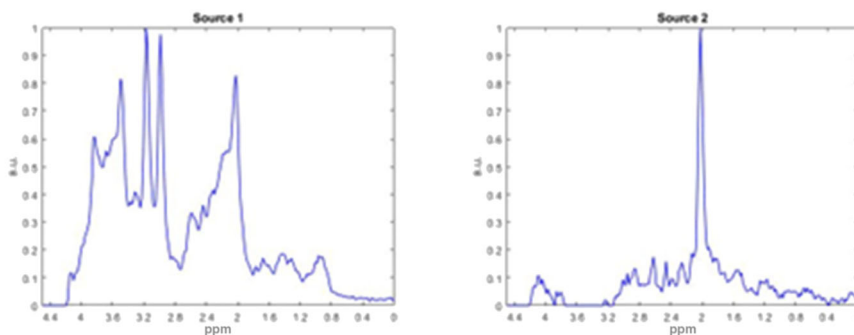


FIGURE 3 Results of two-source extraction for Problem 1 (astrocytoma grade II [A2] vs. normal [NO]), using the different extraction methods. C-NMF, convex non-negative matrix factorization; G-NMU, global non-negative matrix underapproximation; R-NMU, recursive non-negative matrix underapproximation; S-NMU, sparse non-negative matrix underapproximation.

onwards, G-NMU performed worse than the previous two. R-NMU appeared to have a range of sources (between five and six) where it performed the best, whereas if too few sources (4S, BER = 0.240) or too many (7S, BER = 0.193) were extracted, performance dropped notably. Again, for C-NMF, the main drop in BER was for the A2 class, particularly at 3S.

In this problem, the first source (Figure 6) is the characteristic one that would be expected for grade IV tumors, with the main component from necrotic lipids with peaks at 0.9 and 1.28, and lipid-macromolecules between 2 and 2.4 ppm, which would be expected to represent the ME class (see Figure 1). As can be seen, all winning methods extract it in a similar way (visually, the pattern has only small differences in the macromolecules 2–2.4 ppm, for example). The second source of the best performing methods almost always corresponds to the same pattern that appeared using

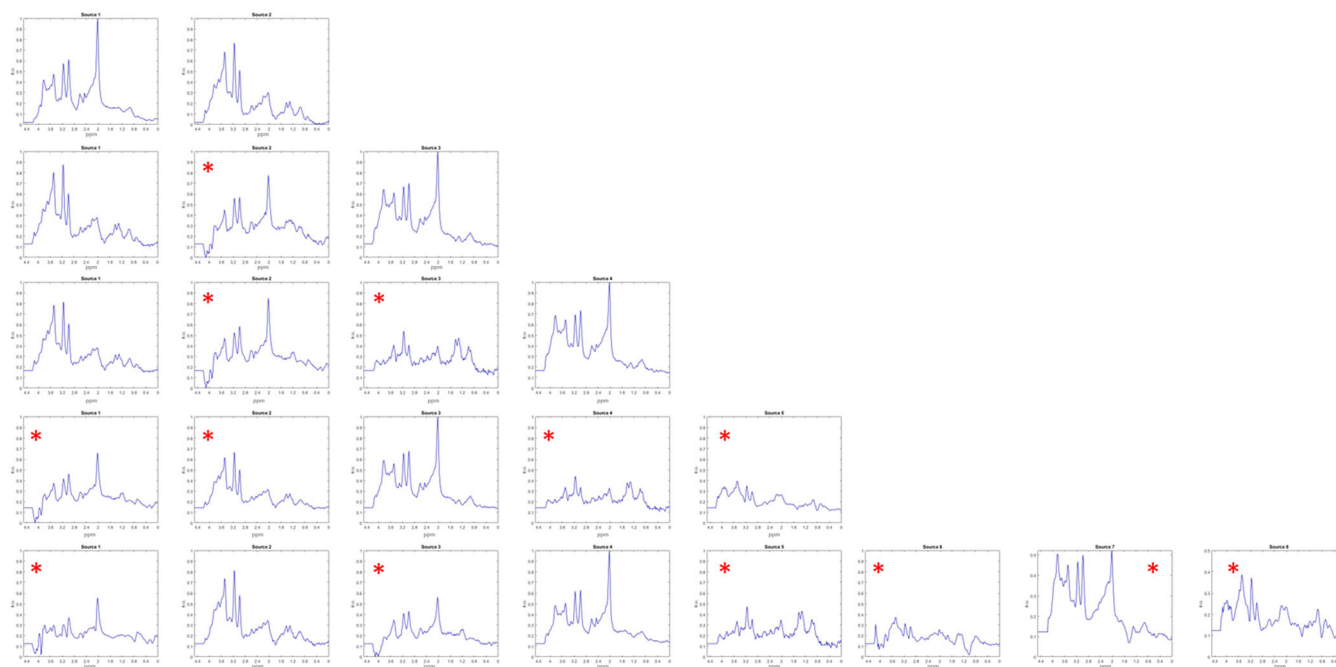


FIGURE 4 Problem 1 solutions using convex non-negative matrix factorization (C-NMF), using two, three, four, five, and eight sources (rows). Red asterisks symbolize intermediate patterns (see text).

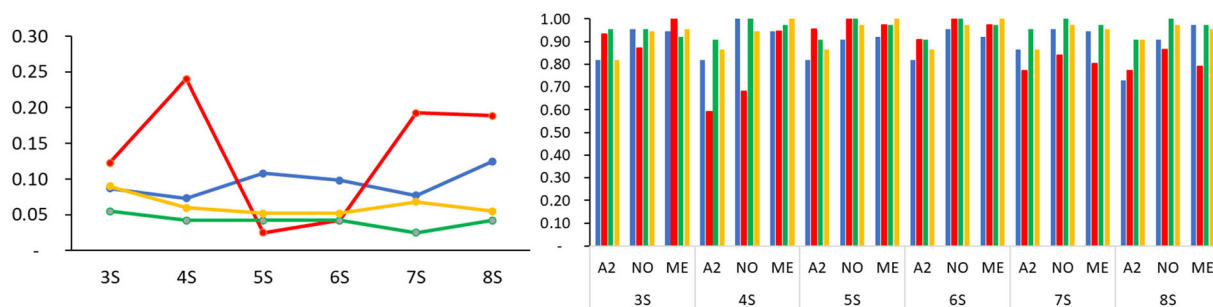


FIGURE 5 Testing performances for Problem 2 (astrocytoma grade II [A2] vs. metastasis [ME] vs. normal [NO]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). See Data S3 for an enlarged version of the figure.

S-NMU for Problem 1 (Figure 3, fourth row): one with peaks of similar height at 3.03 ppm (creatine) and 3.21 (choline-containing compounds) in the same proportion, as well as macromolecules between 2 and 2.4 ppm. The only exception is for four sources (Figure 6), in which the second source corresponds to a very similar pattern to the one already found to be useful for Problem 1: the NAA peak, characteristic of normal brain. Another finding was that S-NMU and R-NMU found sparse sources consisting only of combinations of specific spectral ranges; this can be observed in Figure 6. In those sources marked with a blue asterisk, full contiguous regions have zero values. Note that these “partial” sources always appear after the first two, and if they appear later (third or fourth source), it is after previous “complete” sources (rows 2 and 5 in Figure 6).

3.3 | Problem 3 (A2 vs. GL vs. NO): Can we distinguish grades II and IV while one group (GL) is heterogeneous?

In Problem 3, the important point to consider is that the GL group has high variability in the spectral pattern, which we illustrate in Figure 7, with a few GL resembling class A2 (e.g., case I1098 in Figure 7), in addition to most GL showing the expected pattern, which is very similar to ME

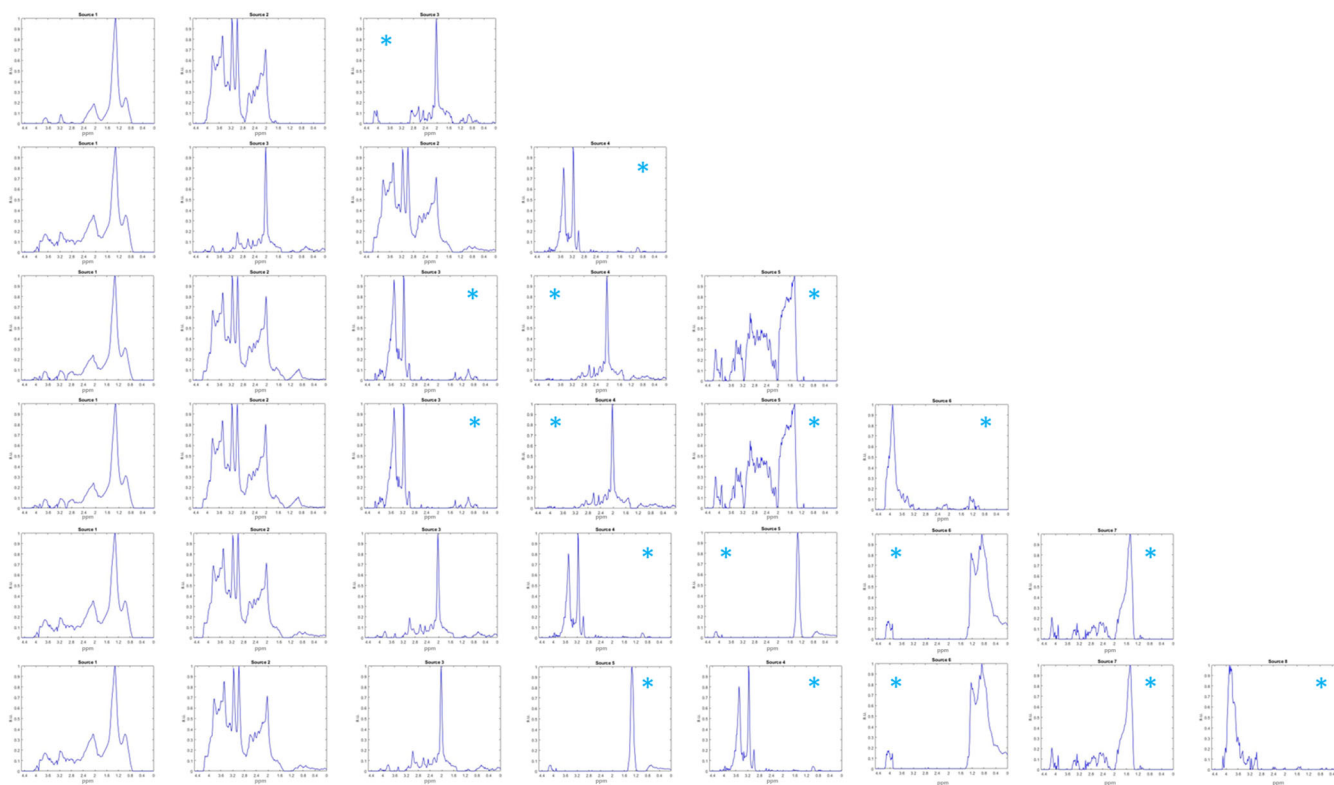


FIGURE 6 Problem 2 (astrocytoma grade II [A2] vs. metastasis [ME] vs. normal [NO]). Best methods for each different number of extracted sources. Each row corresponds to a different source extraction from $K = 3$ to $K = 8$. Rows 1, 2, 5, and 6: sparse non-negative matrix underapproximation (S-NMU). Rows 3 and 4: recursive non-negative matrix underapproximation (R-NMU). Sources that appear to be restricted spectral ranges are marked with a blue asterisk.

(mainly the necrotic lipid pattern at 0.9 and 1.28 ppm, lipid-macromolecules between 2.0 and 2.4 ppm; e.g., case I1032 in Figure 7) and proliferation (choline-containing compounds at 3.21 ppm) in different proportions (e.g., I0407, I0214, I0083, I0183, and I1098 in Figure 7).

S-NMU was the method with the lowest BER for most numbers of extracted sources (Figure 8), except for three sources, requiring at least four to be the winning method. C-NMF was quite constant in BER values across all numbers of sources tested, but was only third best after S-NMU and R-NMU. The worst performing method was G-NMU, particularly for six sources. Interestingly, G-NMU fails to detect NO at five sources (Figure 8, right panel). The first, second, and third sources on the graphs were extremely similar in appearance in all the winning methods for each number of sources extracted by the source-extraction algorithms (Figure 9). The first was the necrotic pattern typical of GL, the second was a pattern essentially of normal brain, perhaps mostly white matter¹ but lacking the NAA peak, which basically was provided by the third source (Figure 9, marked with gray symbol). Again, both S-NMU and R-NMU extracted sparse “partial” spectral ranges at the fourth source and further, which are basically single peaks or sparse parts of spectra, sometimes even redundant from a spectroscopist’s point of view (Figure 9). In fact, according to Figure 8, the gains in BER are marginal for five and six sources in S-NMU and R-NMU. Upon qualitative evaluation, these spectroscopically redundant sources in Figure 9 are indicated with colored dots. The brown dot indicates those sources in which the only contribution is the 1.28 ppm lipid peak (which was also part of the first source). The green dots mark those in which the pattern is again choline-containing compounds and myo-Inositol/glycine (3.55 ppm), which is typical of A2 as well as some GL with low lipids,⁵ and was also present in source 2. The pink dots mark an odd looking (from a spectroscopic point of view) source, again with the necrotic lipid peaks at 0.9 and 1.28 ppm, present in source 1. The blue-dotted ones are more difficult to interpret for a spectroscopist, but in any case, the peak can be assumed to approximately correspond to the second signal of creatine at 3.9 ppm (creatine normally has two peaks, one at 3.03 ppm and the other at 3.9 ppm). Then, two phenomena occurred with different NMU variants: the 3.9 ppm peak was absent from the second source, which apparently was the entire pattern and the two creatine peaks were uncoupled from each other (which does not happen in real spectral signals). Finally, also when the number of sources is eight, the last one (marked with a yellow dot) seems to correspond to the lipid signals that only a few outlier GL spectra contain, with a 0.9 ppm peak of higher intensity than the 1.3 ppm peak (normally it is the reverse). This type of pattern is found in a few cases and has been reported earlier for this dataset.^{27,32} Again, just as in the previous problem and from a spectroscopic point of view, what seems to be happening is that, in order to be the best performing set of sources at each of the numbers extracted, the first two sources capture the most relevant full range

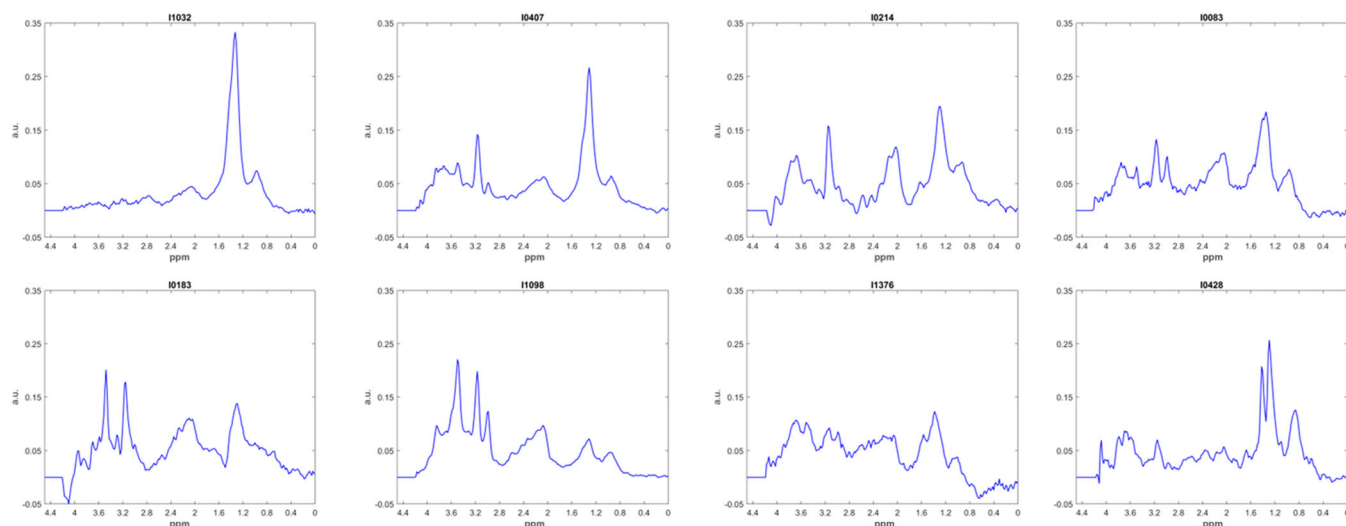


FIGURE 7 Individual examples from the glioblastoma (GL) class (the “lxxxx” code at the top stands for the case code in the INTERPRET database). The first, starting from the left (I1032) on the first row, would be a paradigmatic GL, with necrotic lipid peaks at 0.9 and 1.28 ppm, which is almost impossible to distinguish from a metastasis (ME), from a spectroscopic pattern point of view. The other cases in the first row (I0407, I0214, I0083) show cases with variable contribution from the lipid peaks at 0.9 and 1.28 ppm, lipid-macromolecules (2–2.4 ppm), creatine (3.03 ppm), choline-containing compounds (3.21 ppm), glycine/myo-Inositol (3.55 ppm), or lactate (the doublet between 1.2 and 1.4 ppm present in the example on the left of the bottom row, case I0428).

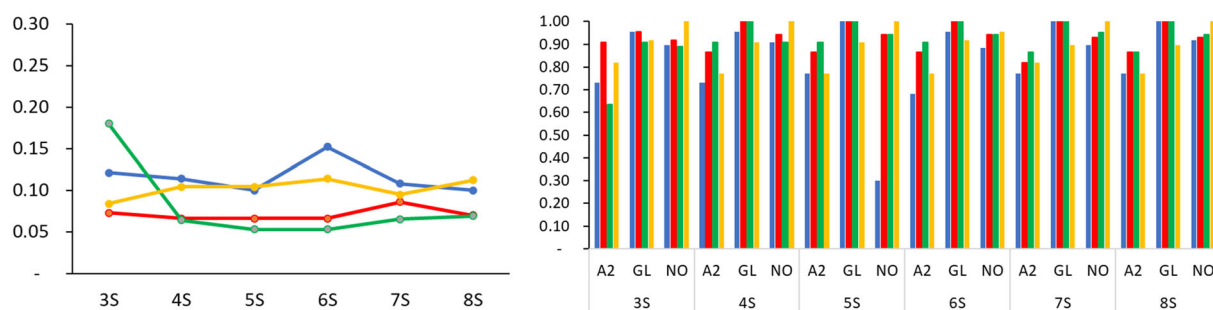


FIGURE 8 Testing performances for Problem 3 (astrocytoma grade II [A2] vs. glioblastoma [GL] vs. normal [NO]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). See Data S3 for an enlarged version of the figure.

(or almost full range) of spectral patterns, and further improvements in BER are provided by NMU variants able to capture specific features (NAA, or sparse spectral regions) that are relevant for fishing out the remaining variability in the dataset.

3.4 | Problem 4 (A2 vs. MM vs. NO): Can low grades of different nature be differentiated?

In Problem 4, we encounter a different situation. To distinguish between types of similar-looking low-grade tumors, three sources do not seem to be sufficient to capture the variability in the dataset, and only using at least four sources, BER values fall below 0.15 for all methods (Figure 10). S-NMU is particularly bad at three sources. For more than four sources, all methods perform similarly overall (BER \leq 0.10), again with G-NMU having an increased BER at six sources. Looking at the right panel of Figure 10, S-NMU particularly fails with either A2 or NO, while C-NMF is the best at only three sources, and is still the best in the class-based accuracies except at seven sources, but only slightly worse than the rest. The analysis of the best sets of sources provides useful information for the spectroscopist (Figure 11: we have labeled some of the sources with different green symbols, according to their resemblance to a typical meningioma pattern). The star symbol is a pattern that looks like the typical meningioma in the whole range (please compare with average class patterns in Figure 1), and it appears as the third source in the five-source

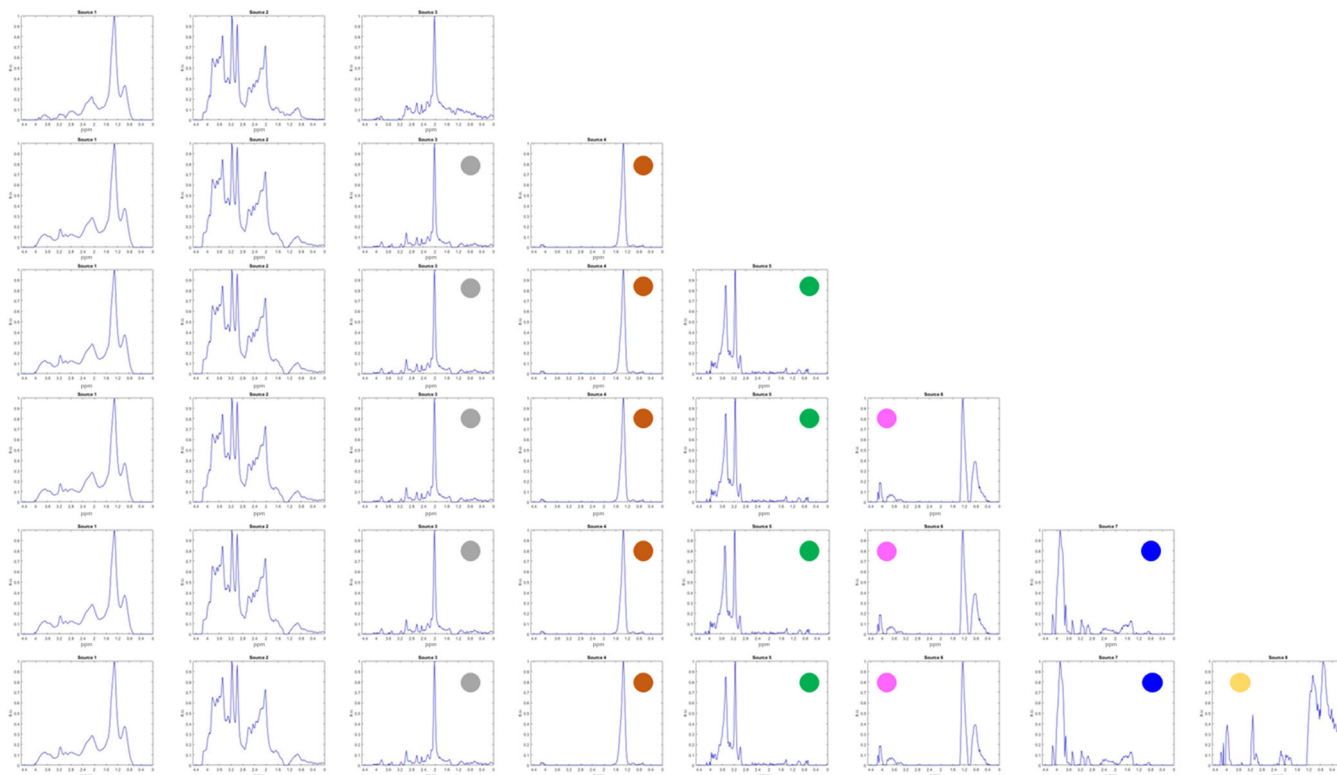


FIGURE 9 Problem 3, astrocytoma grade II (A2) vs. glioblastoma (GL) vs. normal (NO) classification results. Best methods for each different number of extracted sources. Each row corresponds to a different source extraction from $K = 3$ to $K = 8$. Row 1: recursive non-negative matrix underapproximation (R-NMU); rows 2–6: sparse non-negative matrix underapproximation (S-NMU). The colored dots are sources that are mainly nulled for some regions and only represent one or two main peaks. The meaning of the colors of the dots is explained in the main text.

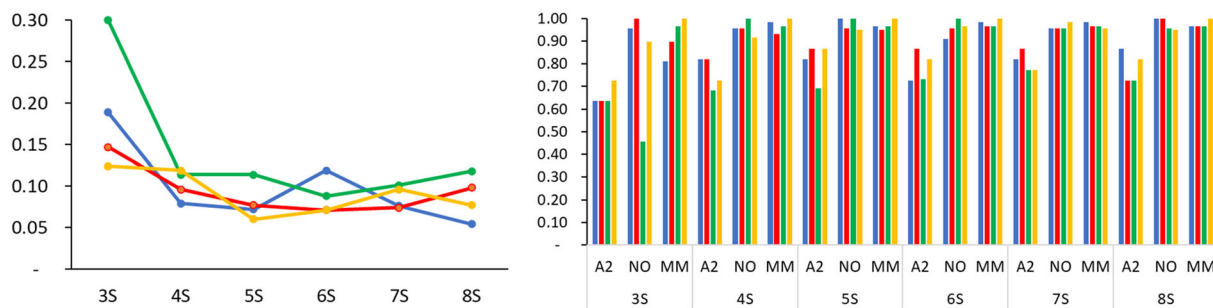


FIGURE 10 Testing performances for Problem 4 (astrocytoma grade II [A2] vs. meningioma [MM] vs. normal [NO]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). See Data S3 for an enlarged version of the figure.

extraction with C-NMF, which is the best for that number of sources. There are two other sources marked with green “plus” symbols, in which the pattern is the whole spectral range, looks very much like meningioma, albeit with too high 0.9 and 1.28 ppm lipid signals for the typical meningioma (this happens in the three-source and again in the five-source extraction, in which C-NMF is the best method). Some more sources are marked with a minus sign, and they appear mainly in the four-, six-, seven-, and eight-source extraction, basically R-NMU and G-NMU; the spectroscopic interpretation is that they capture the 2–2.4 ppm macromolecule region and/or the 3.9 creatine peak region. The five-source solution provided by C-NMF also captures part of the macromolecular feature, but taking all the spectral region. In summary, C-NMF appears to best capture the necessary features for a further successful classification.

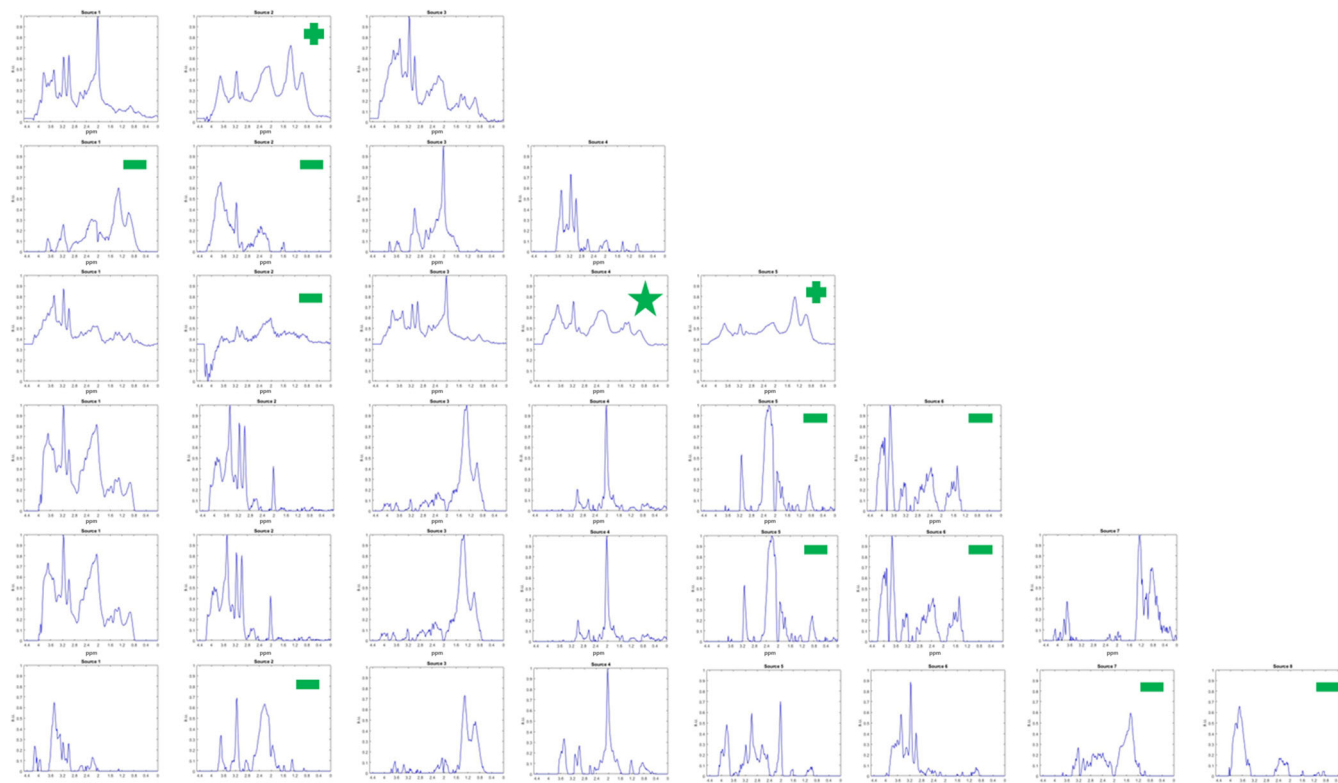


FIGURE 11 Problem 4, astrocytoma grade II (A2) vs. meningioma (MM) vs. normal (NO) classification results—best methods for each different number of extracted sources. Each row corresponds to a different source extraction from $K = 3$ to $K = 8$. Rows 1 and 3: convex non-negative matrix factorization (C-NMF); rows 2 and 6: global non-negative matrix underapproximation (G-NMU); rows 4 and 5: recursive non-negative matrix underapproximation (R-NMU). Green symbols mark those sources that have some resemblance with the pattern of meningiomas; an explanation for each of them is provided in the main text.

3.5 | Problem 5 (A2 vs. AGG vs. NO): Can we distinguish between grades including one heterogeneous class against normal brain?

Problem 5 is a difficult one, as the task involves distinguishing grades, but with one of the classes being naturally heterogeneous. The main finding was that, until seven sources were extracted, there was a remarkable variability in performance ($BER = 0.055$ – 0.298), with local BER minima at four sources (R-NMU $BER = 0.055$, C-NMF $BER = 0.098$, and G-NMU $BER = 0.100$), or five sources ($BER = 0.05$ – 0.26) (Figure 12). At seven sources, results converged to $BER = 0.066$ – 0.095 , with the best method being S-NMU. Eight sources seemed excessive, as BER increased slightly for all the methods.

Results for this problem are similar to those obtained for Problem 3 (Figure 8). Recall that the only difference between Problem 3 and Problem 5 lies in the composition of one of the classes: the AGG tumor class is composed of GL and ME in Problem 5, whereas in Problem 3, we have only GL. In this sense, the aggressive class is still heterogeneous as it contains the heterogeneous GL, but because most ME have the necrotic lipid pattern, we can consider that the AGG class has less heterogeneity in its components, with a higher number of components with the purely mostly necrotic pattern. Still, in Problem 3 we saw that the best method was C-NMF at three sources, S-NMU at four sources, and convergence was reached at seven sources (Figure 8), similar to this problem, in which methods converge to a maximum performance towards seven sources (Figure 12). As expected, we obtained the same type of sources as in Problem 3 (compare Figures 9 and 13, in which we identify several sources that we had also found in Problem 3, note the gray, brown, pink, and blue signs). However, there appears to be one new source (the third one in Figure 13), which did not appear in the best solutions for Problem 3, and again it is a seemingly redundant, odd looking source with necrotic lipids (Figure 13, pink asterisk).

3.6 | Problem 6 (A2 vs. AGG vs. MM): Can we distinguish between grades including one high-grade heterogeneous class and two different low-grade types?

This problem should be considered the most difficult one of those tested in our study: we have the two low grades (A2 and MM) and one heterogeneous class (AGG). As can be seen in Figure 14, the minimum BER at any number of sources is the highest of all the problems, and convergence

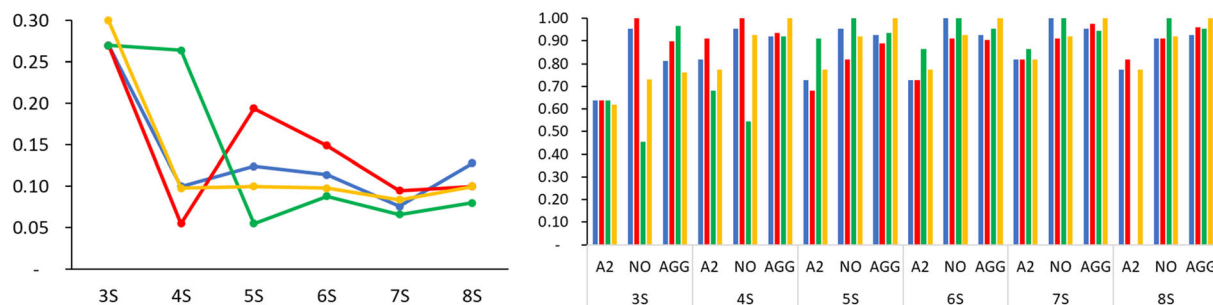


FIGURE 12 Testing performances for Problem 5 (astrocytoma grade II [A2] vs. aggressive [AGG] vs. normal [NO]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). See Data S3 for an enlarged version of the figure.

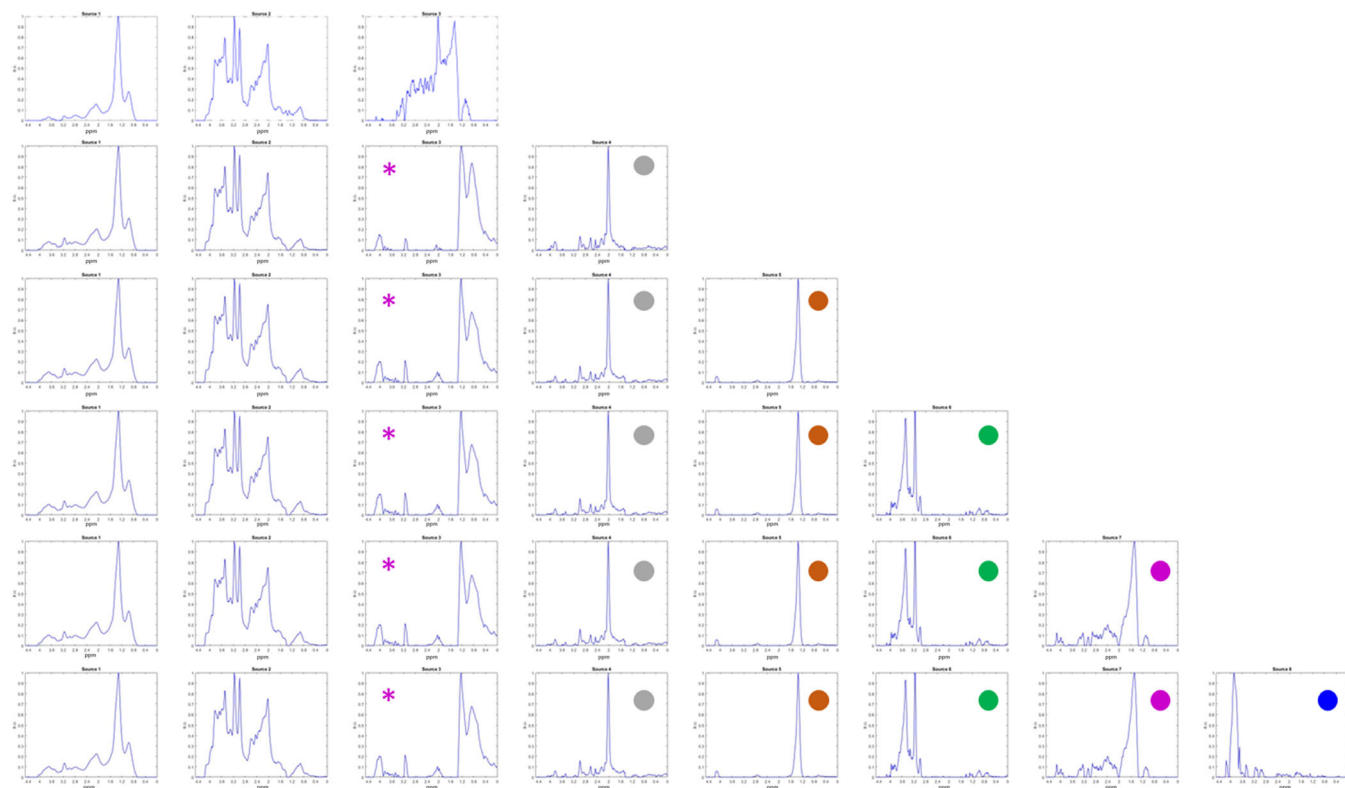


FIGURE 13 Problem 5, astrocytoma grade II (A2) vs. aggressive (AGG) vs. normal (NO) classification results—best methods for each different number of extracted sources. Each row corresponds to a different source extraction from $K = 3$ to $K = 8$. Rows 1 and 2: recursive non-negative matrix underapproximation (R-NMU). Rows 3, 4, 5, and 6: sparse non-negative matrix underapproximation (S-NMU). Colored symbols mark those sources that are explained in the main text.

in performance among methods is not entirely reached. Another finding was that C-NMF was the best performing method overall, except at three and seven sources (3S ranges for BER = 0.223–0.295 and 7S = 0.207–0.143). The best performance of all the combinations was achieved by C-NMF at six sources (BER = 0.126). The greatest variation among methods was at four sources (BER = 0.260–0.160), with C-NMF performing the best. From six to eight sources the worst performing method was R-NMU (6S BER = 0.195, 8S BER = 0.190). In the line graphs, it can also be seen that up to seven sources, each algorithm tends to find their local minima (e.g., the local minimum at seven sources for G-NMU is BER = 0.143). Again, results showed that when one class is MM, the best performance is obtained by C-NMF. Partial decomposition techniques (G-NMU, R-NMU, and S-NMU) failed compared with the full range spectral decomposition method (C-NMF) when one class is highly homogeneous (MM). Regarding the sources extracted, the winning three-source solution is fairly similar in pattern as in Problem 3 (Figure 9), with the

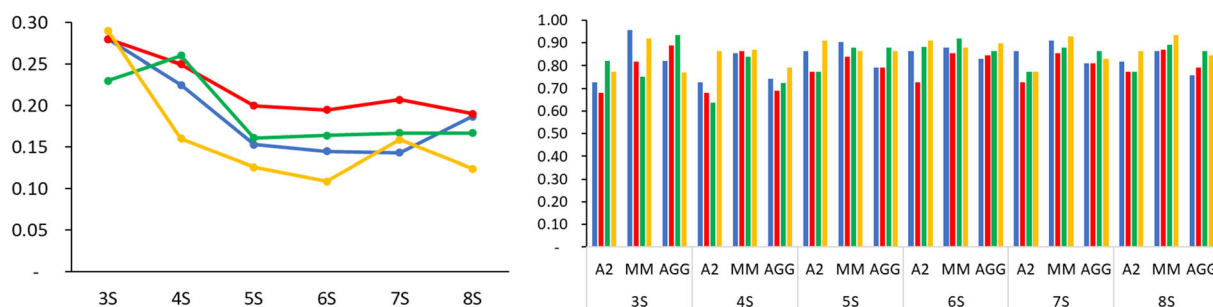


FIGURE 14 Testing performances for Problem 6 (astrocytoma grade II [A2] vs. aggressive [AGG] vs. meningiomas [MM]). Left panel: balanced error rate (BER) for testing set. x-axis = number of extracted sources, y-axis = BER value. Right panel: class-based accuracies, x-axis = classes and number of extracted sources, y-axis = BER value. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow). See Data S3 for an enlarged version of the figure.

necrotic, and low-grade glial patterns, as well as the partial pattern of only NAA, although in Problem 3 the winning method is S-NMU and in Problem 6 it is R-NMU. The four-, five-, six-, and eight-source winning sets of C-NMF solutions contain patterns that resemble the mean meningioma pattern (Figure 15, green star for most similar to the mean spectrum and green diamond for similar to a meningioma pattern), as well as different degrees of necrotic patterns, especially at six sources. There seems to be a repeating trend to extract one low-grade glial pattern plus several meningioma and necrotic patterns as the number of sources increases. At seven sources, G-NMU is the winning method and the patterns extracted are partial, but comprise groups of several nearby peaks (blue asterisk).

To estimate the best number of extracted sources and the best method for all six problems, we applied two-step mean standard deviation (std) analysis. It was found that the five-source solution yielded the lowest mean and std for all problems (Figure 16). Based on this, S-NMU was found to yield the overall lowest BER, as shown in Figure 17. Therefore, S-NMU could be seen as the overall best compromise among all methods and, as a result, the method to be proposed if only one method is to be chosen.

4 | DISCUSSION

4.1 | Performance evaluation of NMU variants against C-NMF

In the current study, we have shown, for the first time, that different NMU variants can perform as well as, or better than C-NMF, for different classification problems using STE MRS data from brain tumor patients. In general, when the number of extracted sources is increased, BER in the test set decreases for NMU variants, whereas C-NMF BER test values tend to increase.

When the minimum number of extracted sources for classification is taken into account, it is shown that R-NMU performs best for Problem 3 (A2 vs. GL vs. NO) with three sources; S-NMU performs best for Problem 2 (A2 vs. ME vs. NO) with three sources; R-NMU performs best for Problem 5 (A2 vs. AGG vs. NO) with four sources; C-NMF performs best for Problem 4 (A2 vs. MM vs. NO) and Problem 6 (A2 vs. AGG vs. MM), with four and three sources, respectively.

In addition, it is shown that, when extracting a number of sources equal to the number of predefined classes for a classification task, NMU methods still outperform C-NMF. The same behavior is observed when the number of sources is increased. These results provide evidence that when the sparsity constraint is added to MRS decomposition methods, the algorithms are able to identify specific ppm regions as sources that best represent the differences between classes, such as in the lipid and NAA regions.

Class-based accuracies in Figures 2, 5, 8, 10, 12, and 14 indicate that there is no optimal combination of number of extracted sources and source extraction algorithm that provides 100% accuracy for a class. However, the methods that most often appear to provide the best performance are NMU variants. S-NMU is the best solution for the A2 and NO classes, while C-NMF is the best for AGG, MM, and ME, and R-NMU for GL.

When using C-NMF, we obtained the same sources (by visual inspection) as in reference,²¹ which were highly correlated with the mean of each class. Comparing the BER performances of all methods, it is found that BER results are not exactly the same values as those reported by Ortega-Martorell et al.²¹ The reason for this is that in Ortega-Martorell et al.,²¹ the authors used the whole INTERPRET set for training the models, and no test set was set aside. By contrast, we applied 5-fold cross-validation, with the training and testing performances being recorded separately. As mentioned by Gillis and Glineur,²⁴ NMU methods are the only ones that are able to extract specific sparse parts of the whole,

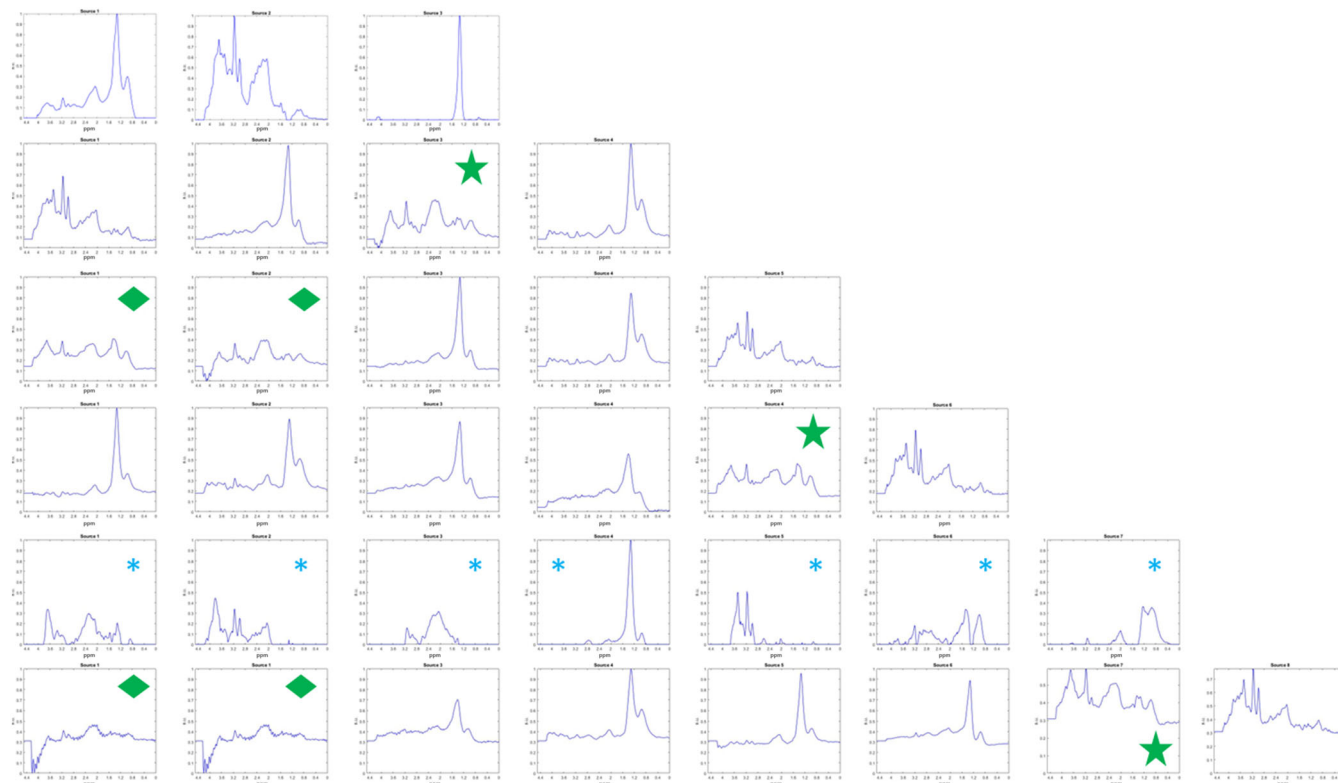


FIGURE 15 Problem 6, astrocytoma grade II (A2) vs. aggressive (AGG) vs. meningiomas (MM) classification results—best methods for each different number of extracted sources. Each row corresponds to a different source extraction from $K = 3$ to $K = 8$. Row 1: sparse non-negative matrix underapproximation (S-NMU); rows 2, 3, 4, and 6: convex non-negative matrix factorization (C-NMF); row 5: global non-negative matrix underapproximation (G-NMU). Green symbols mark those sources that have some resemblance with the pattern of meningiomas; an explanation for each of them is provided in the main text.

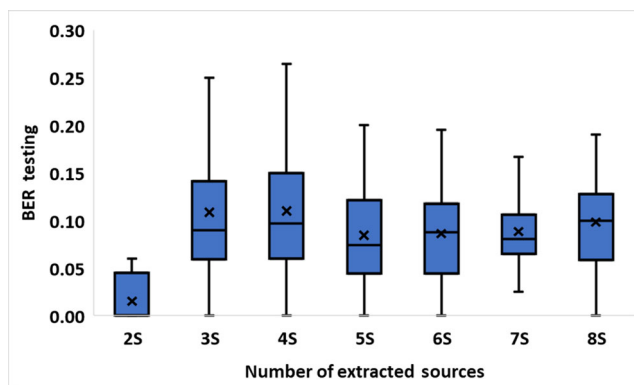


FIGURE 16 Mean and standard deviation for balanced error rate (BER) testing of each number of sources. x-axis = number of extracted sources, y-axis = BER testing value.

whereas NMF methods generate combinations of patterns.²⁴ For this reason, we did not use correlation analysis in our comparisons, as it would be an unfair type of evaluation to compare NMU methods against C-NMF.

In this paper, only STE results are presented, in contrast to the study conducted by Ortega-Martorell et al.²¹ This is because NMU variants require non-negative matrices, therefore using the INTERPRET long echo time dataset, also used by Vilamala et al. and Ortega-Martorell et al.,^{19,21} would also have led to an unfair comparison, because it contains many spectra with negative values (mostly from lactate and alanine, among others³³).

Classification tasks were exclusively performed with LDA. Despite the fact that this is one of the simplest classification algorithms, its use is advisable for two main reasons: first, because it works as a baseline for the problem at hand: good LDA results are a guarantee that the data are

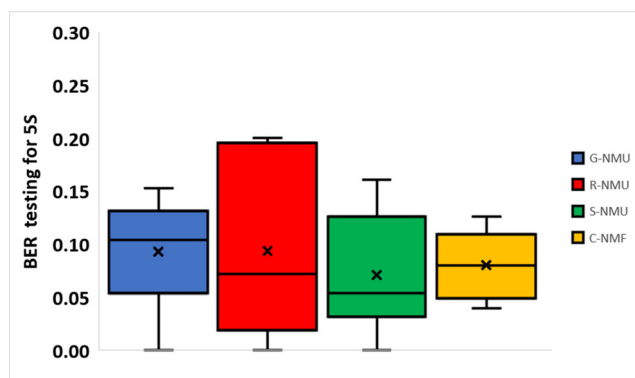


FIGURE 17 Balanced error rate (BER) testing for five-source extraction of each method. The methods are global non-negative matrix underapproximation (G-NMU) (blue), recursive non-negative matrix underapproximation (R-NMU) (red), sparse non-negative matrix underapproximation (S-NMU) (green), and convex non-negative matrix factorization (C-NMF) (yellow).

linearly separable at a reasonable degree; on the contrary, bad results would make the use of more complex nonlinear classifiers advisable. Second, because LDA results are straightforwardly interpretable, which should help the radiologist provide explanations for the diagnostic decisions made by the models. This is key to enable the practical implementations of the models section.³⁴

We can partially compare our results with those reported by Vilamala et al. and Ortega-Martorell et al.,^{19,21} as the same INTERPRET database was analyzed. In Vilamala et al.,¹⁹ the authors evaluated D-C-NMF on the same INTERPRET dataset for the following binary classification tasks: ME versus A2, GL versus A2, GL versus ME, A2 versus NO, ME versus NO, and GL versus NO, at both short and long echo times. At short echo times, the best BER results in cross-validation ranged from 0.03 to 0.13. In the only common experiment (Problem 1), our NMU variants achieved a BER = 0 for all NMU variants. In Vilamala et al. D-C-NMF outperformed C-NMF at both echo times, but it is still a method that finds sources that are not sparse and frequency range-specific, as they span the full spectral range. Regarding the reported results in Ortega-Martorell et al.,²¹ the metric for performance was set to plain accuracy (ACC) for the INTERPRET dataset.^{1,2,26} The number of optimal extracted sources was set to the number of classes. In this aspect, Problem 1 achieved 100% ACC with C-NMF, whereas we achieved 0.00 BER with all NMU variants with two-source extraction methods. In Problem 2, in Ortega-Martorell et al.,²¹ the ACC was 93.8% with three-source C-NMF, and we obtained 0.055 BER with S-NMU three-source extraction. Problem 3 reached 82.1% ACC with three-source C-NMF in Ortega-Martorell et al.,²¹ and in this study 0.073 BER was obtained by R-NMU with the same number of sources. Problem 4 had 95.4% ACC with three-source C-NMF, and we obtained 0.124 BER with the same method and the same number of extracted sources. Problem 5 in²¹ achieved 86.4% ACC with four C-NMF sources, and the same classification task achieved 0.055 BER with R-NMU four-source extraction.

5 | CONCLUSION

In this study, the performance of different NMU model variants for the classification of brain tumors from MRS and in different settings has been assessed for the first time. It has also been compared with the performance of NMF variants more commonly used in this area. Beyond classification as such, we were interested in the potential qualitative differences in the sources extracted from the data by the NMU and NMF models. The unsupervised analysis of SV MRS data from human brain tumors using different NMU methods has shown that their sparsity constraints can provide useful information on specific regions of the MRS spectrum that are relevant for discrimination between tumor types. Importantly, NMU variants achieve better performance than C-NMF for most classification tasks when the source coefficients are used as input for an LDA baseline classifier.

AUTHOR CONTRIBUTIONS

Conceptualization: Alfredo Vellido, Margarida Julià-Sapé, and Carles Arus. *Methodology:* Alfredo Vellido, Margarida Julià-Sapé, and Gulnur Ungan. *Software:* Gulnur Ungan. *Formal analysis:* Gulnur Ungan, Carles Arus, Alfredo Vellido, and Margarida Julià-Sapé. *Data curation:* Margarida Julià-Sapé. *Writing—original draft preparation:* Margarida Julià-Sapé, Alfredo Vellido, Gulnur Ungan, and Carles Arus. All authors have read and agreed to the published version of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORCID

Gulnur Ungan  <https://orcid.org/0000-0002-5436-4665>

Carles Arús  <https://orcid.org/0000-0003-2510-2671>

Alfredo Vellido  <https://orcid.org/0000-0002-9843-1911>

Margarida Julià-Sapé  <https://orcid.org/0000-0002-3316-9027>

REFERENCES

- Tate AR, Underwood J, Acosta DM, et al. Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. *NMR Biomed*. 2006;19:411-434. doi:10.1002/nbm.1016
- Julià-Sapé M, Griffiths JR, Tate AR, et al. Classification of brain tumours from MR spectra: the INTERPRET collaboration and its outcomes. *NMR Biomed*. 2015;28:1772-1787. doi:10.1002/nbm.3439
- Julià-Sapé M, Coronel I, Majós C, et al. Prospective diagnostic performance evaluation of single-voxel 1H MRS for typing and grading of brain tumours. *NMR Biomed*. 2012;25:661-673. doi:10.1002/nbm.1782
- Vellido A, Romero E, Julià-Sapé M, et al. Robust discrimination of glioblastomas from metastatic brain tumors on the basis of single-voxel (1)H MRS. *NMR Biomed*. 2012;25:819-828. doi:10.1002/nbm.1797
- Tate AR, Griffiths JR, Martínez-Pérez I, et al. Towards a method for automated classification of 1H MRS spectra from brain tumours. *NMR Biomed*. 1998;11(4-5):177-191. doi:10.1002/(sici)1099-1492(199806/08)11:4/5%3C177::aid-nbm534%3E3.0.co;2-u
- Cancer of the brain and other nervous system—Cancer Stat Facts, SEER. (n.d.) Accessed February 27, 2023. <https://seer.cancer.gov/statfacts/html/brain.html>
- Wilson M, Andronesi O, Barker PB, et al. Methodological consensus on clinical proton MRS of the brain: review and recommendations. *Magn Reson Med*. 2019;82:527-550. doi:10.1002/mrm.27742
- Lin A, Andronesi O, Bogner W, et al. Minimum reporting standards for in vivo magnetic resonance spectroscopy (MRSinMRS): experts' consensus recommendations. *NMR Biomed*. 2021;34:e4484. doi:10.1002/nbm.4484
- Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw*. 1999;10:626-634. doi:10.1109/72.761722
- Ladroue C, Howe FA, Griffiths JR, Tate AR. Independent component analysis for automated decomposition of in vivo magnetic resonance spectra. *Magn Reson Med*. 2003;50:697-703. doi:10.1002/mrm.10595
- Yang G, Raschke F, Barrick TR, Howe FA. Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering. *Magn Reson Med*. 2015;74(3):868-878. doi:10.1002/mrm.25447
- Hao J, Zou X, Wilson M, et al. A hybrid method of application of independent component analysis to in vivo 1H MR spectra of childhood brain tumours. *NMR Biomed*. 2012;25(4):594-606. doi:10.1002/nbm.1776
- Menze BH, Lichy MP, Bachert P, et al. Optimal classification of long echo time in vivo magnetic resonance spectra in the detection of recurrent brain tumors. *NMR Biomed*. 2006;19:599-609. doi:10.1002/nbm.1041
- Goryawala MZ, Sherif S, Stoyanova R, Maudsley AA. Spectral decomposition for resolving partial volume effects in MRSI. *Magn Reson Med*. 2018;79:2886-2895. doi:10.1002/mrm.26991
- Stoyanova R, Kuesel AC, Brown TR. Application of principal-component analysis for NMR spectral quantitation. *J Magn Reson A*. 1995;115(2):265-269. doi:10.1006/jmra.1995.1177
- Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999;401:788-791. doi:10.1038/44565
- Paatero P, Tapper U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environ*. 1994;5:111-126. doi:10.1002/env.3170050203
- Ding C, Li T, Jordan M. Convex and semi-nonnegative matrix factorizations. *IEEE Trans Pattern Anal Mach Intell*. 2010;32(1):45-55. doi:10.1109/TPAMI.2008.277
- Vilamala A, Lisboa PJG, Ortega-Martorell S, Vellido A. Discriminant convex non-negative matrix factorization for the classification of human brain tumours. *Pattern Recogn Lett*. 2013;34(14):1734-1747. doi:10.1016/j.patrec.2013.05.023
- Sajda P, Du S, Brown TR, et al. Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. *IEEE Trans Med Imaging*. 2004;23(12):1453-1465. doi:10.1109/TMI.2004.834626
- Ortega-Martorell S, Lisboa PJG, Vellido A, Julià-Sapé M, Arús C. Non-negative matrix factorisation methods for the spectral decomposition of MRS data from human brain tumours. *BMC Bioinformatics*. 2012;13(1):38. doi:10.1186/1471-2105-13-38
- Shungu DC, Du S, Mao X, et al. Automated analysis of 1H magnetic resonance metabolic imaging data as an aid to clinical decision-making in the evaluation of intracranial lesions. *Annu Int Conf IEEE Eng Med Biol Soc*. 2007;2007:4327-4330. doi:10.1109/IEMBS.2007.4353294
- Tepper M, Sapiro G. Nonnegative matrix underapproximation for robust multiple model fitting. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Xplore; 2017;655-663. doi:10.1109/CVPR.2017.77
- Gillis N, Glineur F. Using underapproximations for sparse nonnegative matrix factorization. *Pattern Recognition*. 2010;43(4):1676-1687. doi:10.1016/j.patcog.2009.11.013
- Gillis N, Plemmons RJ. Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis. *Linear Algebra Appl*. 2013;438:3991-4007. doi:10.1016/j.laa.2012.04.033
- Julià-Sapé M, Acosta D, Mier M, Arús C, Watson D. INTERPRET consortium. A multi-centre, web-accessible and quality control-checked database of in vivo MR spectra of brain tumour patients. *MAGMA*. 2006;19:22-33. doi:10.1007/s10334-005-0023-x
- García-Gómez JM, Luts J, Julià-Sapé M, et al. Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy. *MAGMA*. 2009;22:5-18. doi:10.1007/s10334-008-0146-y
- Lukas L, Devos A, Suykens JAK, et al. Brain tumor classification based on long echo proton MRS signals. *Artif Intell Med*. 2004;31(1):73-89. doi:10.1016/j.artmed.2004.01.001

29. Luts J, Poulet J-B, Garcia-Gomez JM, et al. Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra. *Magn Reson Med*. 2008;60(2):288-298. doi:[10.1002/mrm.21626](https://doi.org/10.1002/mrm.21626)
30. Devos A, Lukas L, Suykens JAK, et al. Classification of brain tumours using short echo time 1H MR spectra. *J Magn Reson*. 2004;170:164-175. doi:[10.1016/j.jmr.2004.06.010](https://doi.org/10.1016/j.jmr.2004.06.010)
31. Gillis N, Plemmons RJ. Dimensionality reduction, classification, and spectral mixture analysis using nonnegative underapproximation. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVI*. SPIE; 2010:428-440. doi:[10.1117/12.849345](https://doi.org/10.1117/12.849345)
32. Vellido A, Romero E, González-Navarro FF, Belanche-Muñoz LA, Julià-Sapé M, Arús C. Outlier exploration and diagnostic classification of a multi-centre 1H-MRS brain tumour database. *Neurocomputing*. 2009;72(13-15):3085-3097. doi:[10.1016/j.neucom.2009.03.010](https://doi.org/10.1016/j.neucom.2009.03.010)
33. Govindaraju V, Young K, Maudsley AA. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR Biomed*. 2000;13(3):129-153. doi:[10.1002/1099-1492\(200005\)13:3%3C129::aid-nbm619%3E3.0.co;2-v](https://doi.org/10.1002/1099-1492(200005)13:3%3C129::aid-nbm619%3E3.0.co;2-v)
34. Lisboa PJG, Saralajew S, Vellido A, Fernández-Domenech R, Villmann T. The coming of age of interpretable and explainable machine learning models. *Neurocomputing*. 2023;535:25-39. doi:[10.1016/j.neucom.2023.02.040](https://doi.org/10.1016/j.neucom.2023.02.040)

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ugan G, Arús C, Vellido A, Julià-Sapé M. A comparison of non-negative matrix underapproximation methods for the decomposition of magnetic resonance spectroscopy data from human brain tumors. *NMR in Biomedicine*. 2023;e5020. doi:[10.1002/nbm.5020](https://doi.org/10.1002/nbm.5020)