Effective processing of PFG-NMR of mixtures by Blind-Source Separation.

ARTICLE *in* ANALYTICAL CHEMISTRY · OCTOBER 2013

Impact Factor: 5.64 · DOI: 10.1021/ac402085x · Source: PubMed

3 AUTHORS, INCLUDING:

Bruno Torrésani
Aix-Marseille Université
131 PUBLICATIONS **1,849** CITATIONS
SEE PROFILE

Stefano Caldarelli
Aix-Marseille Université
115 PUBLICATIONS **1,494** CITATIONS
SEE PROFILE

analytical chemistry

Article

pubs.acs.org/ac

Effective Processing of Pulse Field Gradient NMR of Mixtures by Blind Source Separation

Ichrak Toumi,[†] Bruno Torrésani,*[,‡] and Stefano Caldarelli*[,†,§]

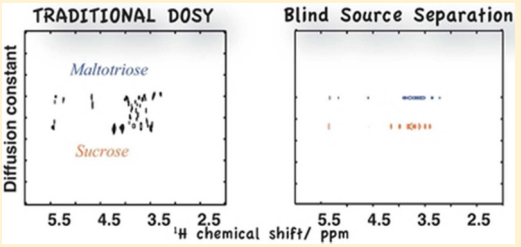
[†]Aix Marseille Université, CNRS, Centrale Marseille, iSm2 UMR 7313, 13397, Marseille, France

[‡]Aix Marseille Université, CNRS, Centrale Marseille, LATP UMR 7353, 13453, Marseille, France

[§]CNRS UPR 2301 ICSN 91190, Gif-sur-Yvette, France

Supporting Information

ABSTRACT: NMR diffusometry and its flagship layout, diffusion-ordered spectroscopy (DOSY), are versatile for studying mixtures of bioorganic and synthetic molecules, but a limiting factor of its applicability is the requirement of a mathematical treatment capable of distinguishing molecules with similar spectra or diffusion constants. We present here a processing strategy for DOSY, a synergy of two high-performance blind source separation (BSS) techniques: non-negative matrix factorization (NMF) using additional sparse conditioning (SC), and the JADE (joint approximate diagonalization of eigenmatrices) declination of independent component analysis (ICA). While the first approach has an intrinsic affinity for NMR data, the latter one can be orders of magnitude computationally faster and can be used to simplify the parametrization of the former.



One-pot analyses of mixtures carry considerable interest. This is particularly true for the associated economy of steps to facilitate fast throughput or to the quest of compounds that are very similar in structure or for all other cases for which appropriate separation protocols are not yet developed. NMR is finding increasing applications in this field, as illustrated in recent reviews.[1] In classic *n*D NMR identification of the mixture components can be achieved by spreading overlapping peaks along a second or further dimension, as, for example, seen in metabolomics.[2] Specific processing alternative to the classic Fourier transform have been devised.[3]

NMR diffusometry,[4] and particularly its diffusion-ordered spectroscopy (DOSY) display,[5] has attracted particular attention for separation purposes for its pictorial and conceptual proximity to liquid chromatography.[6] Recent reviews of the experiment have been published.[7] The performance of the data processing in DOSY is a delicate issue, which is linked to the notorious difficulties of analyzing multiexponential decays.[8] The problem here is associated to the presence of overlapping peaks from species with similar diffusion coefficients. Thus, on one hand, protocols have been proposed to experimentally ameliorate the analysis. The first issue has been addressed using homonuclear decoupling[9] or by adding a selective shift reagent.[10] Spreading the resonances using heteronuclei[11,12] or integrating DOSY into 2D experiment can alleviate the data analysis, but at the cost of longer or more complex experiments.[13,14] The addition of a matrix to selectively enhance the differences in diffusion coefficients of the mixture components has also found increasing applications.[10,15]

At any rate, the processing remains a critical aspect, and Stilbs[16] has recently discussed the issue. Methods based on approximations of the inverse Laplace transform have been suggested,[17,18] but they all tend to produce distributions of solutions around certain diffusion values. While this is not an issue for objects in which such a distribution is an actually property, such as polydispersed polymer samples, it results in an unnecessary low-resolution DOSY spectra for mixtures of molecules.

A consensus for DOSY of discrete mixtures nowadays is around CORE (component-resolved NMR) family of processing,[19] which uses some prior information and an alternate fitting algorithm. Faster and sometimes more accurate results have been achieved by dividing the NMR spectrum in bands, which reduces the number of local minima in the fitting.[16,20,21] Other processing alternatives have been illustrated but did not found widespread application for specific reasons. The DECRA (direct exponential curve resolution algorithm) approach is theoretically very simple[22,23] and computationally fast, but it requires essentially perfect data sets, in terms of negligible contribution from the noise and absence of deviations from the exponential decay, and it requires a specific sampling of the pulsed gradient spin-echo (PGSE) decay, as it creates two data sets that have to be in a strict relation of similarity. Inverse regularized resolvent transform (iRRT)[24] models directly the 2D diagram with the resolution distributed as chosen by the end user. However, it suffers from similar drawbacks as DECRA. Multivariate curve resolution (MCR), a class of algorithms originally developed for chemometrics, has been tested on NMR diffusometry,[25] although its somewhat elaborate implementation may be invoked for not

Received: July 18, 2013
Accepted: October 7, 2013

   dx.doi.org/10.1021/ac402085x | *Anal. Chem.* XXXX, XXX, XXX−XXX

being explored further. Covariance analysis was also tested on DOSY.[26]

In signal analysis, the blind source separation (BSS) is an area of research that encompasses methodologies seeking to decompose a signal mixture into its components by relying on a series of observations and without any knowledge of the shape of these components.

In matrix form this can be represented as

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

where **E** is the matrix of residuals, typically the noise.

Unmixing of NMR spectra of mixtures using DOSY does fall within the scope of BSS, as the signals follow the functional form dictated by the Stejskal–Tanner equation[27]

$$X_{l_\nu k_G} = \sum_i S_{il_\nu} e^{-D_i(\gamma G\delta)^2(\Delta-\varepsilon)} \tag{2}$$

Here, $X_{l_\nu k_G}$ describes the intensity of each frequency ν of the spectrum as a function of the intensity of the gradient pulses, G; S_{il_ν} is the reference spectrum recorded with a very small (ideally null) gradient pulse intensity; D_i are the diffusion coefficients, γ is the gyromagnetic ratio of the observed nucleus, δ is the duration of the gradient pulse, Δ is the time during which diffusion is allowed to take place, and ε is a parameter that depends on the duration of the gradient pulses and their shape. The coefficients D can be recovered by direct modeling or at a later stage by fitting of the **A** matrix.

Following the pioneering work of Nuzillard et al.,[28] BSS has been tested on NMR, including demonstrations on DOSY. The independent component analysis (ICA) approach is probably the most widespread implementation of BSS, due to its lack of parametrization and fast calculation times.[29] Indeed, the first BSS attempt at analyzing NMR data[28] has been SOBI (second-order blind identification), a member of the ICA family.[30] Later, the same authors developed a procedure specifically devised for NMR (linear programming, LP-BSS)[31] and compared its performance to other ICA algorithms, fastICA[32] and JADE (joint approximate diagonalization of eigenmatrices).[33] The interest in ICA as a processing for PGSE-NMR was revisited more recently with a critical evaluation of fastICA for DOSY,[34] that pointed out that this method may have intrinsic resolution much higher than the equivalent fitting, being capable of discriminating diffusion coefficients that differ less than 1% for molecular mobilities typical of small molecules in common solvents.

The very popular PARAFAC (parallel factor analysis)[35] technique was tested in cases in which diffusion was one of the involved dimensions.[36] Sun and co-workers approached the development of BSS techniques specifically adapted to NMR, including possible variations in the signal line widths among the mixture components.[37−40] Two of these methods were tested on model PGSE-NMR examples.[38,40] Finally, a BSS method developed for underdetermined cases was demonstrated, but not on PGSE-NMR.[41]

This large variety of methods suggests that perhaps the perfect and completely general processing method has not been found, if it exists at all. In facts, a typical analysis of DOSY data sets consists often of trying several processing schemes and choosing the one that produces the best results in a given situation. Thus, the general feeling is that the choice of the best processing is problem-dependent and could be a trade-off of computational speed or easiness of application with respect to accuracy of the diffusion measurement and the introduction of artifacts in the calculated "pure" spectra. In the case of BSS, one particular aspect is to assess is the correct number of components, which is often achieved in a systematic fashion, within the limits posed by the size of the data sets (the number of measurements limits the maximum number of components that can be computed) and of the particular algorithm. A second point that remains to be understood is to what degree of complexity each method is best adapted to data at hand. A lack of knowledge on these two issues can induce a time-consuming series of trial-and-error searches of the optimal DOSY processing,

We propose here a strategy for PGSE-NMR analysis that relies on two algorithms from different BSS families: non-negative matrix factorization (NMF, see ref 42 for an application to NMR) with sparse constraints (NNSC), which has never been applied to DOSY, and a revisitation of JADE. In the first approach, sparsity induces an effect similar to the so-called soft thresholding, a concept that is becoming popular in reconstruction of nonuniformly sampled NMR spectra[43] and that found a very recent application to DOSY.[44] The related divergence (i.e., the measure of dissimilarity between the model and the data) reads

$$D(X\backslash AS) = \|X - AS\|^2 + \lambda \sum_r \|S\| \tag{3}$$

The eq 3 bears close resemblance to the one used in an approach introduced very recently,[44] with the noticeable exception that here no modeling of the matrix **A** is performed. NNSC is expected to be particularly well-performing in difficult cases, but relies on proper parametrization, a notoriously difficult task.[18] On the other hand, JADE is likely to produce already good results for simpler problems, and it could be used as a quick preliminary analysis to determine the relevant number of components for more elaborate processing schemes, such as NNSC.

METHODS

Three data sets of ^{1}H PGSE NMR data sets were analyzed: (1) QGC—quinine, geraniol, and camphene, consisting of 30 spectra; (2) DENET—dextran, ethanol, nicotinic acid, ephedrine, and tartrazine, consisting of 30 spectra; (3) SM—sucrose and maltotriose, consisting of 32 spectra.

The first two of these data sets were obtained from the DOSY Toolbox project[45] (from the Web site for the QGC and from the author for DENET) and have already served as a test for past comparisons of the efficiency of the DOSY processing methods. The third data set, a mixture of maltotriose and sucrose, was prepared in deuterated water, about 40 and 50 mmol, respectively. DOSY was acquired using the BPP–LED sequence.[46] Conventional DOSY processing was obtained using the DOSY Toolbox built-in "DOSY" function, which performs a least-squared fit to eq 2 after performing a selection of the relevant peaks in the spectrum above a threshold value selected by the user. BSS analysis was performed in MATLAB, using code made available by the authors for both JADE[47] and NNSC.[48] In both methods the number of components was determined by trial. NNSC requires an adequate choice of the regularization parameter. In the case of the DENET mixture, the spectra were divided in two symmetric halves to accelerate the computation. To all data reference deconvolution[49] was applied, with a Lorentzian broadening of 2 Hz. DOSY charts from JADE and NNSC were constructed calculating first the diffusion coefficients of each source by fitting the estimated mixing matrix **A** to eq 2. The chart was subsequently built by locating the source

dx.doi.org/10.1021/ac402085x | Anal. Chem. XXXX, XXX, XXX–XXX

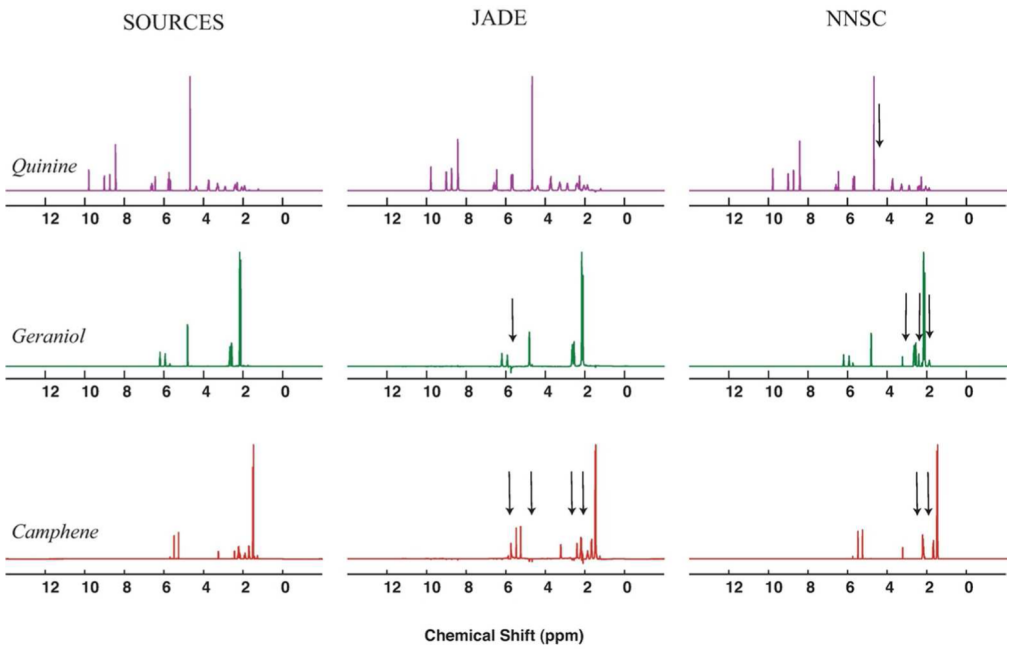


Figure 1. Comparison, for the QGC mixture, of the spectra of the pure components (left) and those calculated by BSS–NNSC (middle) and BSS–JADE (right). The main points of departure of the calculated spectra from the experimental ones are highlighted by the arrows. The calculation time was 0.07 s for JADE and 11 min for NNSC.

spectra at the corresponding value of the diffusion coefficient, with a Gaussian uncertainty equal to the one of the fitting procedure.

RESULTS AND DISCUSSION

One of the difficulties in assessing the resolution quality and the general applicability of a given processing approach of PGSE data is to provide, at the same time, a comparison with the literature and a proper demonstration of challenging cases.

To address this issue, we tested the proposed processing on three data sets. The first two came from an open source repository, associated to the DOSY Toolbox project, and they consist of three and six components, respectively. Since these two data sets present just a moderate degree of spectral overlap, a third example is further analyzed, consisting of the mixture of two oligosaccharides, sucrose and maltotriose, where conversely the spectral overlap is more marked.

A problem in comparing alterative DOSY processing schemes is that, in practice, no simple mathematical criteria have been proposed to this purpose. Often visual analysis is applied, and more specifically detection of distortions in the 2D DOSY diagrams (i.e., NMR peaks of a same molecule calculated to have slightly different diffusion coefficients) or alternatively by identifying the presence of spurious peaks in the case of an unmixing approach. In this latter case, the extra peaks can assume positive or negative values and most often appear as cross-talk at the frequency of peaks of other molecules of the mixture, and are thus the signature of imperfect demixing. Here, we resorted to evaluating the quality of the unmixing by analyzing the number of multiplets correctly estimated in a given spectrum and the artifacts introduced. The values of the calculated diffusion constants are a further indicator of the success of the separation. Figure 1 and Table 1 illustrate the outcome of the NNSC and JADE decomposition on the first test data set, a mixture of quinine, geraniol, and camphor. These three molecules have spectra that only mildly overlap and have made the object of extensive study in DOSY analysis.[21,39,40,45,50] The separation produces good results, with JADE providing almost all peaks of the sources, with the exception of one multiplet for camphene. Artifacts appear in the form of some negative resonances corresponding to peaks of other sources, and they are very rare, and the main ones were observed for camphene at about 2 and 5 ppm, and 5.9 for geraniol. Note that since the outcome of JADE is not constrained to be positive, negative peaks can in principle be also used to accommodate variations in the signal along the PGSE experiment. These demixing artifacts are equivalent to spectral distortions commonly observed in 2D NMR spectroscopy and do not hamper the molecular recognition in this case. Surprisingly, NNSC also provided good unmixing but somewhat less faithful calculated sources (Table 1). Particularly, camphene is not reproduced correctly, with several of its peaks being rather assigned to geraniol as artifacts.

Table 1. Summary of the Performance of the BSS Methods Tested in This Work for the QGC Mixture

	JADE	
compd	% (no.) of correct multiplets	no. of artifacts (>1%)
quinine	100 (19)	1
geraniol	100 (9)	1
camphene	91 (10)	5
	NNSC	
compd	% (no.) of correct peaks	no. of artifacts (>1%)
quinine	89 (17)	0
geraniol	100 (9)	3
camphene	73 (8)	0

The outcome of our BSS approach can be further interpreted by fitting **A** to the Stejskal–Tanner equation[27] (Table 2) and be used to construct a DOSY chart (Supporting Information Figure S1). Note that a more accurate description of the PGSE-NMR experiment, beyond the simple eq 1,[51] can be easily integrated in our approach, since no hypothesis on the form of the mixing

dx.doi.org/10.1021/ac402085x | *Anal. Chem.* XXXX, XXX, XXX–XXX

Table 2. Calculated Values for the Diffusion Coefficients, D ($m^2\ s^{-1}/10^{-10}$), for the QGC Data Set

	quinine	geraniol	camphene
NNSC	7.11	4.63	9.58
JADE	7.2	4.63	9.48
SCORE[a]	8.19	5.06	11.8

[a]From ref 45.

matrix has been assumed. The diffusion values calculated by the two BSS methods are very close (see Table 2 for numerical comparisons) and also very similar to the ones obtained through the classical DOSY fitting routine, with the exception of camphene that is calculated by both BSS methods to displace slower than by the DOSY processing.

The second test data set analyzed (Figure 2), from the same standard collection, contains a larger number of molecules: dextran, tartrazine, ephedrine, nicotinic acid, and ethanol, plus the solvent (HDO).

These compounds have a nice spread of diffusion coefficients, between about 1 and 16×10^{-10} $m^2\ s^{-1}$. The reproduction of the spectra (Figure 2 and Table 3) is very good, with minor artifacts, more abundant in the case of JADE, including the solvent resonance.

The diffusion coefficients are calculated well in agreement with the SCORE procedure (Table 4) for the molecules with slower dynamics. Ethanol mobility is well-estimated by NNSC but not by JADE, which overestimates it. Finally, HDO is also found to be moving faster by JADE, with NNSC assigning a large indetermination to this signal, while remaining in agreement with the DOSY fit (Supporting Information Figure S2).

The final comparison case is a mixture of sucrose and maltotriose, which present a high degree of overlap (indicatively, the two spectra have a correlation value of 41%). Among the cases challenging for the NMR analysis, mixtures of oligosaccharides are notorious,[23] as the anomeric protons are the most likely to provide distinguishable resonances, while the region around 3−4 ppm is often a blurred overlap of complex multiplets. Again, while the JADE decomposition produced sources with a few negative spurious peaks (Figure 3 and Table 5), corresponding to cross-talk from the second source, NNSC reproduced very faithfully the NMR spectra of the pure compounds. Indeed, sucrose spectra are perfectly reproduced while the maltotriose spectrum does still present one minor spurious signal from the other source. The diffusion constants are also calculated in perfect agreement with the DOSY fit (Table 6 and Figure 4).

Table 3. Summary of the Evaluation of the Performance of the BSS Methods Tested in This Work for the DENET Mixture

	JADE	
compd	% (no.) of correct peaks	no. of artifacts (>1%)
dextran	100 (7)	3
tartrazine	100 (3)[a]	0
ephedrine	100 (5)	7
nicotinic acid	100 (4)	6
ethanol	100 (0)	2
	NNSC	
compd	% (no.) of correct peaks	no. of artifacts (>1%)
dextran	100 (7)	0
tartrazine	100 (3)[a]	0
ephedrine	100 (5)	2
nicotinic acid	100 (4)	0
ethanol	100 (2)	0

[a]The spectrum of tartrazine in the repository deviates significantly from the one of the pure compound, and it was estimated from the database http://sdbs.riodb.aist.go.jp/.

Effect of the Signal-to-Noise Ratio on the Decomposition Efficiency. The impact of the noise on the quality of any unmixing procedure cannot be underestimated, but it may be elusive to quantify. This can be achieved by submitting to the separation algorithm an artificial mixing made of known signals with the addition of a given amount of noise. We followed here the approach first proposed by Naanaa and Nuzillard[31] and used to this respect two indexes, introduced, respectively, by Comon[29] and Choi et al.,[52] which both focus on the accuracy of reproduction of the mixing matrix **A**.

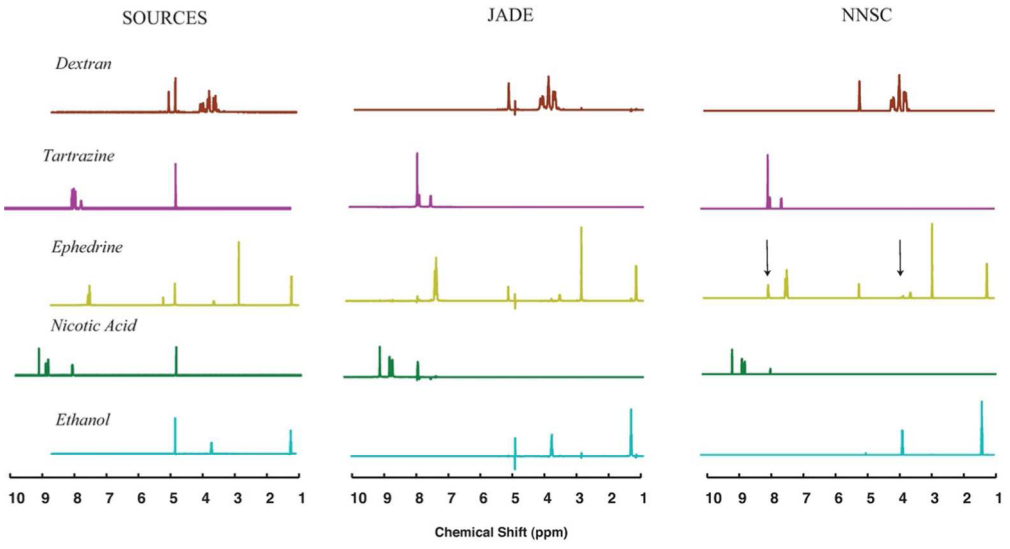


Figure 2. Comparison, for the DENET mixture, of the spectra of the pure components from the repository (left) and those calculated by BSS−JADE (middle) and BSS−NNSC (right). The main points of departure of the calculated spectra from the experimental ones are highlighted by the arrows. The single peak of HDO was not shown for simplicity. The calculation time was 0.08 s for JADE and 11 min for NNSC.

dx.doi.org/10.1021/ac402085x | Anal. Chem. XXXX, XXX, XXX−XXX

Table 4. Calculated and Reference Values for the Diffusion Coefficients, D ($m^2\ s^{-1}/10^{-10}$), for the DENET Data Set

method/source	dextran	ethanol	nicotinic acid	ephedrine	tartrazine	HOD
NNSC	1.02	7.67	5.06	3.31	2.05	15.7
JADE	1.08	9.53	4.92	3.29	2.10	17.1
LOCODOSY[a]	0.79	8.26	5.15	3.36	2.02	16.0

[a]From ref 21.

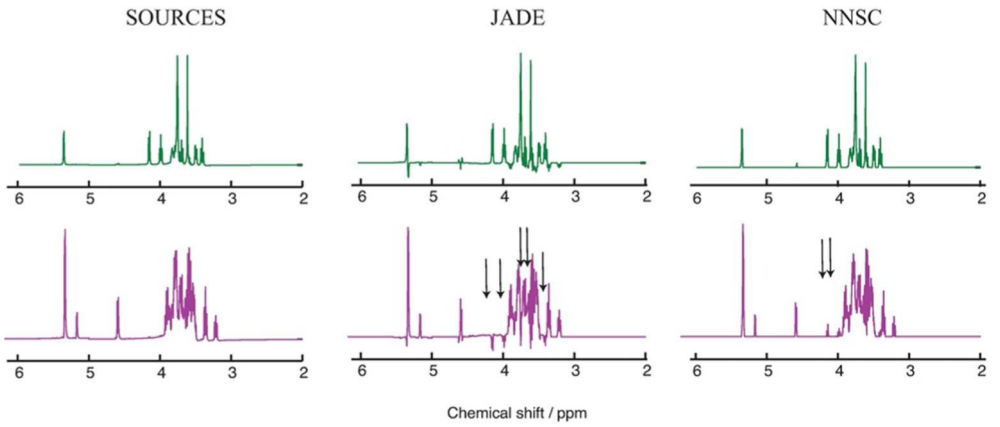


Figure 3. Comparison, for the SM (sucrose and maltotriose) mixture, of the spectra of the pure components (left) and those calculated by BSS–NNSC (middle) and BSS–JADE (right). The main points of departure of the calculated spectra from the experimental ones are highlighted by the arrows. The calculation time was 0.10 s for JADE and 10 min for NNSC.

Table 5. Evaluation of the Performance of the BSS Methods Tested in This Work for the SM Mixture

JADE		
compd	% (no.) of correct peaks	no. of artifacts (>1%)
sucrose	100 (11)	1
maltotriose	100 (8)	3
NNSC		
compd	% (no.) of correct peaks	no. of artifacts (>1%)
sucrose	100 (11)	0
maltotriose	100 (8)	1

Table 6. Calculated Values for the Diffusion Coefficients, D ($m^2\ s^{-1}/10^{-10}$), for the SM Data Set

	sucrose	maltotriose
NNSC	4.24	3.55
JADE	4.15	3.65
SCORE[a]	4.22	3.50

[a]Calculated on the spectra of the pure components.

$$\varepsilon_{\mathrm{Comon}}(A, \hat{A}) = \Sigma_i|\Sigma_j|d_{ij}| - 1|^2 + \Sigma_j|\Sigma_i|d_{ij}| - 1|^2 + \Sigma_i|\Sigma_j|d_{ij}|^2 - 1| + \Sigma_j|\Sigma_i|d_{ij}|^2 - 1| \quad (4)$$

$$\varepsilon_{\mathrm{Choi}}(A, \hat{A}) = \frac{1}{2(n-1)}\Sigma_{i=1}^{n}\left(\sum_{k=1}^{n}\frac{|g_{ik}|^2}{\max_j|g_{ij}|^2} - 1 + \sum_{k=1}^{n}\frac{|g_{ki}|^2}{\max_j|g_{ji}|^2} - 1\right) \quad (5)$$

where g_{ij} and d_{ij} are, respectively, the coefficients of the matrices $G = \hat{A}^{-1}A$ and $D = A^{-1}\hat{A}$, A being the column-normalized version of A.

Since the main objective in the present case is the identification of the mixture component, we added a further measure of performance, based on average distances between the estimated model and experimental spectra of the mixture components, by using the formula

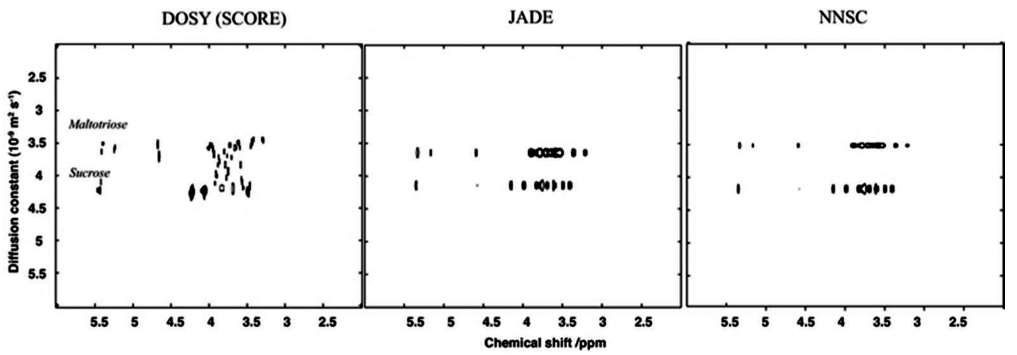


Figure 4. Reconstructed DOSY charts for the SM mixture, obtained by a classic fit, the JADE (middle) and NNSC (right) BSS algorithms. See the text for mode details.

dx.doi.org/10.1021/ac402085x | Anal. Chem. XXXX, XXX, XXX–XXX

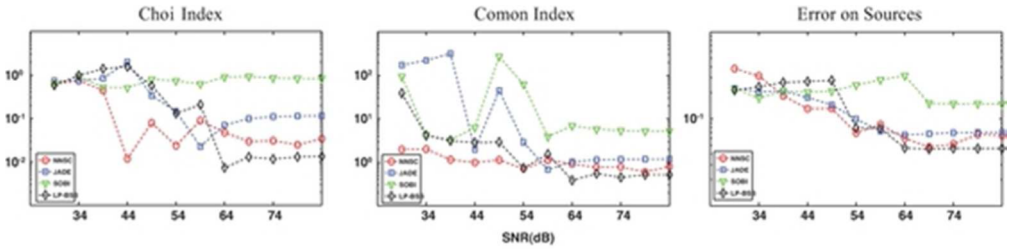


Figure 5. Performance tests of BSS methods applied to a synthetic PGSE-NMR data set of sucrose and maltotriose, with variable white Gaussian noise levels. Left: Comon index, measuring the accuracy of the calculated **A** matrix, see text for details; (middle) the Choi index which also measures the fidelity of the reproduction **A** and (right) of the calculated and original sources as explained in the text.

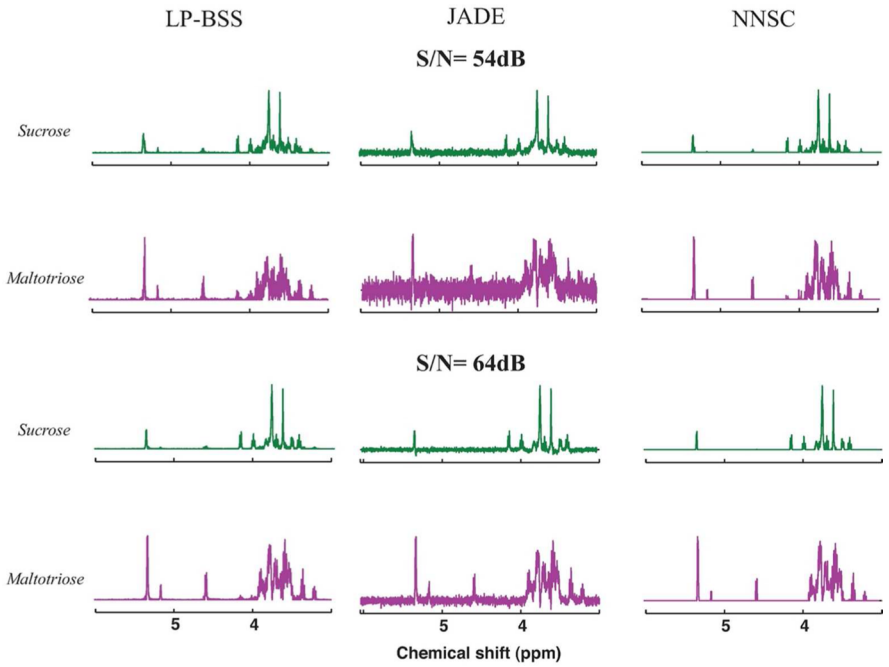


Figure 6. Calculated sources for the demixing test performed on artificial mixtures made of sucrose and maltotriose signal plus a variable amount of noise. The S/N ratio was calculated with respect to the most intense peak in the spectrum, arbitrarily scaled to 1000. The spectra were calculated for a value of S/N of 54 dB (top panels) and 64 dB (bottom panels). Left panels correspond to LP-BSS, middle ones to JADE, and right ones to NNSC.

$$\mathrm{Err} = \log \frac{\| S - \hat{S} \|_p}{\| S \|} \tag{6}$$

where p here is 4, in order to concentrate the comparison to the spectral areas with strong signals. For these tests, synthetic mixtures were created based on the ^{1}H NMR spectra of sucrose and maltotriose, with chosen diffusion coefficients of 4.22 and 3.50×10^{-10} m^2 s^{-1}, respectively. The evaluation was extended to LP-BSS and SOBI, two methods that were providing the best and worst performances in the work of Naanaa and Nuzillard cited above.[31]

The first two panels of Figure 5 represent the Choi and Comon fidelity index, respectively. Although these two do not produce totally coinciding results, some general trends can be inferred. Here, the SOBI approach is confirmed to be the least effective one. The remaining methods perform best for low noise content, starting from around 60 dB. The NNSC method appears to perform best overall, as it is able to produce acceptable results even for slightly lower signal-to-noise levels. On the other hand, for little or no noise content (i.e., S/N > 64 dB), LPBSS is predicted to be the most faithful algorithm. Indeed, the regularization factor introduced in NNSC induces a deviation between real and estimated sources, since all calculated peaks are reduced of an amount proportional to this parameter. Note that the S/N ratio will vary for peaks of different intensity in these spectra, and thus the less intense peak are the one ones that will suffer the highest relative error. These indexes provide a global estimate so that visual inspection (or point by point estimation) can reveal significant distortions in the estimated sources that can go unnoticed.

Figure 6 shows the results obtained by LPBSS, JADE, and NNSC decompositions for S/N equal to 54 and 64 dB. In both cases, the JADE spectra have larger noise content, since no denoising is included in this algorithm. Nevertheless, the reproduction of the sucrose signal is faithful, although with some negative peaks in the case of high S/N. LPBSS produces spectra with good fidelity, although not without artifacts, for S/N = 64 dB. In the case of the lower S/N, the general features of the spectra are well-modeled, but a large number of lower intensity artifacts are observable. NNSC at 64 dB reproduces extremely well both spectra. At higher noise content, the main features are well-reproduced, with a small number of visible artifacts, but less than in the case of LPBSS. Thus, the quality indexes in Figure 5 can be understood as being dominated by the largest values in the spectra, and not representing well the presence of low-intensity artifacts, or noise in the case of JADE for instance.

dx.doi.org/10.1021/ac402085x | Anal. Chem. XXXX, XXX, XXX−XXX

CONCLUSION

The protocol for DOSY analysis using BSS illustrated in this work, and one that can be easily exported to other processing schemes, provided a framework for evaluation of the expected performance of DOSY processing tools.

More specifically for the procedures tested here, the analysis of a series of test cases showed that the very fast JADE algorithm is already a very good tool for processing DOSY of mildly overlapping species, but it does begin to fail for increasing complexity of the separation (i.e., for strongly overlapping signals of moieties with similar diffusion coefficients).

NNSC, tested her for the first time, provides a very good method even for cases with a high degree of complexity, although it is not without imperfections for what should be easy cases. These results point out to the possibility of perfecting the algorithm, work that is currently in progress. The methods share further complementarity, as JADE could be used even in difficult cases to quickly estimate the number of components in the mixture, thus accelerating the NNSC separation procedure. Both tested methods are predicted to fail for too low signal-to-noise levels, with NNSC being more tolerant to this respect.

Finally, it should be noted that the methods demonstrated here could in principle be applied beyond PGSE-NMR, to other mixture analysis, such as metabolomics for instance.

ASSOCIATED CONTENT

Supporting Information

Two figures illustrating the reconstructed DOSY charts for the QGC and DENET mixtures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: s.caldarelli@univ-amu.fr, stefano.caldarelli@cnrs.fr.
*E-mail: bruno.torresani@univ-amu.fr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

ANR (ANR-08-BLAN-273) and Region PACA (APO-G 2009) are gratefully thanked for financial support.

REFERENCES

(1) (a) Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. *Nat. Prod. Rep.* **2011**, *28*, 78−98. (b) McKenzie, J. S.; Donarski, J. A.; Wilson, J. C.; Charlton, A. J. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 336−359.

(2) (a) Beckonert, O.; Coen, M.; Keun, H. C.; Wang, Y.; Ebbels, T. M. D.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Nat. Protoc.* **2010**, *5*, 1019−1032. (b) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536−549.

(3) (a) Robinette, S. L.; Brueschweiler, R.; Schroeder, F. C.; Edison, A. S. *Acc. Chem. Res.* **2012**, *45*, 288−297. (b) Cloarec, O.; Dumas, M. E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. *Anal. Chem.* **2005**, *77*, 1282−1289. (c) Cloarec, O.; Campbell, A.; Tseng, L.-h.; Braumann, U.; Spraul, M.; Scarfe, G.; Weaver, R.; Nicholson, J. K. *Anal. Chem.* **2007**, *79*, 3304−3311. (d) Coen, M.; Hong, Y.-S.; Cloarec, O.; Rhode, C. M.; Reily, M. D.; Robertson, D. G.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Anal. Chem.* **2007**, *79*, 8956−8966.

(4) Stilbs, P. *Anal. Chem.* **1981**, *53*, 2135−2137.

(5) Morris, K. F.; Johnson, C. S. *J. Am. Chem. Soc.* **1992**, *114*, 3139−3141.

(6) Morris, G. A. In *eMagRes* [Online]; John Wiley & Sons, Ltd., 2007, Chichester, U.K.; http://dx.doi.org/10.1002/9780470034590.emrstm0119.pub2 (accessed 11/7/2013).

(7) (a) Callaghan, P. T. *Translational Dynamics & Magnetic Resonance*; Oxford University Press: Oxford, U.K., 2011. (b) Morris, G. In *Multidimensional NMR Methods for the Solution State*; Morris, G. A., Emsley, J. M., Eds.; John Wiley and Sons: Chichester, U.K., 2010; pp 515−532.

(8) Istratov, A. A.; Vyvenko, O. F. *Rev. Sci. Instrum.* **1999**, *70*, 1233−1257.

(9) Nilsson, M.; Morris, G. A. *Chem. Commun.* **2007**, 933−935.

(10) Rogerson, A. K.; Aguilar, J. A.; Nilsson, M.; Morris, G. A. *Chem. Commun.* **2011**, *47*, 7063−7064.

(11) (a) Li, D.; Liu, J.; Williard, P. G. *Abstr. Pap.—Am. Chem. Soc.* **2007**, *233*, 314−314. (b) Li, D.; Hopson, R.; Li, W.; Liu, J.; Williard, P. G. *Org. Lett.* **2008**, *10*, 909−911. (c) Vermillion, K.; Price, N. P. J. *J. Magn. Reson.* **2009**, *198*, 209−214. (d) Botana, A.; Howe, P. W. A.; Caer, V.; Morris, G. A.; Nilsson, M. *J. Magn. Reson.* **2011**, *211*, 25−29.

(12) (a) Stilbs, P.; Moseley, M. E. *Chem. Scr.* **1980**, *15*, 215−216. (b) Moseley, M. E.; Stilbs, P. *Chem. Scr.* **1980**, *16*, 114−116.

(13) Trefi, S.; Gilard, V.; Balayssac, S.; Malet-Martino, M.; Martino, R. *Magn. Reson. Chem.* **2009**, *47*, S163−S173.

(14) (a) Newman, J. M.; Jerschow, A. *Anal. Chem.* **2007**, *79*, 2957−2960. (b) Vitorge, B.; Jeanneat, D. *Anal. Chem.* **2006**, *78*, 5601−5606. (c) Gozansky, E. K.; Gorenstein, D. G. *J. Magn. Reson., Ser. B* **1996**, *111*, 94−96. (d) Jerschow, A.; Muller, N. *J. Magn. Reson., Ser. A* **1996**, *123*, 222−225. (e) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson., Ser. A* **1996**, *121*, 88−91. (f) Barjat, H.; Morris, G. A.; Swanson, A. G. *J. Magn. Reson.* **1998**, *131*, 131−138. (g) Lucas, L. H.; Otto, W. H.; Larive, C. K. *J. Magn. Reson.* **2002**, *156*, 138−145. (h) Nilsson, M.; Gil, A. M.; Delagadillo, I.; Morris, G. A. *Chem. Commun.* **2005**, 1737−1739. (i) Viel, S.; Caldarelli, S. *Chem. Commun.* **2008**, 2013−2015.

(15) (a) Armstrong, D. W.; Ward, T. J.; Berthod, A. *Anal. Chem.* **1986**, *58*, 579−582. (b) Price, W. S. *Concepts Magn. Reson.* **1997**, *9*, 299−336. (c) Price, W. S. *Concepts Magn. Reson.* **1998**, *10*, 197−237. (d) Johnson, C. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203−256. (e) Viel, S.; Ziarelli, F.; Caldarelli, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9696−9698. (f) Pages, G.; Delaurent, C.; Caldarelli, S. *Anal. Chem.* **2006**, *78*, 561−566. (g) Hoffman, R. E.; Arzuan, H.; Pemberton, C.; Aserin, A.; Garti, N. *J. Magn. Reson.* **2008**, *194*, 295−299. (h) Evans, R.; Haiber, S.; Nilsson, M.; Morris, G. A. *Anal. Chem.* **2009**, *81*, 4548−4550. (i) Tormena, C. F.; Evans, R.; Haiber, S.; Nilsson, M.; Morris, G. A. *Magn. Reson. Chem.* **2010**, *48*, 550−553. (j) Adams, R. W.; Aguilar, J. A.; Cassani, J.; Morris, G. A.; Nilsson, M. *Org. Biomol. Chem.* **2011**, *9*, 7062−7064. (k) Asaro, F.; Savko, N. *Magn. Reson. Chem.* **2011**, *49*, 195−198. (l) Pemberton, C.; Hoffman, R.; Aserin, A.; Garti, N. *J. Magn. Reson.* **2011**, *208*, 262−269. (m) Pemberton, C.; Hoffman, R. E.; Aserin, A.; Garti, N. *Langmuir* **2011**, *27*, 4497−4504. (n) Manjunatha Reddy, G. N.; Ballesteros-Garrido, R.; Lacour, J.; Caldarelli, S. *Angew. Chem., Int. Ed.* **2013**, *52*, 3255−3258.

(16) Stilbs, P. *Eur. Biophys. J.* **2012**, *42*, 25−32.

(17) (a) Provencher, S. W.; Vogel, R. H. In *Numerical Treatment of Inverse Problems in Differential and Integral Equations*; Deuflhard, P., Hairer, E., Eds.; Birkhaüser: Boston, MA, 1983; pp 304−319; (b) Provencher, S. W. *Comput. Phys. Commun.* **1982**, *27*, 229−242. (c) Delsuc, M. A.; Malliavin, T. E. *Anal. Chem.* **1998**, *70*, 2146−2148.

(18) Day, I. J. *J. Magn. Reson.* **2011**, *211*, 178−185.

(19) Stilbs, P.; Paulsen, K.; Griffiths, P. C. *J. Phys. Chem.* **1996**, *100*, 8180−8189.

(20) Stilbs, P. *J. Magn. Reson.* **2010**, *207*, 332−336.

(21) Colbourne, A. A.; Morris, G. A.; Nilsson, M. *J. Am. Chem. Soc.* **2011**, *133*, 7640−7643.

(22) Windig, W.; Antalek, B. *Chemom. Intell. Lab. Syst.* **1997**, *37*, 241−254.

(23) Antalek, B. *Concepts Magn. Reson.* **2002**, *14*, 225−258.

(24) (a) Armstrong, G. S.; Loening, N. M.; Curtis, J. E.; Shaka, A. J.; Mandelshtam, V. A. *J. Magn. Reson.* **2003**, *163*, 139−148. (b) Thureau, P.; Thevand, A.; Ancian, B.; Escavabaja, P.; Armstrong, G. S.; Mandelshtam, V. A. *ChemPhysChem* **2005**, *6*, 1510−1513.

(25) (a) Van Gorkom, L. C. M.; Hancewicz, T. M. *J. Magn. Reson.* **1998**, *130*, 125−130. (b) Huo, R.; Wehrens, R.; van Duynhoven, J.; Buydens, L. M. C. *Anal. Chim. Acta* **2003**, *490*, 231−251. (c) Huo, R. ; Wehrens, R.; Buydens, L. M. C. *J. Magn. Reson.* **2004**, *169*, 257−269. (d) Huo, R.; van de Molengraaf, R.; Pikkemaat, J. A.; Wehrens, R.; Buydens. *J. Magn. Reson.* **2005**, *172*, 346−358. (e) Huo, R.; Geurts, C.; Brands, J.; Wehrens, R.; Buydens, L. M. C. *Magn. Reson. Chem.* **2006**, *44*, 110−117. (f) Huo, R.; Wehrens, R.; Buydens, L. M. C. *Chemom. Intell. Lab. Syst.* **2007**, *85*, 9−19.
(26) Smith, L. M.; Maher, A. D.; Cloarec, O.; Rantalainen, M.; Tang, H.; Elliott, P.; Stamler, J.; Lindon, J. C.; Holmes, E.; Nicholson, J. K. *Anal. Chem.* **2007**, *79*, 5682−5689.
(27) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *41*, 288−292.
(28) Nuzillard, D.; Bourg, S.; Nuzillard, J. M. *J. Magn. Reson.* **1998**, *133*, 358−363.
(29) Comon, P. *Signal Process.* **1994**, *36*, 287−384.
(30) Belouchrani, A.; Abed-Meraim, K.; Cardoso, J.-F.; Moulines, E. *IEEE Trans. Signal Process.* **1987**, 443−444.
(31) Naanaa, W.; Nuzillard, J. M. *Signal Process.* **2005**, *85*, 1711−1722.
(32) Hyvarinen, A. *IEEE Trans. Neural Networks* **1999**, *10*, 626−634.
(33) Cardoso, J.-F.; Souloumiac, A. *IEE Proc. F* **1993**, *140*, 362−370.
(34) Zhong, J.; DiDonato, N.; Hatcher, P. G. *J. Chemom.* **2012**, *26*, 150−157.
(35) Harshman, R. A. *UCLA Working Papers in Phonetics* **1970**, *16*, 1−84.
(36) (a) Bro, R.; Hansen, P. I.; Viereck, N.; Dyrby, M.; Pedersen, H. T.; Engelsen, S. B. A new principle for unique spectral decomposition of 2D NMR data. In *Magnetic Resonance in Food Science: The Multivariate Challenge*; Engelsen, S. B., Belton, P. S., Jakobsen, H. J., Eds.; Royal Society of Chemistry: Cambridge, U.K., 2005; pp 195−203. (b) Pedersen, H. T.; Dyrby, M.; Engelsen, S. B.; Bro, R. *Annu. Rep. NMR Spectrosc.* **2006**, *59*, 207−233. (c) Bro, R.; Viereck, N.; Toft, M.; Toft, H.; Hansen, P. I.; Engelsen, S. B. *TrAC, Trends Anal. Chem.* **2010**, *29*, 281−284. (d) Nilsson, M.; Botana, A.; Morris, G. A. *Anal. Chem.* **2009**, *81*, 8119−8125. (e) Khajeh, M.; Botana, A.; Bernstein, M. A.; Nilsson, M.; Morris, G. A. *Anal. Chem.* **2010**, *82*, 2102−2108. (f) Björnerås, J.; Botana, A.; Morris, G.; Nilsson, M. *J. Biomol. NMR* **2013**, 1−7.
(37) Sun, Y.; Ridge, C.; del Rio, F.; Shaka, A. J.; Xin, J. *Signal Process.* **2011**, *91*, 1838−1851.
(38) Sun, Y.; Xin, J. *J. Sci. Comput.* **2011**, *51*, 733−753.
(39) Sun, Y.; Xin, J. *Siam J. Imaging Sci.* **2012**, *5*, 886−911.
(40) Sun, Y.; Xin, J. In *Computer Analysis of Images and Patterns: 14th International Conference, Caip 2011, Part 2*; Real, P., DiazPernil, D., MolinaAbril, H., Berciano, A., Kropatsch, W., Eds.; Springer: Berlin Heidelberg, Germany, 2012, pp 81−88.
(41) (a) Kopriva, I.; Jeric, I.; Smrecki, V. *Anal. Chim. Acta* **2009**, *653*, 143−153. (b) Kopriva, I.; Jeric, I. *Anal. Chem.* **2010**, *82*, 1911−1920.
(42) Snyder, D. A.; Zhang, F.; Robinette, S. L.; Bruschweiler-Li, L.; Brueschweiler, R. *J. Chem. Phys.* **2008**, *128*, 052313/1−052313/4.
(43) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. *J. Biomol. NMR* **2012**, *52*, 315−327.
(44) Urbanczyk, M.; Bernin, D.; Kozminski, W.; Kazimierczuk, K. *Anal. Chem.* **2013**, *85*, 1828−1833.
(45) Nilsson, M. *J. Magn. Reson.* **2009**, *200*, 296−302.
(46) Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson., Ser. A* **1995**, *115*, 260−264.
(47) Cardoso, J.-F. http://perso.telecom-paristech.fr/~cardoso/guidesepsou.html.
(48) Hoyer, P. O. http://www.cs.helsinki.fi/u/phoyer/software.html.
(49) Morris, G. A. B.,H.; Horne, T. J. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *31*, 197−257.
(50) Sun, Y.; Xin, J.; arXiv:1110.1676 [math.NA], 2011.
(51) Evans, R.; Deng, Z. X.; Rogerson, A. K.; McLachlan, A. S.; Richards, J. J.; Nilsson, M.; Morris, G. A. *Angew. Chem., Int. Ed.* **2013**, *52*, 3199−3202.
(52) Choi, S.; Cichocki, A.; Park, H.; Lee, S. *Neural Inform. Process. Lett. Rev.* **2005**, 1−57.