

Water removal in MR spectroscopic imaging with Casorati singular value decomposition

Amirmohammad Shamaei^{1,2}   | Jana Starcukova¹ | Rudy Rizzo^{3,4,5} | Zenon Starcuk Jr¹ 

¹Institute of Scientific Instruments of the Czech Academy of Sciences, Brno, Czech Republic

²Department of Electrical and Software Engineering, Schulich School of Engineering, University of Calgary, Calgary, Alberta Canada

³MR Methodology, Department of Interventional Neuroradiology, University of Bern, Bern, Switzerland

⁴Department of Biomedical Research, University of Bern, Bern, Switzerland

⁵Translational Imaging Center (TIC), Swiss Institute of Translational Entrepreneurial Medicine, Bern, Switzerland

Correspondence

Amirmohammad Shamaei, Magnetic Resonance Group, Institute of Scientific Instruments of the Czech Academy of Sciences, Brno, Czech Republic.
Email: amirmohammad.shamaei@ucalgary.ca

Funding information

Czech Academy of Sciences, Institute of Scientific Instruments, Grant/Award Number: RVO:68081731; ISI-MR Facility of the Czech-BioImaging Infrastructure, Grant/Award Numbers: CZ.02.1.01/0.0/0.0/18_046/0016045, LM2018129, LM2023050; The Marie Skłodowska-Curie Grant Agreement (INSPIRE-MED), Grant/Award Number: 813120

Abstract

Purpose: Water removal is one of the computational bottlenecks in the processing of high-resolution MRSI data. The purpose of this work is to propose an approach to reduce the computing time required for water removal in large MRS data.

Methods: In this work, we describe a singular value decomposition-based approach that uses the partial position-time separability and the time-domain linear predictability of MRSI data to reduce the computational time required for water removal. Our approach arranges MRS signals in a Casorati matrix form, applies low-rank approximations utilizing singular value decomposition, removes residual water from the most prominent left-singular vectors, and finally reconstructs the water-free matrix using the processed left-singular vectors.

Results: We have demonstrated the effectiveness of our proposed algorithm for water removal using both simulated and in vivo data. The proposed algorithm encompasses a pip-installable tool (<https://pypi.org/project/CSVD/>), available on GitHub (<https://github.com/amirshamaei/CSVD>), empowering researchers to use it in future studies. Additionally, to further promote transparency and reproducibility, we provide comprehensive code for result replication.

Conclusions: The findings of this study suggest that the proposed method is a promising alternative to existing water removal methods due to its low processing time and good performance in removing water signals.

KEYWORDS

functional MRS, low-rank approximations, MR spectroscopic imaging, water removal, water suppression

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 Ustav pristrojove techniky Akademie ved Ceske republiky and University of Bern. *Magnetic Resonance in Medicine* published by Wiley Periodicals LLC on behalf of International Society for Magnetic Resonance in Medicine.

1 | INTRODUCTION

In proton magnetic resonance spectroscopy (MRS) and magnetic resonance spectroscopic imaging (MRSI), water signals can seriously interfere with the detection and quantification of low molecular weight metabolites. In most experiments, some physical water-suppression technique, such as chemical shift selective or the variable power and optimized relaxations delays methods,^{1,2} is employed to selectively suppress the water signal and thus enhance the detection of other metabolites.¹

However, achieving complete suppression is often difficult due to the combined effect of the overwhelming concentration of water, the inhomogeneity of the static (B_0) or the excitation (B_1) fields, or the distribution of water T_1 values.^{2,3} Residual water signals are a frequent problem, especially in regions with poor B_0 homogeneity,³ because the variability of their shape may cause errors in metabolite quantification. Thus, postprocessing methods of water suppression play an important role in improving the accuracy and reliability of MRS data analysis.³

According to experts' consensus recommendations,³ the residual water signal can be handled in two distinct ways. The first involves eliminating the remaining water signal before conducting spectral analysis. This can be achieved by fitting the peak to a lineshape function, typically Gaussian, Lorentzian, or a combination of both known as *Voigt*. Alternatively, the peak can be fitted to a set of lineshape components using singular value decomposition (SVD), and the obtained fit can then be subtracted from the spectrum. The second approach involves retaining the water peak but utilizing a fitting model that includes either the water peak itself or the sloping baseline caused by the residual water peak.

SVD-based approaches such as HLSVDPRO⁴ are extensively utilized for modeling and removing the water signal on the premise that MRS signals in the time domain can be characterized by a model consisting of a small number of exponentially damped complex exponentials. After finding the signal model, they subtract the components found within a particular water frequency range from the original signal.^{1,4,5}

The calculation of the SVD is computationally demanding. Moreover, in large MRS datasets, such as MRSI data, the SVD of the Hankel matrix should be calculated for each signal. Thus, water removal is one of the computational bottlenecks in the processing of high-resolution MRSI data.⁶ In this work, we propose an SVD-based approach to reduce the computing time required for water removal in large MRS data.

Our proposed algorithm exploits the spatiotemporal partial separability and the linear predictability of MRSI data.⁷ Our approach arranges MRSI data in a

Casorati matrix form, applies low-rank approximations utilizing SVD, removes residual water from the most prominent left-singular vectors, and finally reconstructs the water-free matrix using the processed left-singular vectors.

We compared our proposed method, Casorati singular value decomposition (CSVD), against the Hankel-Lanczos implementation of the SVD method⁴ in terms of accuracy and time efficiency utilizing simulated data with known ground truth (GT) values and the in vivo data.

We provide a pip-installable Python tool for our proposed algorithm and make codes for generating our results available online at <https://github.com/amirshamaei/CSVD>

2 | METHODS

2.1 | CSVD algorithm for residual water signal removal

Given measured MRSI signals, the proposed algorithm treats the set of spectroscopic signals together before the individual-voxel water models are subtracted from the original signals. It is described in the following steps:

Step 1. Arrange signals ($s(x_i, t_j)$) from voxel x_i at time t_j) in a Casorati matrix C as follows:

$$C = \begin{bmatrix} s(x_1, t_1) & s(x_2, t_1) & \cdots & s(x_M, t_1) \\ s(x_1, t_2) & s(x_2, t_2) & \cdots & s(x_M, t_2) \\ \cdots & \cdots & \cdots & \cdots \\ s(x_1, t_N) & s(x_2, t_N) & \cdots & s(x_M, t_N) \end{bmatrix}, \quad (1)$$

where N and M are the number of time points and the number of voxels, respectively. In other words, the matrix C is constructed by stacking the signal from each voxel in its respective column.

Step 2. Compute the SVD of C to obtain U , Σ , and V , that is,

$$C = U\Sigma V^*, \quad (2)$$

where U is an $n \times m$ complex unitary matrix; Σ is an $n \times m$ rectangular diagonal real matrix with nonnegative real numbers on the diagonal in descending order (the singular values of C); V is an $n \times m$ complex unitary matrix; and $*$ is the conjugate transpose. The columns of U and V are then called the left- and the right-singular vectors of C , respectively.

Step 3. Determine the rank, r , of C . In this work, the optimal singular value hard thresholding method⁸ was utilized for automatically estimating the rank of matrix C .

Step 4. Apply HLSVDPRO⁴ in each of the first r columns of U (r -left-singular vectors); that is, arrange the N data points of column i of U in a Hankel matrix, and remove residual water signal using HLSVDPRO method within a particular water frequency range. We will refer to matrix U as \hat{U} after removing water from the first r columns.

Step 5. Reconstruct the water-free matrix \hat{C} as follows:

$$\hat{C} = \hat{U}\Sigma V^* \quad (3)$$

The proposed CSVD-based water removal algorithm is computationally efficient due to the low rank of C ; that is, we apply the HLSVDPRO algorithm only on the few columns (r) of U instead of all signals ($r \ll$) (Figure 1).

2.2 | Datasets

2.2.1 | Simulated dataset

In this work, a simulated MRSI dataset was generated by a linear combination of amplitude-scaled metabolite signals, baseline, and noise. The model describing a time-domain MRS signal $S(t)$ as a combination of several metabolite profiles defined by

$$S(t) = \left(\sum_{m=1}^M A_m X_m(t) + A_b B(t) \right) e^{(\Delta\alpha + 2i\pi\Delta f)t} e^{i\Delta\theta} + WF \times W(t) + \xi(t), \quad (4)$$

where A_m and $X_m(t)$ are a scaling factor and a basis signal for the m -th metabolite, respectively; $\Delta\alpha$, Δf , and $\Delta\theta$ are a global damping factor, a global frequency shift, and a global phase shift, respectively; and M is the number of metabolites. A_b and $B(t)$ are a scaling factor and a signal for macromolecules, respectively; WF and $W(t)$ are a water amplification factor and exponential decay water components, respectively; and $\xi(t)$ is noise.

The Internet Brain Segmentation Repository was utilized to produce simulated MRSI data, which provides manually guided expert segmentation of white matter, gray matter, and CSF along with MR brain image data.

Each MRI volume has a size of $256 \times 256 \times 128$ voxels. The middle coronal slice of one subject was selected, and then the corresponding segmentation mask was down-sampled from the size of 256×256 voxels to the size of 64×64 voxels.

An MR spectroscopic FID was generated for each voxel in gray or white matter using Equation (4) and the publicly available metabolite basis set from the International Society for Magnetic Resonance in Medicine MRS

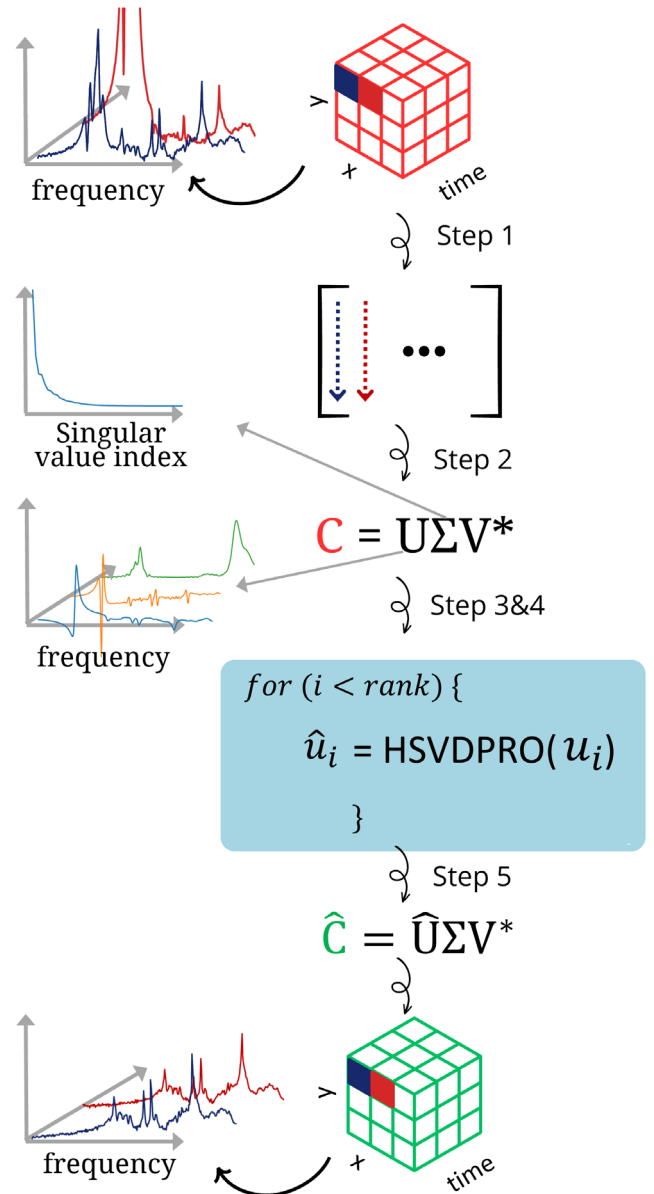


FIGURE 1 A schematic of the CSVD algorithm. The algorithm takes the measured MRSI signals (red cube) and processes them in the following steps: First, the signals are arranged in a Casorati matrix C by stacking the signals from each voxel in the columns. Then, the SVD of C is computed, and the rank of C is determined. Next, HLSVDPRO algorithm is applied to each of the first r columns of U (u_i) to remove residual water signals. Finally, the water-free MRSI signals (green cube) are reconstructed. CSVD, Casorati singular value decomposition; SVD, singular value decomposition.

study group's fitting challenge⁹ (19 metabolites signals and one macromolecule signal). Regional concentrations of metabolites and regional damping factors were determined randomly according to literature.¹⁰⁻¹²

According to Lin et al.,¹³ a set of residual water peaks (parameters are listed in Table 1) were added to signals, randomly and independently. Voxel-dependent frequency

TABLE 1 Parameters used to generate the residual water components.

Signal components	Center chemical shift (ppm)	Damping (s^{-1})	Phase ($^{\circ}$)	Amplitude distribution
$W_1(t)$	4.90	30	-45	$\mathcal{N}(30, 1)$
$W_2(t)$	4.82	25	30	$\mathcal{N}(80, 1)$
$W_3(t)$	4.70	40	20	$\mathcal{N}(180, 1)$
$W_4(t)$	4.62	20	-60	$\mathcal{N}(70, 1)$
$W_5(t)$	4.50	35	120	$\mathcal{N}(70, 1)$

Abbreviation: $W(t)$, exponential decay water components.

shifts induced by residual field inhomogeneity (ranges from -10 to 10 Hz) and Gaussian-distributed zeroth-order phase shifts ($\mathcal{N}(0, 5.7^{\circ})$) were added to the simulated data.

The signals were generated for a bandwidth of 4000 Hz and had 2048 points. Random complex Gaussian white noise was added to the signals, resulting in SNRs (time origin magnitude to noise standard deviation (SD)) of 196 ± 20 . Three sets of signals were simulated with WF of 5, 25, and 125.

2.2.2 | In vivo MRSI data

LTE-WS-MRSI dataset

Publicly accessible long TE brain MRSI data with water suppression (abbreviated as LTE-WS-MRSI and)¹³ from a healthy volunteer acquired on a Philips Achieva 3T scanner (Philips Healthcare, Best, The Netherlands; field of view = $230 \times 172.5 \text{ mm}^2$; slice matrix = 32×24 ; slice thickness = 15 mm; TE = 140 ms; TR = 3.2 s; sample bandwidth = 2000 Hz; sample points = 512; water suppression enhanced through T1 effect scheme for water suppression; 32-channel head coil).

2D-MRSI-sLASER dataset

Four healthy volunteers were scanned on a Siemens 3T MR system (Siemens Healthineer, Erlangen, Germany) equipped with a 64-channel head coil with supraventricular volume of interest positioning. A metabolite-cycled 2D-MRSI-semi-localization by adiabatic selective refocusing (sLASER) scheme (a weighted Cartesian k-space encoding and a 16×16 FOV grid with a resolution of $200 \times 160 \text{ mm}^2$) was utilized for acquisition of MRSI dataset (abbreviated as 2D-MRSI-sLASER). The MRSI-volume of interest had a resolution of $80 \times 60 \times 15 \text{ mm}^3$, and sequence timings were set at TR/TE 1600/35 ms. Spectra were recorded with a spectral width of 4 kHz and acquired with 4096 datapoints, totaling four weighted acquisitions in a 7-min scan time. The dataset will be publicly available in NIFTI-MRS format.¹⁴

2.2.3 | Functional MRS dataset

Our method demonstrates versatility because it can be effectively applied not only to MRSI data but also other large MRS datasets. To validate its efficacy, we thoroughly tested and evaluated our method specifically for functional proton MRS (fMRS) data. The fMRS is publicly accessible fMRS signals from 18 healthy participants acquired on a Philips Achieva 3T scanner¹⁵ (single-channel transmit-receive head coil, PRESS, TE/TR = 22/4000 ms, number of signal averages = 16, scan time = 22:4 min, anterior cingulate cortex, voxel size = $30 \times 25 \times 15 \text{ mm}^3$, second-order shimming, 16-step phase cycle with water suppression using the Excitation option—a Philips variant of chemical shift selective).

All scans were performed in accordance with the ethical review boards, and informed consent was obtained from all investigated subjects.

2.3 | Implementation details

We have developed the CSVD algorithm in Python programming language and released it as an open source pip-installable Python tool.¹⁶ We carried out simulations and data processing utilizing our in-house programs developed with Python on a PC with a 2.7-GHz Intel-Core-i7 processor (Intel Corp., Santa Clara, CA) and 16 GB of memory.

We used the HLSVDPRO method (a widely used HLSVDPRO-based method) for removing residual water signals in the fourth step of our proposed algorithm. We compared the proposed CSVD method with the HLSVDPRO method for removing residual water signals from the simulated and in vivo datasets.

We set the water removal range to 4.7 ± 0.5 ppm in HLSVDPRO and CSVD algorithms (step 4) to keep the peak at 3.9 ppm corresponding to the creatine visible.

We employed a peak integration approach³ that estimates metabolite signal intensities by calculating the area under the signal in the range of 1 to 4 ppm (in the

frequency domain) after the water removal. The consistency of metabolite signal intensities between CSVD and HLSVDPRO algorithms for the in vivo and the fMRS datasets was evaluated with Bland–Altman analysis. The residual spectra, GT-CSVD and GT-HLSVDPRO, were calculated by subtracting the water-free spectra, CSVD and HLSVDPRO, from the ground truth spectra.

We performed a Student *t*-test for pairwise group comparison. A *p*-value of <0.005 was considered statistically significant. To quantify the accuracy between the two methods, we calculated the coefficient of determination (R^2) and the symmetric mean absolute percentage error (SMAPE) as additional evaluation metrics.

We analyzed the effects of two key parameters, the rank and the model order, on the performance of CSVD on simulated data with WF of 25. This analysis sweeps over a range of ranks from 5 to 20, and model orders from 10 to 80. For each combination, the resulting accuracy and computation time are measured. Accuracy is quantified by the mean and SD of squared error between the reconstructed and original data.

3 | RESULTS

3.1 | Simulated dataset

Results of the three different water removal methods on the simulated datasets with WF of 5, 25, and 125 are shown in Figures 2 and 3.

Ranks $r = 9, 12,$ and 16 were selected for WF of 5, 25, and 125, respectively, by the optimal singular value hard thresholding algorithms, which were much lower than the original dimensionality of data ($32 \times 32 = 1024$).

For each water level WF and both CSVD and HLSVDPRO algorithm, a spectral mean squared error (MSE) was calculated using the ground truth spectra data.

Figure 2A shows the error map for CSVD and HLSVDPRO methods at three WF levels. The spectral MSEs produced by CSVD and HLSVDPRO across all voxels are 1.3 ± 0.45 and 2.6 ± 1.2 at WF = 5, 2.6 ± 0.97 and 2.3 ± 0.97 at WF = 25, and 2.1 ± 1 and 2.2 ± 1.5 WF = 125, respectively. The results show that at lower WF (=5) CSVD outperforms HLSVDPRO but becomes slightly less effective at higher WFs (=25 and 125).

Significant pairwise differences were not seen between MSE produced by CSVD and HLSVDPRO at all WFs ($p < 0.001$).

Figure 2B reveals tradeoffs between accuracy and computational efficiency as rank and order vary. An optimal balance (MSE = 1.28, SD = 0.62, and duration = 25.9 s) is achieved for a moderate rank (15) and high-model orders

(80), where accuracy remains high without excessive runtime.

Figure 3 illustrates original spectra from selected voxels. They are shown along with the corresponding water-free and residual spectra. The HLSVDPRO method is effective in removing the extremely strong water signal at higher WFs, resulting in minimal residual; however, it is inadequate in eliminating weak water signals (WF = 5). By contrast, the CSVD method exhibits superior performance by removing the water signal across all water levels, leading to negligible residual.

The residual spectra, GT-CSVD and GT-HLSVDPRO, were calculated by subtracting the water-free spectra, CSVD or HLSVDPRO, from the ground truth spectra.

Processing times for HLSVDPRO and CSVD methods for the simulated dataset were 201, 14, and 8 s, respectively. The model order (the number of components) was set to 60 and 30 for CSVD and HLSVDPRO algorithms, respectively.

3.2 | In vivo MRSI data

3.2.1 | LTE-WS-MRSI dataset

Figure 4 illustrates the results of CSVD and HLSVDPRO methods on the LTE-WS-MRSI dataset. Original spectra from selected voxels were shown along with the corresponding water-free spectra. Both the CSVD and the HLSVDPRO work well; however, the HLSVDPRO cannot adequately remove the residual water signal, resulting in baseline distortion.

Figure 5 shows plots of the Bland–Altman analysis of the agreement between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms. The mean difference line is at -12 , and the ± 1.96 SD of the difference lines are at -32 and 8 , respectively. This means that on average, CSVD gives lower metabolite signal intensities than the HLSVDPRO by 10 units (SMAPE = 2.36%). For the LTE-WS-MRSI dataset, significant pairwise differences were seen between metabolite signal intensity obtained from CSVD and HLSVDPRO algorithms ($p < 0.001$). The coefficients of determination (R^2) between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms is 0.5. It suggests a moderate-to-high agreement between CSVD and HLSVDPRO methods. Processing times for HLSVDPRO and CSVD methods for the LTE-WS-MRSI dataset were 271 and 13 s, respectively. A rank $r = 22$ was selected by the optimal singular value hard thresholding algorithm. The model order was set to 30 for CSVD and HLSVDPRO algorithms.

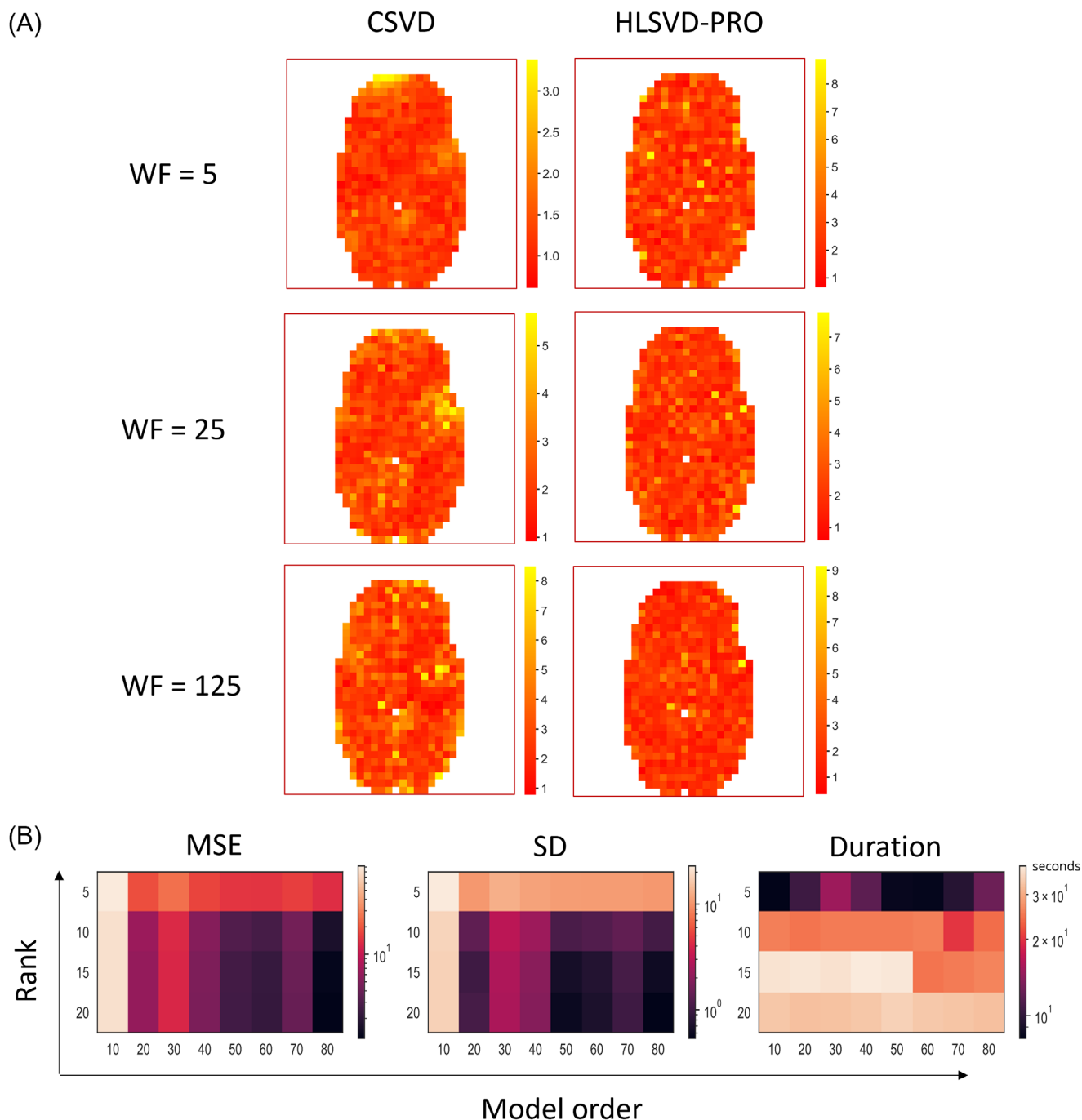


FIGURE 2 (A) The error map (spectral mean squared error) for CSVD and HLSVDPRO methods obtained from the simulated dataset at three WF levels. (B) Heatmap visualization of the effects of rank and model order on the performance of CSVD. The MSE, SD of error, and runtime are shown for ranks ranging from 5 to 20 and model orders from 10 to 80. Lower MSE and SD indicate more accurate and precise results. MSE, mean squared error; WF, water amplification factor.

3.2.2 | 2D-MRSI-sLASER

Figure 6 displays the spectra from chosen voxels of four subjects, both in their original form and after the removal of water; and Figure 7 depicts Bland–Altman analysis plots comparing the metabolite signal intensities derived from CSVD and HLSVDPRO algorithms.

For all four subjects, significant pairwise differences were not seen between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms in ($p < 0.001$). The R^2 between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms for subjects 1 to 4 are 0.92, 0.89, 0.97, and 0.94, respectively. The SMAPE between CSVD and HLSVDPRO algorithms for

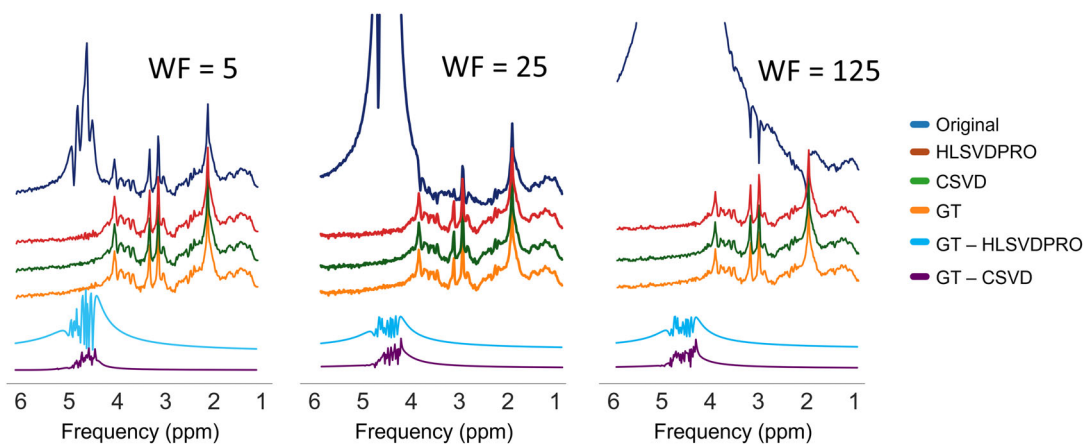
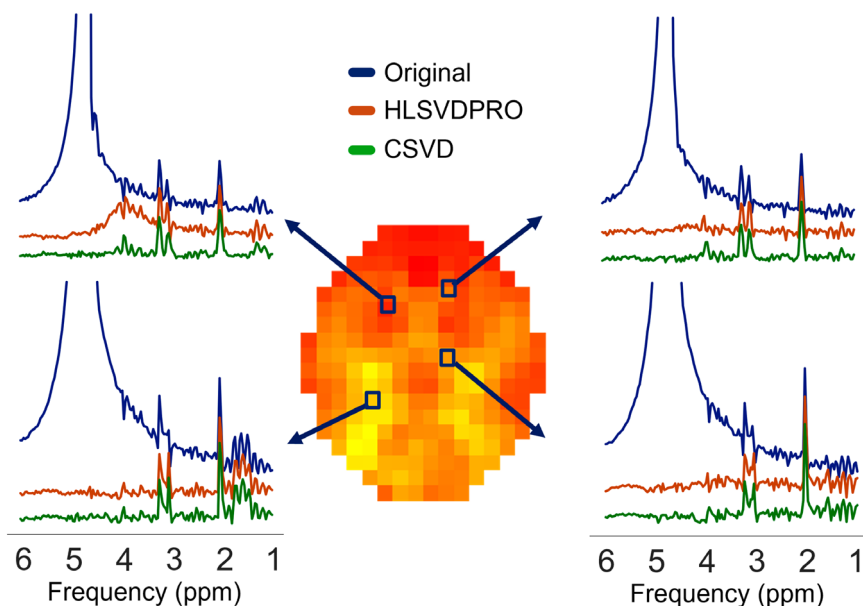


FIGURE 3 Comparison of CSVD and HLSVDPRO methods for the simulated dataset. Representative original spectra at three water levels (WF = 5, 25, and 125) from the selected voxels along with the corresponding water-free, ground truth, and residual spectra. The residual spectra, GT-CSVD and GT-HLSVDPRO, were calculated by subtracting the water-free spectra, CSVD and HLSVDPRO, from the ground truth spectra.

FIGURE 4 Comparison of HLSVDPRO and CSVD methods for the LTE-WS-MRSI dataset. Representative original spectra from the selected voxels along with the corresponding WF spectra. The displayed image visualizes the magnitude of the first FID point from the original dataset. This magnitude value corresponds to the integral of the spectral resonance peaks over the acquisition bandwidth at each voxel location.



subjects 1 to 4 are 2.06%, 1.66%, 2.36%, and 2.05%, respectively. These suggest a high agreement between CSVD and HLSVDPRO methods for all subjects.

The average processing times for CSVD and HLSVDPRO methods across all subjects were 141 and 685 s, respectively. A rank $r = 57$ was selected by the optimal singular value hard thresholding algorithm. The model order was set to 60 and 30 for CSVD and HLSVDPRO algorithms, respectively.

3.3 | fMRS

Results of the fMRS dataset using the CSVD and HLSVDPRO methods are shown in Figures 8 and 9. Figure 8

displays original spectra from specific voxels alongside their corresponding water-free spectra. Plots depicting the results of the Bland–Altman analysis comparing the metabolite signal intensities obtained from the CSVD and HLSVDPRO algorithms are displayed in Figure 9. The mean difference line is at around zero (SMAPE = 1.82%), which means that there is no systematic bias between the two methods, and 95% confidence interval lines are very close to zero as well (−0.0011 and 0.0013), which means that there is very little random variation between the two methods. This suggests that the two methods are very consistent and agree well with each other.

For fMRS, significant pairwise differences were not seen between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms ($p < 0.001$). The R^2

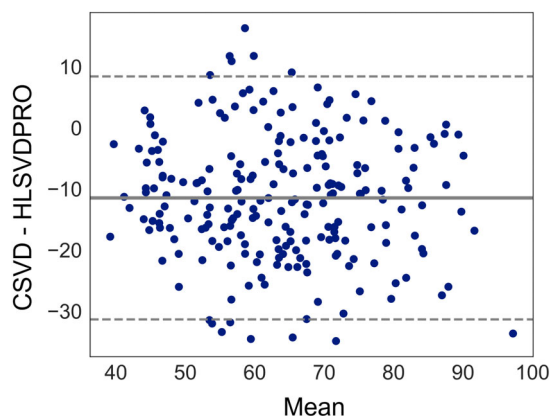


FIGURE 5 The plots of the Bland–Altman analysis of the agreement between metabolite signal intensities of the masked LTE-WS-MRSI dataset obtained from CSVD and HLSVDPRO algorithms. The y-axis displays the difference, whereas the x-axis represents the average of the methods. The limits of agreement are determined as mean \pm 1.96 SD of the difference and displayed on the plot as dashed lines. The solid line on the plot represents the average difference. LTE-WS-MRSI, long TE brain MRSI.

between metabolite signal intensities obtained from CSVD and HLSVDPRO algorithms is 0.81. It suggests a high agreement between CSVD and HLSVDPRO methods.

Processing times for HLSVDPRO and CSVD methods for the fMRS dataset were 68 and 7 s, respectively. A rank $r = 29$ was selected by the optimal singular value hard thresholding algorithm. The model order was set to 15 for CSVD and HLSVDPRO algorithms.

4 | DISCUSSION

The presented results demonstrate a comprehensive evaluation of our proposed (CSVD) and HLSVDPRO water removal methods with simulated and in vivo MRS and MRSI datasets. In the simulated dataset, CSVD outperforms HLSVDPRO at lower water fractions but becomes slightly less effective at higher water fractions. However, the pairwise differences between the MSE produced by CSVD and HLSVDPRO are not significant.

Overall, the processing time for CSVD is much faster than HLSVDPRO, indicating its practical usability for real-time analysis. The processing time of HLSVDPRO depends on the number of exponential components in the signal, which can vary depending on the experimental conditions and parameters.^{4,6} On the other hand, the processing time of CSVD depends on the rank of the Casorati matrix and the number of exponential components in step 4.

The SVD of an $n \times m$ matrix requires $\mathcal{O}(nm^2 + n^3)$ floating-point operations (FLOPs). Given a dataset with

2048 time points and 2048 voxels with a rank of 16 and forming a Hankel matrix with a size of 1024×1024 , our approach reduces the required FLOPs by $\sim 98\%$. However, more efficient algorithms for computing the SVD such as Lanczos⁴ may reduce the required FLOPs.

For the in vivo MRSI dataset (the LTE-WS-MRSI dataset), both CSVD and HLSVDPRO work well, but the HLSVDPRO did not adequately remove the residual water signal, resulting in baseline distortion. However, significant pairwise differences were not seen between the metabolite signal intensities. A possible explanation for this discrepancy is that HLSVDPRO overestimates metabolite signal intensities.

We tested our method against a fMRS dataset to show the capability of our method in the water removal of non-MRSI datasets from different subjects. Our method showed an excellent performance with a high agreement with the HLSVDPRO method.

For the in vivo MRSI dataset (the 2D-MRSI-sLASER dataset), CSVD and HLSVDPRO methods show a high agreement for all subjects. We also tried to accelerate our method in a way that we combine spectra from all four subjects to form a big matrix and applied our proposed method for water removal (Figure S1).

Apart from employing the CSVD method to eliminate water, it is also feasible to remove lipid signals concurrently by properly defining frequency ranges. The outcomes of simultaneous water and lipid removal within a range of 1.3 ± 0.6 ppm on long-TE MRSI data are presented in Figure S2.

Water suppression techniques have several disadvantages, such as increasing scan time and systematic errors in metabolite quantification.

Numerous prior studies have successfully shown the feasibility of in vivo non-water-suppressed MRSI, wherein both metabolite and water signals are concurrently measured.^{17,18}

This approach offers several advantages, such as having access to the complete water signal as a reference to correct lineshape distortion caused both by eddy currents induced in the metal surface of the magnet, correcting voxel-to-voxel frequency drifts, and employing an internal reference for metabolite quantification. Furthermore, it circumvents potential issues linked to water suppression pulses, including partial suppression of nearby resonances and magnetization transfer effects on the Cr signal. Despite its benefits, water unsuppressed experiments face sideband artifacts originating from the water signal and the difficulty of accurately quantifying the much smaller metabolite peaks against the significantly larger water signal. A software solution can be the extraction of the water signal without interfering with the metabolite content and introducing baseline distortions and artifacts.¹⁷

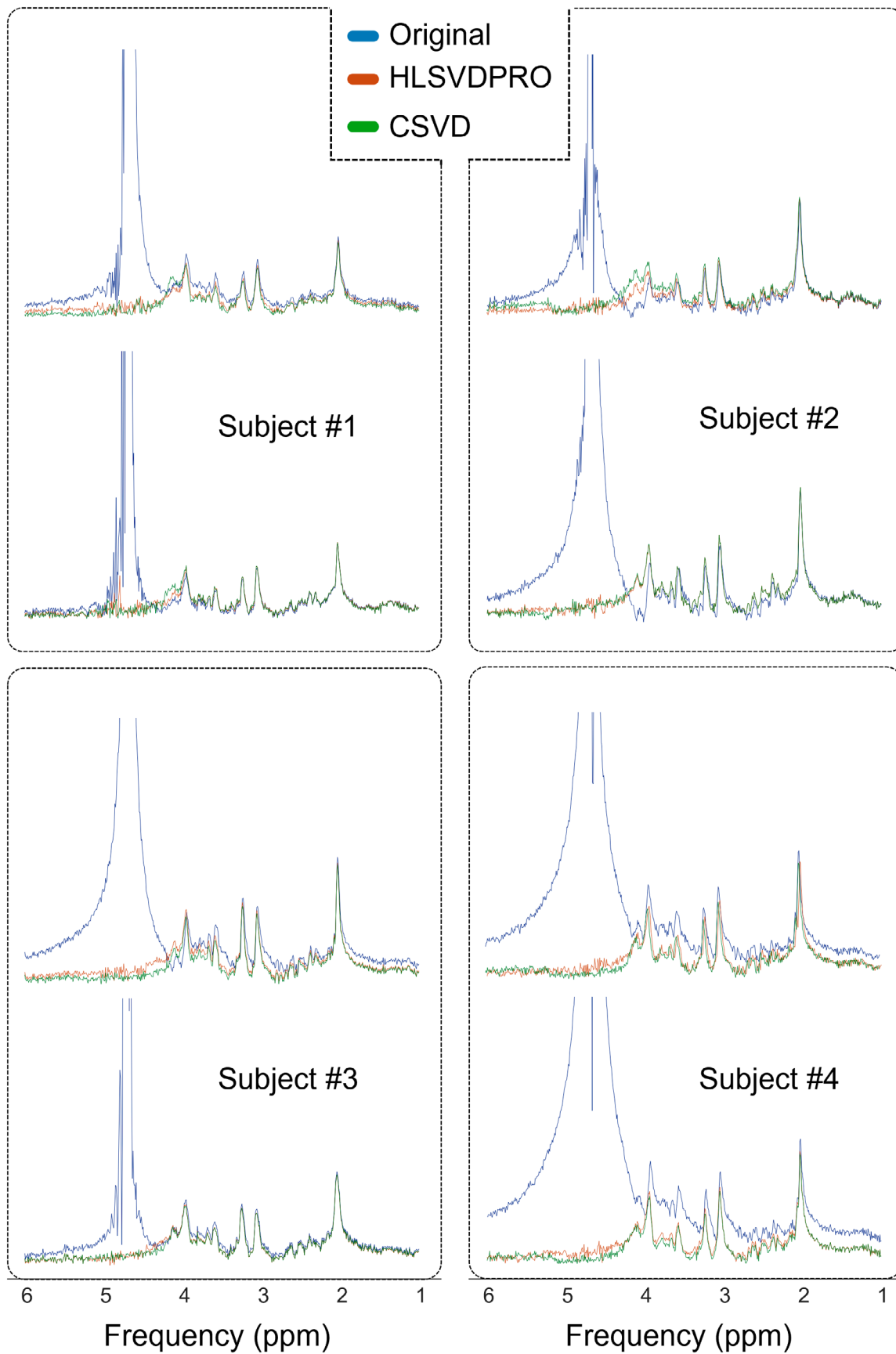


FIGURE 6 Comparison of HLSVDPRO and CSVD methods for the 2D-MRSI-sLASER dataset. Two representative original spectra from each of the four subjects along with the corresponding WF spectra.

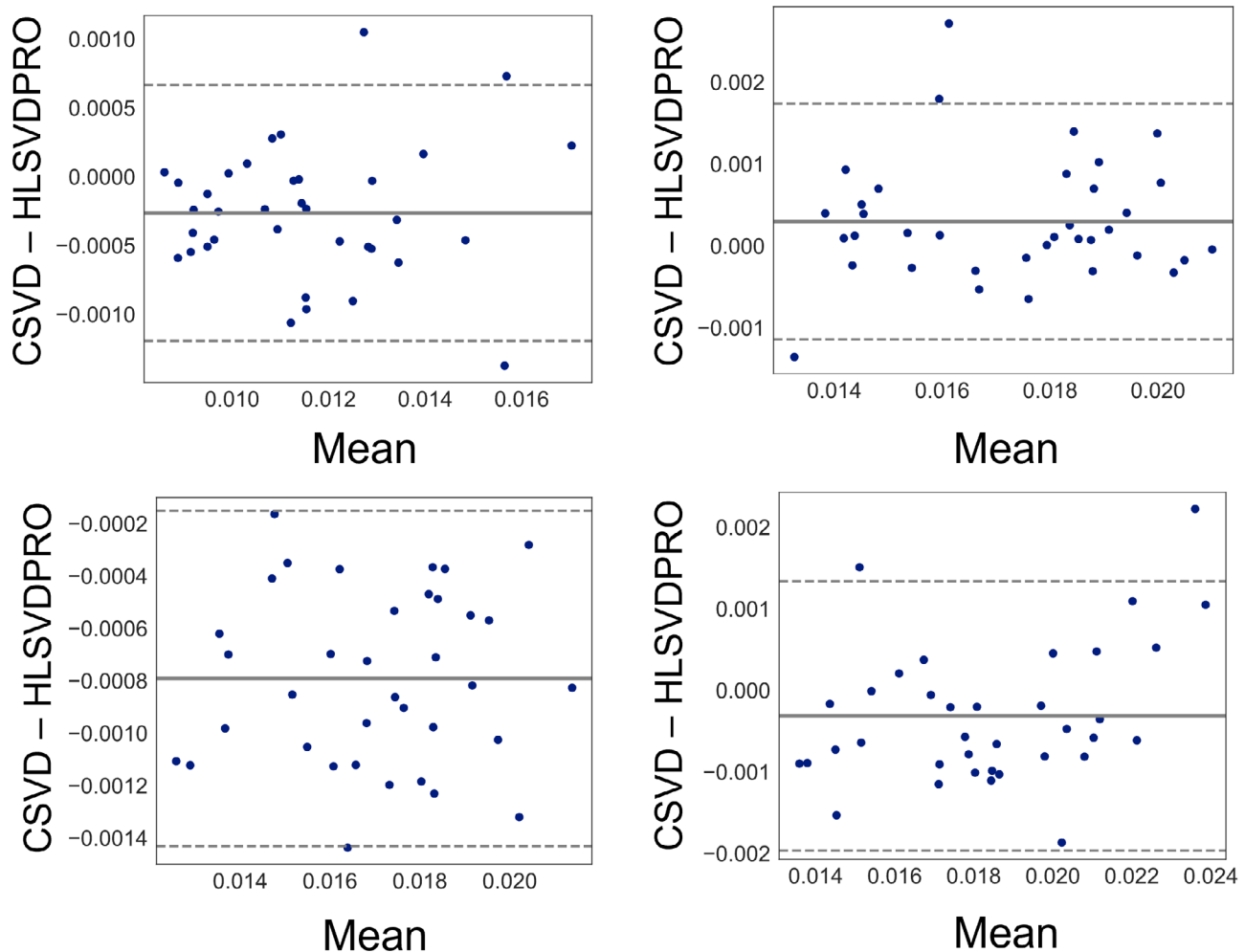


FIGURE 7 The plots of the Bland–Altman analysis of the agreement between metabolite signal intensities of the masked 2D-MRSI-sLASER dataset obtained from CSVD and HLSVDPRO algorithms. In each plot, the y-axis displays the difference, whereas the x-axis represents the average of methods. The limits of agreement are determined as mean \pm 1.96 SD of the difference and displayed on the plot as dashed lines. The solid line on each plot represents the average difference.

SVD-based algorithms were proposed for water signal extraction.¹⁷ Figure S3 shows the result of water removal for non-water-suppressed MRSI data using CSVD method. Our proposed method showed acceptable water removal for the non-suppressed water data in a shorter time.

4.1 | Limitations

Our study has some limitations that should be acknowledged. First, we used a limited set of data to evaluate the performance of CSVD and HLSVDPRO methods. Although we tried to mimic realistic scenarios by varying WF levels and adding noise, our data may not fully capture the complexity and variability of different real-life spectra. Therefore, further studies are needed to validate our findings using experimental data from different sources and scanners.

Second, we only compared two methods for water removal in this study. There may be other methods that can achieve better results or have different advantages and disadvantages.^{19,20} Future research should explore other alternatives or combinations of methods for the water removal problem, for example, utilizing the incorporation of physical knowledge in deep neural networks^{21–23}—or a recent alternative proposed by Lin et al., which implements an L2-based water removal method that can accelerate the water removal procedure by a factor of 50, and requires prior knowledge such as the range of damping factors, phases, and frequencies about water components.¹³

The third point to consider is that the HLSVDPRO-based methods assume the MRS signals in the time domain are limited to a few exponential decaying oscillators. If the actual residual water signal deviates significantly from a Lorentzian lineshape, the results produced by HLSVDPRO-based methods may not be satisfactory.

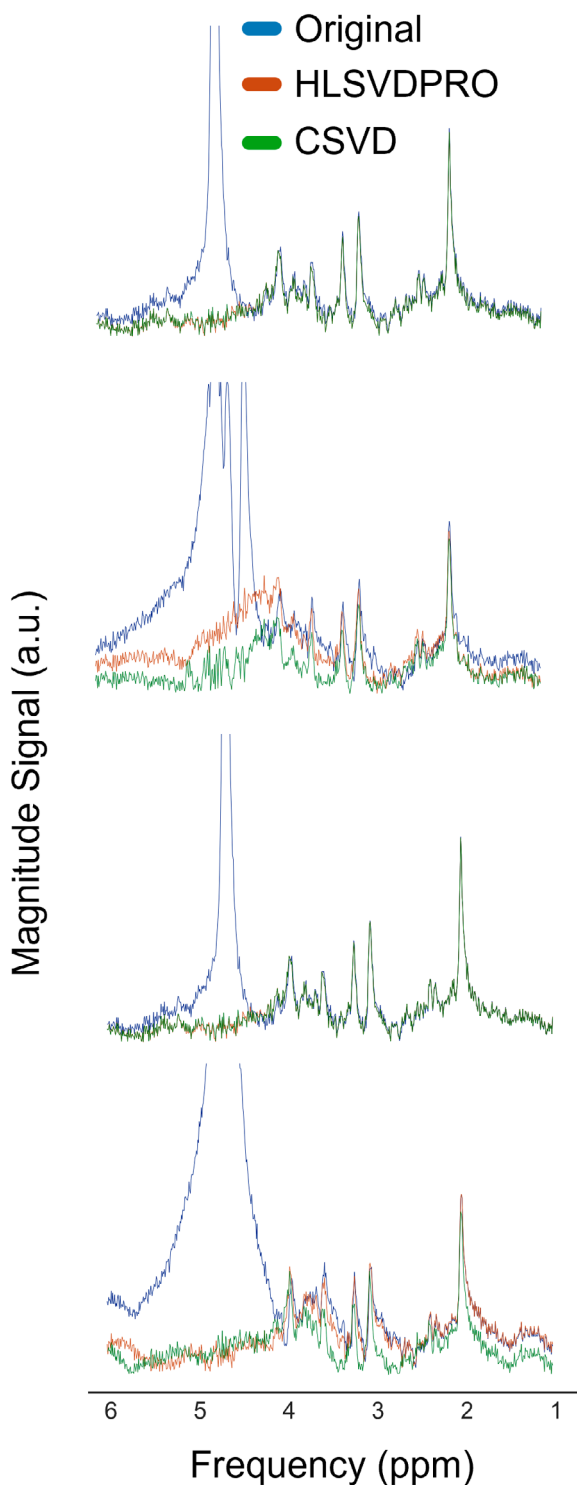


FIGURE 8 Comparison of HLSVDPRO and CSVD methods for the fMRS dataset. Representative original spectra from four subjects along with the corresponding WF spectra.

Step 4 of our proposed algorithm is needed to afford more flexibility in future work.

Finally, the choice of rank is a critical consideration when using low-rank approximation techniques such as the proposed CSVD method. Selecting the appropriate

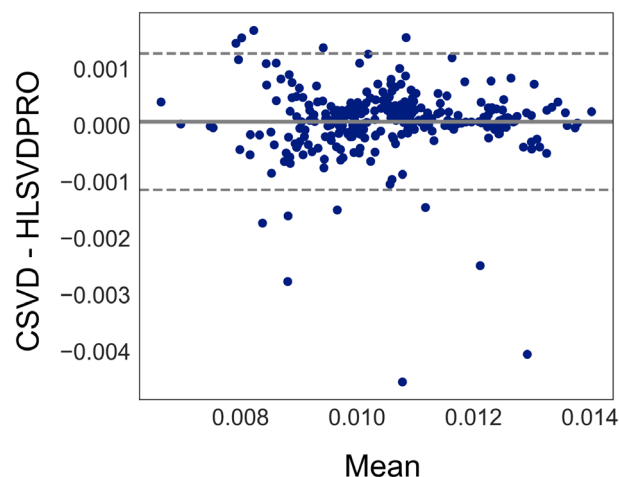


FIGURE 9 The plots of the Bland–Altman analysis of the agreement between metabolite signal intensities of the fMRS dataset obtained from CSVD and HLSVDPRO algorithms. The y-axis displays the difference, whereas the x-axis represents the average of methods. The limits of agreement are determined as mean \pm 1.96 SD of the difference and displayed on the plot as dashed lines. The solid line on each plot represents the average difference.

rank has a major impact on balancing approximation accuracy and model complexity. If the rank is too low, important information may be lost. On the other hand, excessively high ranks increase computation time and may cause overfitting to noise. In this work, we utilized the optimal hard thresholding algorithm⁸ to automatically estimate the rank of the Casorati matrix. This approach essentially thresholds the singular values at an optimal point to minimize the expected reconstruction error. The optimal hard thresholding technique is well studied and provides a near-optimal estimate of the matrix rank in an efficient and easy-to-implement manner, without requiring manual parameter tuning. However, other automated parameter selection methods could also be considered for determining the rank in the proposed low-rank approximation. For example, approaches based on Stein's unbiased risk estimate²⁴ and the Marchenko-Pastur²⁵ distribution have been extensively utilized in the MR field for selecting regularization parameters and estimating the number of components.^{25,26} Although we include these approaches in the CSVD python package,¹⁶ investigating these and other data-driven selection procedures represents an interesting area for further improving the rank estimation in the current work. The optimal rank depends on the intrinsic dimensionality and complexity of the dataset. The addition of noise and modeling errors means a higher rank is needed to explain the same amount of variance compared to the noiseless case. Thus, the optimal rank must balance retaining useful information against overfitting. As data complexity increases, for example, due

to greater variability in water suppression or metabolite mixing, higher rank representations become necessary to maintain modeling accuracy. In our simulated experiments, the estimated rank increased at higher water fractions, although the number of underlying metabolite signals was unchanged. This implies the data became more complex with increasing water. The additional water signal likely introduced extra variance needing additional components to model.

5 | CONCLUSIONS

The presented results suggest that CSVD is a promising alternative to existing water removal methods due to its low processing time and good performance in removing water signals across different water fractions. However, HLSVDPRO methods may be preferable in specific cases where extremely strong water signals need to be removed, although they may result in larger residuals. The evaluation of different water removal methods with different datasets provides valuable insights into the strengths and weaknesses of each method, helping researchers and clinicians select the most suitable method for their specific application.

ACKNOWLEDGMENTS


This work is part of the project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (no. 813120; INSPiRE-MED); institutional support from the Czech Academy of Sciences, Institute of Scientific Instruments (ISI) (grant RVO:68081731); and was provided with staff and equipment from the ISI Core Facility Magnetic Resonance (ISI-MR) facility of the Czech-BioImaging infrastructure (grants CZ.02.1.01/0.0/0.0/18_046/0016045, LM2018129, and LM2023050 of The Ministry of Education, Youth and Sports of the Czech Republic [MEYS CR]).

DATA AVAILABILITY STATEMENT

The source code is freely available at (<https://github.com/amirshamaei/CSVD>). For questions, please contact the corresponding author. The proposed algorithm encompasses a pip-installable tool available at (<https://pypi.org/project/CSVD/>). The script for generating the simulated data is available at (https://github.com/amirshamaei/CSVD_paper). LTE-WS-MRSI dataset is openly available at (https://github.com/mrshub/Wat_Lip_Removal_L2). 2D-MRSI-sLASER dataset will be available on MRSHub (<https://mrshub.org/>). Functional-MRS dataset is openly available at (https://www.nitrc.org/projects/fmrs_2020/).

ORCID

Amirmohammad Shamaei  <https://orcid.org/0000-0001-8342-3284>

Zenon Starcuk Jr  <https://orcid.org/0000-0002-1218-0585>

TWITTER

Amirmohammad Shamaei  AmirLand6

REFERENCES

- Kreis R, Boer V, Choi IY, et al. Terminology and concepts for the characterization of in vivo MR spectroscopy methods and MR spectra: background and experts' consensus recommendations. *NMR Biomed.* 2020;34:e4347.
- Tkác I, Dreher W, Kumaragamage C, et al. Water and lipid suppression techniques for advanced ¹H MRS and MRSI of the human brain: experts' consensus recommendations. *NMR Biomed.* 2021;34:1-25.
- Near J, Harris AD, Juchem C, et al. Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR Biomed.* 2019;2020:1-23.
- Laudadio T, Mastronardi N, Vanhamme L, Van Hecke P, Van Huffel S. Improved Lanczos algorithms for blackbox MRS data quantitation. *J Magn Reson.* 2002;157:292-297.
- Cabanes E, Le FY, Simond G, Cozzone PJ. Optimization of residual water signal removal by HLSVD on simulated short echo time proton MR spectra of the human brain. *J Magn Reson Imag.* 2001;125:116-125.
- Vanhamme L, Fierro RD, Van Huffel S, De Beer R. Fast removal of residual water in proton spectra. *J Magn Reson.* 1998;132:197-203.
- Nguyen HM, Peng X, Do MN, Liang ZP. Denoising MR spectroscopic imaging data with low-rank approximations. *IEEE Trans Biomed Eng.* 2013;60:78-89.
- Gavish M, Donoho DL. The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Trans Inf Theory.* 2013;60:5040-5053.
- Marjańska M, Deelchand DK, Kreis R, 2016 ISMRM MRS Study Group Fitting Challenge Team. Results and interpretation of a fitting challenge for MR spectroscopy set up by the MRS study group of ISMRM. *Magn Reson Med.* 2022;87:11-32.
- Michaelis T, Merboldt KD, Bruhn H, Hancic W, Frahm J. Absolute concentrations of metabolites in the adult human brain in vivo: quantification of localized proton MR spectra. *Radiology.* 1993;187:219-227.
- Tra F, Block W, Lamerichs R, Schild HH. ¹H metabolite relaxation times at 3.0 Tesla: measurements of T₁ and T₂ values in normal brain and determination of regional differences in transverse relaxation. *J Magn Reson Imaging.* 2004;545:537-545.
- Wang Y, Li SJ. Differentiation of metabolic concentrations between gray matter and white matter of human brain by in vivo ¹H magnetic resonance spectroscopy. *Magn Reson Med.* 1998;39:28-33.
- Lin L, Považan M, Berrington A, Chen Z, Barker PB. Water removal in MR spectroscopic imaging with L2 regularization. *Magn Reson Med.* 2019;82:1278-1287.

14. Clarke WT, Bell TK, Emir UE, et al. NifTI-MRS: a standard data format for magnetic resonance spectroscopy. *Magn Reson Med.* 2022;88:2358-2370.
15. Archibald J, MacMillan EL, Graf C, Kozlowski P, Laule C, Kramer JLK. Metabolite activity in the anterior cingulate cortex during a painful stimulus using functional MRS. *Sci Rep.* 2020;10:19218.
16. PyPI. CSVD. Available at: <https://pypi.org/project/CSVD/> Accessed March 3, 2023.
17. Dong Z. Proton MRS and MRSI of the brain without water suppression. *Prog Nucl Magn Reson Spectrosc.* 2014;86-87:65-79.
18. Emir UE, Burns B, Chiew M, Jezzard P, Thomas MA. Non-water-suppressed short-echo-time magnetic resonance spectroscopic imaging using a concentric ring k-space trajectory. *NMR Biomed.* 2017;30:e3714.
19. Nagaraja BH, Debals O, Sima DM, Himmelreich U, De Lathauwer L, Van Huffel S. Tensor-based method for residual water suppression in 1 H magnetic resonance spectroscopic imaging. *IEEE Trans Biomed Eng.* 2019;66:584-594.
20. Ma C, Lam F, Johnson CL, Liang ZP. Removal of nuisance signals from limited and sparse 1 H MRSI data using a union-of-subspaces model. *Magn Reson Med.* 2016;75:488-497.
21. Gurbani SS, Sheriff S, Maudsley AA, Shim H, Cooper LAD. Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting. *Magn Reson Med.* 2019;81:3346-3357.
22. Shamaei A, Starcukova J, Pavlova I, Starcuk Z. Model-informed unsupervised deep learning approaches to frequency and phase correction of MRS signals. *Magn Reson Med.* 2023;89:1221-1236.
23. Shamaei AM, Starcukova J, Starcuk Z. Physics-informed deep learning approach to quantification of human brain metabolites from magnetic resonance spectroscopy data. *Comput Biol Med.* 2023;158:106837. doi:10.1016/j.combiomed.2023
24. Candes EJ, Sing-Long CA, Trzasko JD. Unbiased risk estimates for singular value thresholding and spectral estimators. *IEEE Trans Signal Proc.* 2013;61:4643-4657.
25. Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. *Neuroimage.* 2016;142:394-406.
26. Clarke WT, Chiew M. Uncertainty in denoising of MRSI using low-rank methods. *Magn Reson Med.* 2022;87:574-588.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

FIGURE S1. Comparison of HLSVDPRO and CSVD methods for the combined spectra from all four subjects of 2D-MRSI-sLASER dataset.

FIGURE S2. Results of simultaneous water and lipid removal within a range of 1.3 ± 0.6 ppm for the LTE-WS-MRSI dataset.

FIGURE S3. Results of water removal for the water unsuppressed experiment. Representative original spectra from the selected voxels along with the corresponding water-free spectra.

How to cite this article: Shamaei A, Starcukova J, Rizzo R, Starcuk Z. Water removal in MR spectroscopic imaging with Casorati singular value decomposition. *Magn Reson Med.* 2024;1-13. doi: 10.1002/mrm.29959