Uncertainty in denoising of MRSI using low-rank methods

William T. Clarke, **Mark Chiew**

Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

Abstract

Purpose—Low-rank denoising of MRSI data results in an apparent increase in spectral SNR. However, it is not clear if this translates to a lower uncertainty in metabolite concentrations after spectroscopic fitting. Estimation of the true uncertainty after denoising is desirable for downstream analysis in spectroscopy. In this work, the uncertainty reduction from low-rank denoising methods based on spatiotemporal separability and linear predictability in MRSI are assessed. A new method for estimating metabolite concentration uncertainty after denoising is proposed. Automatic rank threshold selection methods are also assessed in simulated low SNR regimes.

Methods—Assessment of denoising methods is conducted using Monte Carlo simulation of proton MRSI data and by reproducibility of repeated in vivo acquisitions in 5 subjects.

Results—In simulated and in vivo data, spatiotemporal based denoising is shown to reduce the concentration uncertainty, but linear prediction denoising increases uncertainty. Uncertainty estimates provided by fitting algorithms after denoising consistently underestimate actual metabolite uncertainty. However, the proposed uncertainty estimation, based on an analytical expression for entry-wise variance after denoising, is more accurate. It is also shown automated rank threshold selection using Marchenko-Pastur distribution can bias the data in low SNR conditions. An alternative soft-thresholding function is proposed.

Conclusion—Low-rank denoising methods based on spatiotemporal separability do reduce uncertainty in MRS(I) data. However, thorough assessment is needed as assessment by SNR measured from residual baseline noise is insufficient given the presence of non-uniform variance. It is also important to select the right rank thresholding method in low SNR cases.

Keywords

denoising; low rank; MRS; MRSI; spectroscopy

Correspondence to: William T. Clarke.

Correspondence: William T. Clarke, Wellcome Centre for Integrative Neuroimaging, FMRIB, University of Oxford, Level 0, John Radcliffe Hospital, Oxford, OX3 9DU, UK. william.clarke@ndcn.ox.ac.uk.

Twitter
William T. Clarke @wt_clarke

1 Introduction

Increasing the SNR of MRSI allows for faster acquisitions, more reliable quantification, or higher resolution acquisitions. One way SNR can be increased is by reducing the noise variance through computational post-processing (ie, by "denoising"). Low-rank denoising methods achieve this in spectroscopic imaging data either by exploiting the linear predictability, or the spatiotemporal separability of the spectroscopic data, or both.[1] Low-rank denoising is a data-driven technique and does not incorporate prior knowledge or physical models of the data. These methods have also recently been applied to MR imaging techniques that use an additional dimension of encoding, such diffusion encoding direction in diffusion-weighted MRI,[2,3] and time in functional MRI.[4] Low-rank models have also been applied directly in the reconstruction of fast MRSI acquisitions.[5,6] The application of low-rank models to these techniques aims to exploit signal correlations across these encoding dimensions.

High levels of apparent denoising are consistently achieved by denoising algorithms in MRS with additional encoding dimensions,[7,8] or MRSI.[1,9,10] However, it is not clear whether there is an overall reduction in uncertainty of final dynamic model parameters,[11] or metabolite concentrations (the typical output of MRSI). Denoising will lower the apparent noise in any given spectrum, but can introduce systematic model-based errors affecting the bias and variance of the output, ultimately resulting in lower reproducibility and higher mean squared error. Furthermore, because spectroscopic signals are typically converted to metabolite concentrations by fitting of an explicit spectroscopic model to the data,[12] it is not clear whether denoising before fitting is statistically advantageous. Finally, uncertainty estimation of the fitted metabolite concentrations by fitting packages typically assumes a uniform spectral (or time-domain) noise profile. Usually a frequency independent noise variance is estimated from a signal-free region or noise pre-scans. After low-rank denoising, the residual noise cannot be assumed to be independent and identically distributed (complex) Gaussian noise. Therefore, uncertainty estimates cannot be trusted without validation. For the same reason, SNR is an inadequate metric for the evaluation of denoising algorithms.

In this work, we assess actual uncertainty reduction achieved by low-rank algorithms based on linear predictability (LP) of the time domain signal,[13] spatiotemporal (ST) separability,[14] and a combination of the 2 using the low rank approximations method (LORA).[1] We do this using Monte Carlo (MC) simulations of a toy problem, in simulated ^{1}H-MRSI data, and finally by assessing reproducibility in vivo ^{1}H-MRSI of the human brain. We additionally assess a new variance propagation method for data truncated by singular value decomposition for accurate estimation of the fitting uncertainties in denoised MRSI data.[15,16]

In MRSI, multiple metabolite signals are present in each voxel; the amplitudes of each signal are expected to vary across space, primarily driven by changes in metabolite concentration. Where there is pathologically driven change, metabolite concentrations can vary dramatically across short distances. To maximize noise variance reduction, low-rank methods aim to explain the data using as few rank components as possible. However, using few components may not accurately capture the full variation of rapidly changing metabolite

concentrations. Automated parameter selection methods such as those based on Stein's unbiased risk estimate (SURE) minimize the mean squared error (MSE) of the denoised data,[17,18] whereas Marchenko-Pastur (MP)-based methods aim to separate noise from signal by estimating which components arise from the pure noise matrix.[2] Minimization of spectral data MSE is not equivalent to minimization of fitted concentration MSE. In very noisy data, or in data where the target signal variance is on the order of the noise variance, methods that minimize data MSE are unlikely to preserve true signal variation.

Therefore, in this work, we first assess the effect of different rank estimation methods and a different thresholding operation in ST denoising of simulated data.

2 Theory

2.1 Low-rank denoising methods

In this work, truncation to a fixed rank is achieved using the singular value decomposition. Given an observation model

$$M_{ij} = M_{ij}^0 + E_{ij},$$

(1)

the Eckart-Young-Minsky theorem states the best rank-r approximation to M_{ij} can be derived from a truncated singular value decomposition retaining only the r highest singular values,

$$\widehat{M} = \hat{U}\,\hat{\Sigma}\,\hat{V}^{\dagger}.$$

(2)

$\hat{U} \in \mathbb{R}^{i \times r}$ and $\hat{V} \in \mathbb{R}^{j \times r}$ contain the truncated left and right singular vectors of $\widehat{M}$, $\hat{\Sigma}$ is a truncated diagonal matrix of singular values, and $\dagger$ is the Hermitian (conjugate) transpose. The noise term in Equation (1), E_{ij}, may either be zero-mean independent and identically distributed (i.i.d.) complex Gaussian (for ST denoising), or otherwise structured (for LP denoising).

2.2 Linear predictability denoising

LP denoising aims to exploit the low-rankness of a Hankel matrix formed from the single-voxel time domain signal.[13] This low-rankness arises from the sparsity of the equivalent spectral information.

This method is applied voxel-wise to the MRSI data, with no data shared across voxels. A Hankel matrix is formed from the time domain data of a single voxel *s*,

$$H = \begin{bmatrix} s[1] & \cdots & s[K] \\ s[2] & \cdots & s[K+1] \\ \cdots & \cdots & \ldots \\ s[M-K+1] & \cdots & s[M] \end{bmatrix}.$$

The denoised Hankel matrix $\bar{H}$ is formed by minimizing

$$\overline{H} = \text{argmin} \parallel H - \widehat{H} \parallel_F \,,$$
$$\text{rank}(\widehat{H}) = R$$

(3)

where $\parallel \bullet \parallel_F$ is the Frobenius norm. The denoised time-domain data is then reformed from the first row and last column of $\overline{H}$, effectively "softening" the low-rank filtering by not enforcing the Hankel structure in Equation (3).[1]

2.3 Spatiotemporal denoising

ST denoising exploits the partial separability of spatiotemporal modes because of correlated spectral information across space. For example, as in eq. 2 in Nguyen et al[1] the signal can be seen as L^{th}-order separable between the spatial basis $c_l(r)$ and the spectral basis $\psi_1(f)$

$$\rho(\boldsymbol{r}, f) = \sum_{l=1}^{L} c_l(\boldsymbol{r})\Psi_l(f)\,.$$

(4)

Equation (4) may equivalently be written in terms of basis in the temporal domain

$$s(\boldsymbol{r}, t) = \sum_{l=1}^{L} c_l(\boldsymbol{r})\widehat{\Psi}_l(t)\,.$$

To perform ST denoising a Casorati matrix is formed

$$C = \begin{bmatrix} s(t_1, r_1) & s(t_1, r_2) & \cdots & s(t_1, r_M) \\ s(t_2, r_1) & s(t_2, r_2) & \cdots & s(t_2, r_M) \\ \ldots & \ldots & \ldots & \ldots \\ s(t_N, r_1) & s(t_N, r_2) & \cdots & s(t_N, r_M) \end{bmatrix},$$

and the denoised matrix $\overline{C}$ is formed by minimizing

$$\overline{C} = \text{argmin} \parallel C - \widehat{C} \parallel_F \,.$$
$$\text{rank}(\widehat{c}) = R$$

(5)

2.4 LORA

The LORA method combines ST and LP denoising sequentially, applying ST to all voxels simultaneously before applying LP denoising voxel-wise.

2.5 Patch based methods (ST and LORA)

ST and LORA can be applied globally to all voxels in a data set, or locally in an overlapping patch-wise manner. In the local method, ST denoising is applied within each patch, and the signals from voxels belonging to multiple (overlapping) patches are averaged. Both methods may be applied to a restricted range of voxels identified by a mask. The patch is

characterized by a 3D patch size and a stride parameter, which dictates the amount of patch overlap.

2.6 Automatic rank selection methods: SURE and MP

Two methods for automatically estimating the rank threshold from the noisy data are assessed in this work. In addition, a different thresholding operation, "soft" or singular value thresholding (SVT), is introduced for 1 framework.

2.7 "Soft thresholding"

SVT or soft thresholding linearly shrinks the singular values above a threshold otherwise setting values below the threshold to 0. For example, the SVT estimated denoised matrix is defined as in eq. 2 in Candès et al.[17]

$$SVT_\lambda(\boldsymbol{M}) = \sum_{i=1}^{min(M, N)} (\sigma_i - \lambda)_+ \boldsymbol{u}_i \boldsymbol{v}_i^\dagger,$$ (6)

where $x_+ = \max(x, 0)$, λ is the threshold, and σ_i is the ith singular value.

Singular value hard thresholding (SVHT) is implemented for all other cases in this work and may be written,

$$SVHT_\lambda(\boldsymbol{M}) = \sum_{i=1}^{min(M, N)} H\big(\sigma_i - \lambda\big)\sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\dagger,$$ (7)

where H is the Heaviside function, or equivalently as a function of rank threshold R

$$SVHT_R(M) = \sum_{i=1}^{R} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\dagger.$$ (8)

2.8 SURE

For the model in Equation (1), SURE constructs an unbiased estimate of the risk (MSE) without requiring knowledge of the ground truth. For example, for the SVT function

$$MSE(\lambda) = \mathbb{E}\left\| \boldsymbol{M}_0 - SVT_\lambda(\boldsymbol{M})\right\|_F^2.$$ (9)

Using SURE, it is therefore, straightforward to find the singular value threshold λ, which minimizes the MSE of the denoised data. SURE for the (soft) SVT operation for complex data is given by eq. 7 in Candès et al.[17] The equivalent expression for a hard thresholding (SVHT) is given by eqs. 2 and 3 in Ulfarsson and Solo.[18] The equations as implemented in this work are provided in the Supporting Information.

2.9 MP

The second method assessed uses the upper limit of the MP distribution, which models the distribution of singular values of random matrices, as a singular value threshold. Introduced

to MRI denoising problems by Veraart et al,[2] this method truncates to 0 all eigenvalues below a threshold given by

$$\lambda_+ = \sigma^2(1 + \sqrt{\gamma})^2,$$

(10)

with $\gamma = M/N$ (the ratio of the dimensions of matrix M) and σ^2 the variance of the i.i.d. noise of the observed noisy data.

2.10 Uncertainty of low-rank denoised data

According to Chen et al,[15] under moderate-to-high SNR conditions, the variance of an element $\hat{v}_{ij}$ of "denoised" matrix $\widehat{M}$ is

$$\hat{v}_{ij} = \sigma^2\left(\left\|\widehat{U}_{i,\,\cdot}\right\|_2^2 + \left\|\widehat{V}_{j,\,\cdot}\right\|_2^2\right),$$

(11)

where σ^2 is the variance of the i.i.d. noise of the observed data, and $\hat{U}_{i\cdot}$ and $\hat{V}_{j\cdot}$ are the ith and j^{th} row of the left and right truncated singular vectors, respectively.[15] One way to interpret this equation is that because both U and V have normalized columns, and the low-rank constraint results in fewer columns than rows, then looking at the L2-norm of the entries of any given row of U and V will typically result in values <1. This captures the expected reduction in variance after rank truncation, although it is clear that the resulting variance estimates are signal dependent, which is a consequence of the non-linear rank thresholding operation. In contrast, conventional linear least squares estimators have signal-independent output variance.

Song et al[16] extended the above framework to incorporate patch-based low-rank denoising methods. The variance of elements having undergone patch denoising and averaging is given by eq. 20 in Song et al.[16] This extension of Chen et al's[15] method incorporates calculation of variance cross terms in the presence of mutually shared information between patches.

2.11 Uncertainty of metabolite concentrations from denoised data

The variance of the denoised data may be non-uniform in the time domain, and there may be significant covariance between the denoised time domain data points. When this is the case the conventional estimation of metabolite concentration uncertainty by using a signal free region to estimate noise variance is insufficient. However, propagation of non-uniform and correlated variance through the non-linear fitting process analytically is difficult. Therefore, in this work, a bootstrapping approach is proposed. Repeated fits are made of each voxel's data perturbed with complex correlated noise, created using the estimated variance and covariance of the denoised data. The off-diagonal elements of the covariance of the time domain data are approximated as the inner product of the corresponding rows of the left truncated singular vectors scaled by the input variance, $\sigma^2\widehat{U}_{i,\,\cdot}\,\widehat{U}_{i',\,\cdot}^{\dagger}$. The full covariance matrix is estimated by combination with the element-wise variance (Equation 11)

$$cov(i, i')_j = \begin{cases} \sigma^2 \left(\left\| \widehat{U}_{i, \cdot} \right\|_2^2 + \left\| \widehat{V}_{j, \cdot} \right\|_2^2 \right), i = i' \\ \\ \sigma^2 \widehat{U}_{i, \cdot} \cdot \widehat{U}_{i', \cdot}^{\dagger}, i \neq i' \end{cases},$$ (12)

where $cov(i, i')_j$ is the covariance of time point i with time point i' in voxel j, † denotes the Hermitian (conjugate) transpose operation, and σ^2 is the variance of the i.i.d. noise of the observed noisy data. The covariance of patch-based methods is estimated as the mean of the covariance of all overlapping patches. A qualitative assessment of this approximation for spectral data is made in the Supporting Information section "Covariance approximation."

3 Methods

All fitting in this work was carried out using version 1.1.2 of FSL-MRS. [19] Denoising was carried out using version 0.0.2 of the Python "mrs-denoising-tools" package. Data and analysis code for this paper are available online; please see the data availability statement.

3.1 In vivo ^{1}H MRSI data

In vivo ^{1}H-MRSI data was acquired from the brains of 5 healthy subjects (3 female, 26.6 ± 3.2 years old, 69.8 ± 7.2 kg) at 3T (Magnetom Prisma, Siemens Healthineers, Germany) using a previously reported density-weighted concentric ring (CONCEPT) MRSI sequence with semi-LASER localization.[20,21] Reconstructed data formed a $48 \times 48 \times 1$ grid at a spatial resolution of $5 \times 5 \times 15$ mm^3. Each subject was scanned with the CONCEPT sequence 10 times sequentially. Each separate acquisition was identical, baring a frequency adjustment between acquisitions, and took 4:30 min to acquire.

Reconstruction and preprocessing of the data used in-house custom MATLAB (The MathWorks, Natick, MA) scripts designed specifically for the density-weighted sequence. These steps were as follows:

1. Loading and reordering of the non-Cartesian k-space data.

2. Regridding with the adjoint 2D non-uniform fast Fourier transform (NUFFT).[22]

3. Coil combination using the wSVD algorithm with weights calculated from the water signal.[23]

4. Processing of metabolite-cycled acquisitions to form water-suppressed and unsuppressed data.

5. Phase and frequency correction across the 10 averages using cross-correlation.[24]

6. Frequency correction across voxels using B_0 shifts measured from the water reference data.

7. A combined high-SNR average was constructed as the mean of the 10 sequential acquisitions in each subject. This formed data with acquisition time equivalent to 45 min.

8. Data was saved in the NIfTI-MRS format with masks derived from hard thresholding of the water reference data, selecting a $16 \times 16 \times 1$ block in each subject.[25]

All volunteers were recruited and scanned in compliance with local ethical and legal requirements.

3.2 Bias and rank selection in simulated data

To examine the efficacy of the 2 rank threshold selection methods in low SNR data, synthetic MRSI data was prepared with an explicitly constructed rank-3 Casorati matrix. This was achieved by generating 3 spectral peaks described in the time domain as a sum of decaying exponentials:

$$\hat{x}_n = \sum_{k=1}^{K} c_k \exp((\alpha_k + j\omega_k)nt_s)\exp(j\phi_k), \tag{13}$$

with $k = 3$. The peaks were offset (at $\omega_k = -200$, 0, and 300 Hz) with different spatially varying concentrations (varying c_k) in an $8 \times 8 \times 1$ grid. Phase (ϕ_k) and linewidth (α_k) were constant. This numeric phantom is shown in Supporting Information Figure S1. Subsequently the phantom was generated with 50 MC repetitions at 6 different noise levels. The highest noise levels were picked to ensure that the MP and SURE rank threshold estimation techniques underestimated the rank of the data (Supporting Information Figure S2). Simulated SNR levels spanned a range of in vivo data quality: lowest spectral SNR = 2.1, highest = 44.5.

Each MC repetition in each noise level was denoised using 5 different approaches to rank threshold selection combined with local ST denoising. The rank threshold was either selected using the MP method, fixed at rank 2, fixed at rank 3, selected using the SURE SVHT algorithm, or the SURE SVT algorithm (that also applied soft thresholding in the denoising algorithm). The patch size was $3 \times 3 \times 1$ voxels with a stride of 1 for all algorithms.

All denoised data and the original noisy data was fitted using an explicit 3-peak "AMARES style" algorithm (equivalent to Equation [13] with K = 3; implementing bounds and manual starting values, but no "prior-knowledge" constraints) and optimized using the Scipy "curve_fit" trust region reflective algorithm.[26]

For each noise level and algorithm, a spectral RMSE was calculated using the true noiseless data, and a fitted concentration RMSE was calculated from the input metabolite concentration maps (Supporting Information Figure S1A).

$$RMSE_{\text{Spectral}} = \sqrt{\frac{\sum_r^N s(r,t) - \hat{s}(r,t)}{N}}, \tag{14}$$

and

$$RMSE_{\text{Concentration}} = \sqrt{\frac{\sum_k^{k=3}\sum_r^N c_k(r) - \hat{c}_k(r)}{3N}} \ . \tag{15}$$

In addition, the voxel-wise concentration error was calculated, and the mean, standard deviation (SD), and skewness of the resulting distribution calculated at each noise level. All metrics were calculated using all voxels and all MC repetitions ($N = 8 \times 8 \times 50 = 3200$ voxels).

3.3 Uncertainty in denoising of a uniform single-peak simulation

The performance of the LP, global ST, local ST, and LORA denoising algorithms were evaluated in a simple MC test using a single on-resonance Lorentzian peak. Data for this analysis was generated as an $8 \times 8 \times 1$ grid of voxels each containing identical signal. The signal was formulated as a decaying exponential with unit amplitude and 10 Hz spectral linewidth (full width at half maximum, [fwhm]), that is, following the form of Equation (13) with $K = 1$, $c_1 = 1$, $\omega_1 = 0$, $a_1 = -10\pi$, and $\phi_1 = 0$. Fifty instantiations of these data were created with independent and identically distributed complex Gaussian noise at each of 6 noise levels with SDs of 0.5, 0.1, 0.05, 0.01, 0.005, and 0.001.

At each noise level, for each MC repetition (separate noise instantiation), data was denoised using the LP, global ST, local ST, and global LORA algorithms. For this explicitly constructed rank = 1 case all algorithm rank thresholds were set to 1. All denoised data and original noisy data were fit using non-linear least squares (Scipy "curve_fit" Levenberg-Marquard algorithm[27]) using the same model as was used to generate the data (Equation 13). Each fit produced an estimate of the peak amplitude and the amplitude uncertainty. The Scipy "curve_fit" amplitude uncertainty is derived from the diagonal elements of the covariance matrix, which in turn, is derived from the numerically estimated Jacobian.

For the global and local ST methods the denoised time-domain variance and covariance were estimated using the proposed method (Equations 11 and 12), and subsequently 100 bootstrap fitting repetitions were carried out to estimate the amplitude uncertainty after fitting. The variance and covariance estimated using the proposed method were qualitatively compared to the MC estimated variance and covariance.

The "actual uncertainty" of the fitted amplitude for each denoising case and noise level was calculated using the SD across the MC repetitions. The "conventional uncertainty" was calculated as the mean of the fitting derived uncertainty. The approximation to the actual uncertainty was calculated as the mean of the bootstrap derived uncertainty. The uncertainty expected for the average of all the voxels in the $8 \times 8 \times 1$ grid was also calculated. All uncertainties were then summarized across all noise levels as a ratio to the actual (MC derived) uncertainty of the noisy data.

3.4 Evaluation of denoising methods in simulated ^{1}H MRSI

Denoising performance and uncertainty estimation were also assessed in realistic simulated ^{1}H-MRSI data. Simulated data was constructed from the median concentration, line

shape, frequency shift, and noise variance of the 5 high SNR (45-min equivalent) in vivo acquisitions. The high SNR acquisitions were fitted using the FMRIB Software Library (FSL)-MRS "Newton" algorithm using a basis set with 17 metabolites (ascorbate [Asc], aspartate [Asp], creatine [Cr], γ-aminobutyric acid [GABA], glucose [Glc], glutamine [Gln], glutamate [Glu], glycerophosphocholine [GPC], glutathione [GSH], myo-inositol [Ins], lactate [Lac], N-acetylaspartate [NAA], N-acetylaspartateglutamate [NAAG], phosphocholine [PCho], phosphocreatine [PCr], scyllo-inositol [scyllo], taurine [Tau]), default FSL-MRS macromolecules and a 3rd-order baseline.[19] Only the 256 voxels within the mask of each subject were fitted and used to construct the simulated data. Noiseless simulated data was constructed using the FSL-MRS Voigt line shape model with baseline and phase parameters set to 0.

From the noiseless simulated data, 50 different noise instantiations (MC repetitions) were created for each of 5 noise levels (Figure 1A). The noise levels were equivalent to 90, 45, 9, 4.5, and 2.25 min of scanning. Data was saved in NIfTI-MRS format with a single mask to identify the 256 voxels containing data and water reference scan for consistent metabolite concentration scaling.

A single noise level (9-min equivalent) was denoised using LP, global ST, local ST, global LORA, and local LORA (Figure 1B). All other noise levels were denoised using just the global and local ST methods. Rank thresholds were selected using the MP method, except for the LP and second stages of LORA, which used a fixed threshold of 20, which was heuristically estimated from the known number of spectral peaks.[1] Local versions of ST and LORA used a patch size of $4 \times 4 \times 1$ voxels with a stride of 1. Denoising (and subsequent fitting) was only applied to the $16 \times 16 \times 1$ voxel masked region. Each denoised data set and the original noisy data was fitted using fsl_mrsi using the parameters and basis set described at the start of this section, but with no baseline.[19] The concentrations of highly correlated metabolites were combined (namely Glu+Gln, Glc+Tau, PCho+GPC, NAA+NAAG, and Cr+PCr).

Variance and covariance for the ST denoised-data, estimated by the proposed method, was compared with that measured using the 50 MC repetitions. Five of the generated data sets at each noise level underwent bootstrap fitting. For each voxel of the selected data sets, additional correlated noise was added using the estimated covariance matrix before fitting was carried out. This process was repeated 40 times. The reported concentrations and uncertainties are calculated from the mean and SD of the 40 fits.

Forty repetitions were chosen based on additional simulations carried out on a single voxel of simulated data and the time taken for the fitting process. Different numbers of samples, ranging from 10 to 200, were repeatedly drawn from a pool of 500 bootstrapped fits. The resultant estimates were compared with the ground truth. Forty repetitions showed no appreciable bias up to intermediate noise levels and low SD. The average bootstrapped fit execution time for 1 MRSI data set was 6.3 h (running parallelized on 6 threads).

For analysis, 2 groups of metabolites were defined: "high signal" metabolites include Glu+Gln, Glc+Tau, PCho+GPC, NAA+NAAG, Cr+PCr, and Ins; whereas "all unique" contains all fitted metabolites (combined as described) including the macromolecular peaks.

For the analysis of all denoising methods (at a single noise level), mean concentration uncertainties were calculated for all methods and expressed as a fraction of the noisy data uncertainty. The means was calculated across all voxels and metabolites in the respective sets ("high signal" and "all unique"). RMSE was calculated and normalized using the fitted concentrations of the noiseless synthetic data. RMSE were calculated across all voxels and metabolites in the respective sets ("high signal" and "all unique").

For the analysis of ST methods (applied at all noise levels) the MC "actual uncertainty," FSL-MRS "conventional uncertainty," and the bootstrap estimated uncertainty were calculated for all voxels and metabolites in the respective sets ("high signal" and "all unique"). These values are expressed as ratios to the MC "actual uncertainty" of the noisy data. A ratio <1 indicates lower uncertainty (ie, better performance).

3.5 Reproducibility of in vivo ^{1}H MRSI

For each of the 5 subjects, the 10 sequential in vivo acquisitions were denoised as described above for the simulated data. The resultant denoised spectra, the original noisy spectra, and the combined high SNR (45-min equivalent) spectra were then fit. Fitting was performed as described above using fsl_mrsi (3rd order baseline, 17 metabolites, default FSL-MRS macromolecules, "Newton" algorithm with Voigt line shape).[19] Fitting was performed over the 256 masked voxels. Concentrations were references to the unsuppressed water reference scans for each subject, but no relaxation correction was carried out (ie, $T_E = 0$ ms, $T_R = 15$ s). Therefore, metabolite concentrations are expressed in "institutional-units."

The effect of each denoising method was then assessed by comparing the RMSE of the metabolite concentrations of denoised 4.5-min acquisitions to those from the 45-min acquisition. Additionally, reproducibility was assessed as the SD of the fitted metabolite concentrations across the 10 repeated acquisitions for each subject. This value was normalized for each metabolite (by the median concentration) and expressed as a ratio to the median of the voxel-wise reproducibility of the original noisy spectra.

A subset of the denoised in vivo data underwent the proposed bootstrap fitting approach to validate the fit against the measured changes in reproducibility. Bootstrap fitting was run on the noisy, and the local and global ST denoised instances of a single 4.5-min acquisition from each subject. The bootstrapped estimated uncertainties of each metabolite in each voxel of the ST-denoised data were expressed relative to the noisy uncertainties. These relative uncertainties were linearly regressed (OLS, stats-model package V0.12.2) against the matched metabolite-wise and voxel-wise relative reproducibility.

4 Results

4.1 Bias and rank selection in simulated data

Figure 2A,B shows the spectral RMSE and concentration RMSE, respectively. SURE-optimized SVHT produced the lowest spectral RMSE, whereas MP produced a consistently low concentration RMSE. At high noise levels (indicated by vertical dotted line in Figure 2B) SURE SVT and the fixed $R = 2$ algorithm achieved lower concentration RMSE than MP.

Voxel-wise error also showed that SURE SVT had lower mean error, skew, and equivalent SD to MP at high noise levels, despite relatively poor performance at low noise levels (Figure 3D-F). MP thresholding has high bias and skew at the highest noise levels, but outperforms SURE SVHT at most points as measured by concentration RMSE.

4.2 Uncertainty in denoising of a uniform single-peak simulation

MC measured frequency-domain SD showed significant non-uniformity for all denoising methods, with the highest variance coinciding with regions of highest signal (Figure 4).

Qualitative evaluation of the proposed variance estimation method for the ST algorithm resulted in good agreement with the MC estimated variance (Figure 5) and the covariance (Supporting Information Figure S3).

For the single peak simulations, all algorithms except the LP algorithm decreased the uncertainty in fitted amplitude (Figure 6). The LP algorithm increased the uncertainty. Furthermore, the conventional uncertainty significantly underestimated actual uncertainty for all algorithms ("Fit" vs. "MC" in Figure 6). The fitting algorithm estimated the conventional uncertainty for the global ST method is equal to the uncertainty predicted for the spatial average of all voxels (vertical dashed line). The bootstrap method was able to accurately estimate the actual uncertainty ("Bootstrap" in Figure 6). These results were consistent across all noise levels examined (Supporting Information Figure S4).

4.3 Evaluation of denoising methods in simulated ^{1}H-MRSI

For the 9-min synthetic MRSI data, all denoising algorithms (except LP) reduced the actual uncertainty (Figure 7A). Furthermore, in all cases the estimated conventional uncertainty was lower than the MC measured actual uncertainty, although that for the local ST method only underestimated by 17% (vs. 64% for global ST).

Specifically, the mean ($\pm$SD) actual uncertainty ratio of the "high signal" metabolites for the global ST, local ST, LP, global LORA, and local LORA methods were 0.67 ± 0.20, 0.56 ± 0.15, 1.07 ± 0.20, 0.68 ± 0.21, and 0.57 ± 0.15, respectively. The values were 0.57 ± 0.42, 0.57 ± 0.34, 1.22 ± 2.40, 0.58 ± 0.38, and 0.65 ± 2.55 for "all unique" metabolites (Supporting Information Figure S5). The median normalized RMSE showed the same pattern (Figure 7B), with RMSE of 0.056, 0.050, 0.068, 0.057, and 0.050 (noisy = 0.062) for the global ST, local ST, LP, global LORA, and local LORA methods, respectively.

Qualitative evaluation of the variance estimated by the proposed method against that measured by MC simulation showed that the proposed method slightly underestimated the variance, but predicted its non-uniform time (or spectral) dependence, as well as capturing features of the covariance structure (Supporting Information Figures S6 and S7).

The local and global ST methods were assessed at all 5 noise levels. Both methods reduced the actual uncertainty in metabolite concentrations at all noise levels (Figures 8 and S8), with the local ST algorithm outperforming the global algorithm, and achieving a reduction of more than half at the highest noise level.

The bootstrapping uncertainty accurately predicted the actual uncertainty of the global ST method at all but the lowest noise levels, where it over estimated (bootstrapping = 107% of the actual uncertainty, conventionally estimated = 41%). For the local ST method, the bootstrapping method was as accurate as the conventional estimate: bootstrapping = 90% of the actual uncertainty, conventional = 90% of the actual uncertainty. Figure 8B illustrates this for each noise level for the "high signal" metabolites. See Supporting Information Figure S8 for the "all unique" metabolites.

4.4 Reproducibility of ^{1}H-MRSI

In the in vivo data metabolite concertation RMSE was reduced by all ST and LORA methods, except for the highest signal metabolites: tNAA, tCr, tCho (Figure 9A). Although RMSE was lowered for these metabolites by the local patch-wise denoising implementations. The LP algorithm raised the RMSE in nearly all cases, and the ST algorithm outperformed the LORA algorithm.

Relative reproducibility (Figure 9B) showed similar results with median (±SD) ratios of 0.95 ± 0.53, 0.94 ± 0.53, 1.03 ± 0.59, 0.90 ± 0.48, 0.86 ± 0.52 for global ST, local ST, LP, global LORA, and local LORA, respectively.

The bootstrap fitting estimated uncertainty plotted against the relative reproducibility is plotted in Figure 9C (each point is a voxel-wise measurement of the high signal metabolites). The 2 independent measures of uncertainty showed moderate-to-high correlation (r = 0.58 and r = 0.71 for the ST global and ST local algorithms, respectively). Regression coefficients >1 (1.02 and 1.41, respectively) indicate an underestimation of the range of reproducibility by the bootstrap fitting approach.

5 Discussion and Conclusion

This work has demonstrated that low-rank denoising based on spatiotemporal separability can reduce uncertainty of estimated metabolite concentrations in both synthetic and in vivo ^{1}H-MRSI data. However, the reduction in uncertainty is typically not as great as is apparent, because the non-uniform residual variance is underestimated using conventional baseline noise inspection methods. Furthermore, denoising based on spectral linear predictability confers no benefit, as seen with increased uncertainty after LP denoising in all tests, and the similar performance of LORA when compared to ST denoising alone. In simulated ^{1}H-MRSI data, it is also apparent that patch based (local) methods perform better than

global methods. Local methods provided no improvement in the single-peak case or for the in vivo data, although the latter may be masked by inherent inter-scan variability, which cannot be removed by denoising, and optimal patch sizes are likely data-dependent. It should be noted that LORA might still be advantageous for situations where higher apparent denoising is desirable (eg, generating basis spectra).

Our results indicated that LP denoising is not effective on its own as a means of reducing metabolite estimation uncertainty (although it does reduce apparent spectral noise). This seems to be consistent with the literature, as evidenced by the assertion that spatiotemporal filtering should be performed before LP denoising in the original LORA work[1] to generate higher SNR inputs for the LP denoising stage, suggesting that the LP approach is not robust at low SNR. Other recent studies found the LP method (referred to as "HSVD") the least effective spectroscopic denoising method among those evaluated.[11] One reason for this is low-rank filtering is not simply a separation of signal and noise components, and the retained components are still biased by the noise.[29] Consequently, non-linear denoising by projecting data onto subspaces estimated from noisy signals results in signal-dependent variance.

Furthermore, it is important to note that, in this study, we focused on the downstream metabolite estimates as our measure of interest, rather than spectral RMSE, and our results, therefore, reflect the complex interaction between the denoising method and the spectroscopic fitting process. For example, in Figure 2 we show that, whereas rank thresholds chosen by SURE have optimal spectral RMSE, they often do not have optimal metabolite RMSE. One potentially promising direction of future development is the joint optimization of signal/image reconstruction, post-processing (eg, denoising), and analysis, which considers the spectroscopic processing from end-to-end. An area of particular importance is addressing the influence of structured (non i.i.d.) noise on fitting. Grage and Akke[30] note that unhandled correlation in NMR signals will result in less efficient estimation and underestimated uncertainties. By either directly modelling the noise correlation during fitting, or by incorporating a pre-whitening step, efficient parameter and accurate uncertainty estimation may be achieved.

The proposed uncertainty estimation was not applied to the LP denoising method, nor the LORA methods. Equation (11) requires the noise to be i.i.d., 0-mean, and normally distributed (eq. 3 in Chen et al).[15] This assumption is violated by the lifting of the data into a Hankel matrix for LP denoising, and in the second stage of LORA. Furthermore, the nonuniform variance created in the first stage of LORA cannot be propagated forward into the second stage using Equation (11). In simulations, it was found that Equation (11) only captured the broad structure of the signal-dependent nonuniform variance produced by LP denoising, with significant errors present in the estimation.

The proposed uncertainty estimation method (via bootstrapping) was found to be highly accurate in the case of a single peak. Although not shown in this work, the same was found to be the case when extended to simple multi-peak simulation. In simulated 1H-MRSI data, denoised by global ST, the bootstrap uncertainty was much closer to the MC-measured actual uncertainty than the conventional uncertainty estimate (107% vs. 41%). For the local

ST method, both the uncertainty estimated conventionally, and the proposed method were relatively accurate (both 90% of actual). This may be because the repeated averaging of patches removes the nonuniform variance and covariance that is apparent in the global case. This work has not assessed how many overlapping patches are needed to achieve this, or its dependence on patch stride or size. This may become important for data with fewer dimensions, for example, one dimensional dynamic (functional or diffusion weighted) MRS. The uncertainty of in vivo data predicted by the proposed method showed moderate-to-high correlation with inter-scan reproducibility. Without an in vivo ground truth, it is impossible to assess whether the unexplained variance in Figure 9C arises from the proposed method, or genuine inter-scan variability. Nevertheless, the proposed method still predicts the voxel-wise rank ordering of the uncertainty after denoising.

If more accurate covariance estimation was available, the bootstrapping uncertainty estimation could be improved. In simulated ^{1}H-MRSI data, the variance and covariance estimation method underestimated the denoised data variance and covariance (Supporting Information Figures S6 and S7). Our estimate of the off-diagonal covariance elements accurately captures the structure, but not the magnitude in a range of data (Supporting Information Figure S7 and supporting text "Covariance approximation"). Chen et al[15] note that the variance estimation (Equation 11) deteriorates at low SNR and has a dependence on condition number. Bootstrap fitting of MRSI data is slow and only 40 repetitions were used here in the simulated ^{1}H-MRSI data, more might also lead to higher accuracy.

In this work, we introduce the use of SURE and soft thresholding to the denoising of MRSI data. Although SURE selection of a hard threshold minimizes spectral RMSE, it does not also minimize the RMSE of the fitted metabolite concentrations. Both SURE and MP automated threshold selection lead to biased data in low SNR cases (because the rank threshold tends to 1 and signal variance is lost). The lower bias and skewness of the SURE SVT (soft thresholding) at high noise levels indicates that it or other thresholding functions may be optimal for low SNR regimes (or where trying to detect small signal fluctuations), such as phosphorus-31 MRS or diffusion-weighted MRS.[9,31] In fact, recent work investigating optimal matrix denoising[29] highlights the increased performance of soft thresholding in lower SNR regimes and also proposes optimized singular value shrinkage functions that outperform hard or soft thresholding in all SNR regimes. Evaluation of this parameter-free optimized singular value shrinkage method in both simulation and in vivo data would certainly be an interesting direction for future work.

We have not assessed the impact of algorithm parameters: undoubtedly patch size and stride for the local methods are important. Currently there is no automated way to select these. In addition, it is difficult to assess the true uncertainty reduction in in vivo data. Here, we have demonstrated modest (10%) decreases in inter-scan variability, but the true amount might be masked by inter-scan variability not caused by thermal noise (eg, physiological noise, or scanner drift).

The denoising tools used in this work have been made available for use as the Python package "mrs_denoising_ tools" and operate on 4 dimensional MRSI data stored in the NIfTI format.

In conclusion, low-rank spectroscopic denoising methods based on spatiotemporal (or dynamic temporal) separability do reduce uncertainty in MRS(I) data. However, thorough assessment of the method and use case should be made. It is important to select the right thresholding method in low SNR cases, and assessment simply by SNR measured from residual baseline noise is insufficient given the possibility of non-uniform variance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data Availability Statement

The source code and data used to generate the results presented in this manuscript can be found at https://git.fmrib.ox.ac.uk/wclarke/low-rank-denoising-uncertainty (#106c36a27061be4b1eb10edbf32e1c45bd1d2fa5). The specific MRSI denoising tools have been released as a Python package "mrs-denoising-tools" available from the Pypi and Conda package managers. The source code is available from https://git.fmrib.ox.ac.uk/wclarke/low-rank-denoising-tools. This work was created using version 0.0.2 of mrs-denoising-tools. Fitting was performed using version 1.1.2 of FSL-MRS, which is also available from the Conda package manager and https://git.fmrib.ox.ac.uk/fsl/fsl_mrs.

References

1. Nguyen HM, Peng X, Do MN, Liang Z. Denoising MR spectroscopic imaging data with low-rank approximations. IEEE Trans Biomed Eng. 2013; 60 :78–89. [PubMed: 23070291]

2. 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. [PubMed: 27523449]

3. Moeller S, Pisharady PK, Ramanna S, et al. NOise reduction with DIstribution Corrected (NORDIC) PCA in dMRI with complex-valued parameter-free locally low-rank processing. Neuroimage. 2021; 226 117539 [PubMed: 33186723]

4. Vizioli L, Moeller S, Dowdle L, et al. A paradigm change in functional brain mapping: suppressing the thermal noise in fMRI. bioRxiv. 2020 2020.11.04.368357

5. Lam F, Liang Z-P. A subspace approach to high-resolution spectroscopic imaging. Magn Reson Med. 2014; 71 :1349–1357. [PubMed: 24496655]

6. Lam F, Li Y, Guo R, Clifford B, Liang Z-P. Ultrafast magnetic resonance spectroscopic imaging using SPICE with learned subspaces. Magn Reson Med. 2020; 83 :377–390. [PubMed: 31483526]

7. Francischello R, Geppi M, Flori A, Vasini EM, Sykora S, Menichetti L. Application of low-rank approximation using truncated singular value decomposition for noise reduction in hyperpolarized 13C NMR spectroscopy. NMR Biomed. 2021; 34 e4285 [PubMed: 32125739]

8. Brender JR, Kishimoto S, Merkle H, et al. Dynamic imaging of glucose and lactate metabolism by 13 C-MRS without hyperpolarization. Sci Rep. 2019; 9 :3410. [PubMed: 30833588]

9. Froeling M, Prompers JJ, Klomp DWJ, van der Velden TA. PCA denoising and Wiener deconvolution of 31P 3D CSI data to enhance effective SNR and improve point spread function. Magn Reson Med. 2021; 85 :2992–3009. [PubMed: 33522635]

10. Abdoli A, Stoyanova R, Maudsley AA. Denoising of MR spectroscopic imaging data using statistical selection of principal components. Magn Reson Mater Phys Biol Med. 2016; 29 :811–822.

11. Rowland BC, Sreepada L, Lin AP. A comparison of denoising methods in dynamic MRS using pseudo-synthetic data. medRxiv. 2021 2021.02.23.21252282

12. Near J, Harris AD, Juchem C, et al. Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. NMR Biomed. 2021; 34 e4257 [PubMed: 32084297]

13. Cadzow JA. Signal enhancement-a composite property mapping algorithm. IEEE Trans Acoust Speech Signal Process. 1988; 36 :49–62.

14. Liang, Z. Spatiotemporal imaging with partially separable functions; 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 2007. 988–991.

15. Chen Y, Fan J, Ma C, Yan Y. Inference and uncertainty quantification for noisy matrix completion. Proc Natl Acad Sci. 2019; 116 :22931–22937. [PubMed: 31666329]

16. Song J, Xia S, Wang J, Patel M, Chen D. Uncertainty quantification of hyperspectral image denoising frameworks based on sliding-window low-rank matrix approximation. IEEE Trans Geosci Remote Sens. 2021 :1–12.

17. Candès EJ, Sing-Long CA, Trzasko JD. Unbiased risk estimates for singular value thresholding and spectral estimators. IEEE Trans Signal Process. 2013; 61 :4643–4657.

18. Ulfarsson MO, Solo V. Selecting the number of principal components with SURE. IEEE Signal Process Lett. 2015; 22 :239–243.

19. Clarke WT, Stagg CJ, Jbabdi S. FSL-MRS: an end-to-end spectroscopy analysis package. Magn Reson Med. 2021; 85 :2950–2964. [PubMed: 33280161]

20. Chiew M, Jiang W, Burns B, et al. Density-weighted concentric rings k-space trajectory for 1H magnetic resonance spectroscopic imaging at 7 T. NMR Biomed. 2018; 31 e3838

21. Steel A, Chiew M, Jezzard P, et al. Metabolite-cycled density-weighted concentric rings k-space trajectory (DW-CRT) enables high-resolution 1 H magnetic resonance spectroscopic imaging at 3-Tesla. Sci Rep. 2018; 8 :7792. [PubMed: 29773892]

22. Fessler JA, Sutton BP. Nonuniform fast Fourier transforms using min-max interpolation. IEEE Trans Signal Process. 2003; 51 :560–574.

23. Rodgers CT, Robson MD. Coil combination for receive array spectroscopy: are data-driven methods superior to methods using computed field maps? Magn Reson Med. 2016; 75 :473–487. [PubMed: 25820303]

24. Bolan PJ, Henry P-G, Baker EH, Meisamy S, Garwood M. Measurement and correction of respiration-induced B0 variations in breast 1H MRS at 4 Tesla. Magn Reson Med. 2004; 52 :1239–1245. [PubMed: 15562472]

25. Clarke, WT; Bell, T; Emir, UE; , et al. NIfTI MRS: a standard format for spectroscopic data; Proc. Intl. Soc. Mag. Reson. Med. 29; 2021.

26. Branch MA, Coleman TF, Li Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM J Sci Comput. 1999; 21 :1–23.

27. More, JJ. Levenberg–Marquardt Algorithm: Implementation and Theory. Argonne National Lab; IL (USA): 1977.

28. Seabold, S; Perktold, J. statsmodels: Econometric and statistical modeling with python; 9th Python in Science Conference; 2010.

29. Gavish M, Donoho DL. Optimal shrinkage of singular values. IEEE Trans Inf Theory. 2017; 63 :2137–2152.

30. Grage H, Akke M. A statistical analysis of NMR spectrometer noise. J Magn Reson. 2003; 162 :176–188. [PubMed: 12762994]

31. Jelescu, IO; Veraart, J; Cudalbu, C. MP-PCA denoising dramatically improves SNR in large-sized MRS data: an illustration in diffusion-weighted MRS; Proc. Intl. Soc. Mag. Reson. Med; 2020. 0395
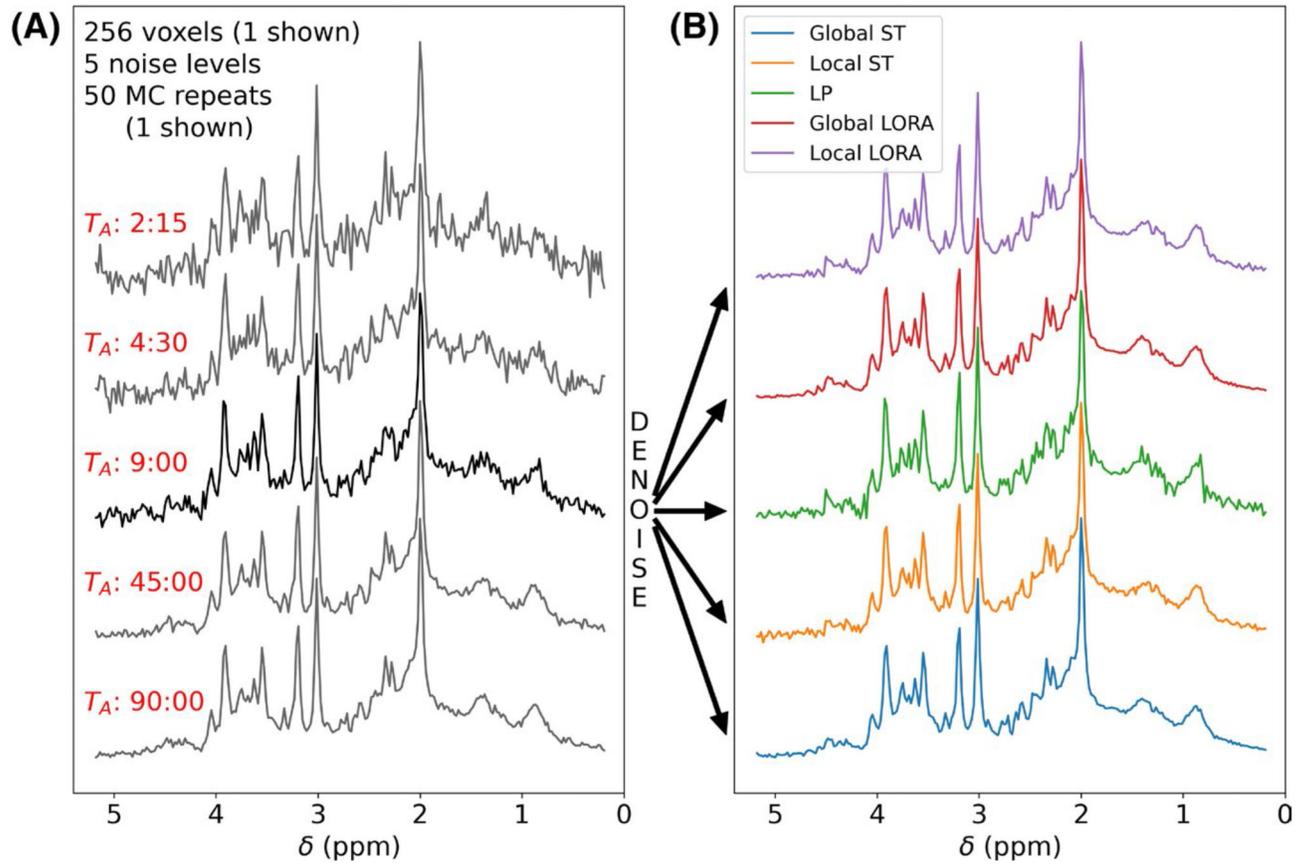
Figure 1.
(A) Example synthetic data at each of the 5 noise levels. (B) The 9-min equivalent data is
shown denoised by each of the 5 methods
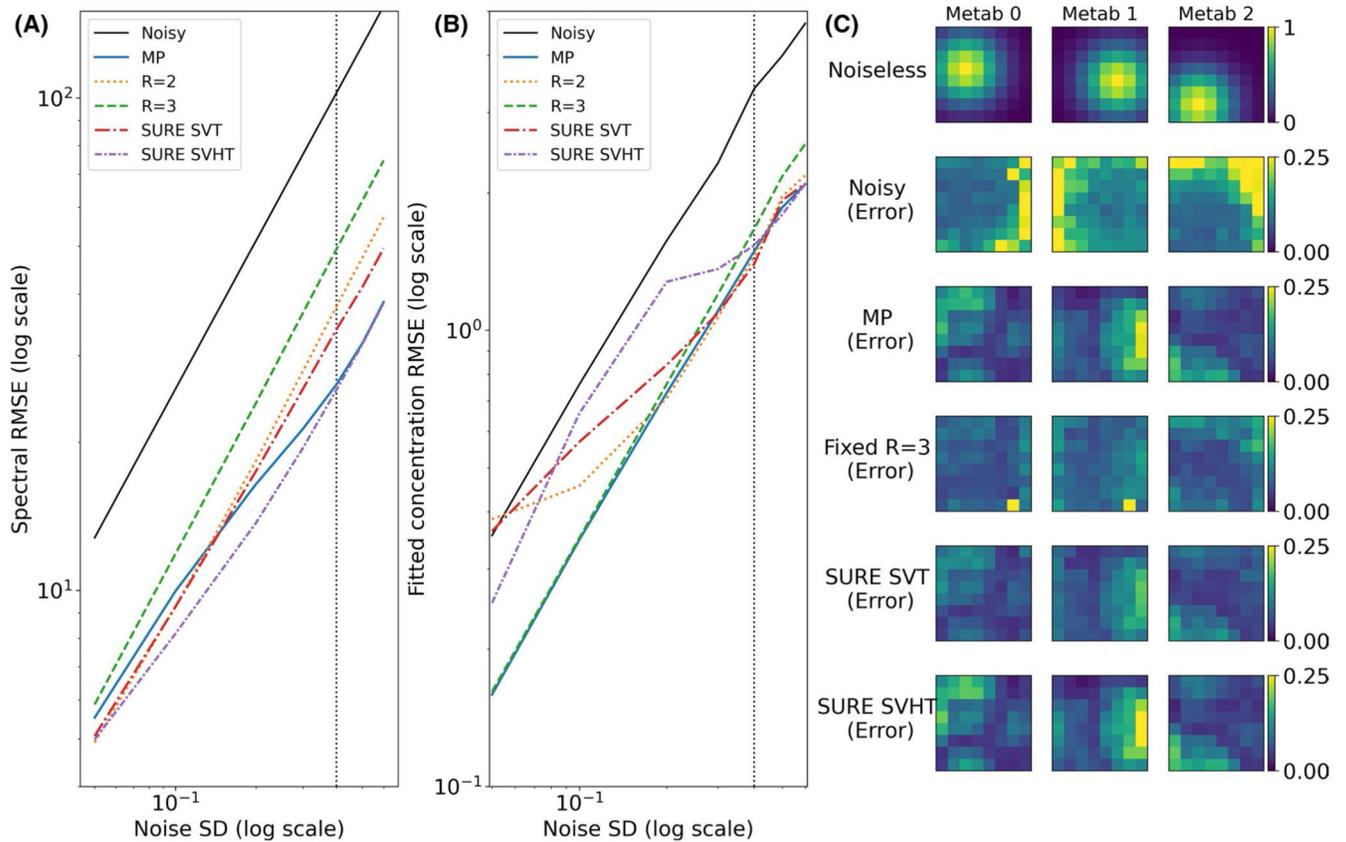
Figure 2.
Spectral RMSE (A) and concentration RMSE (B) of the 5 different thresholding approaches of the simulated rank = 3 data, compared to the noisy data. (C) The noiseless concentration maps for each of the 3 "metabolites" and the absolute error for each method at a single noise level (dotted vertical line in A and B). High bias is observed for SURE SVHT and MP, with less seen in SURE SVT and R = 3. However, the fixed R = 3 shows higher average error
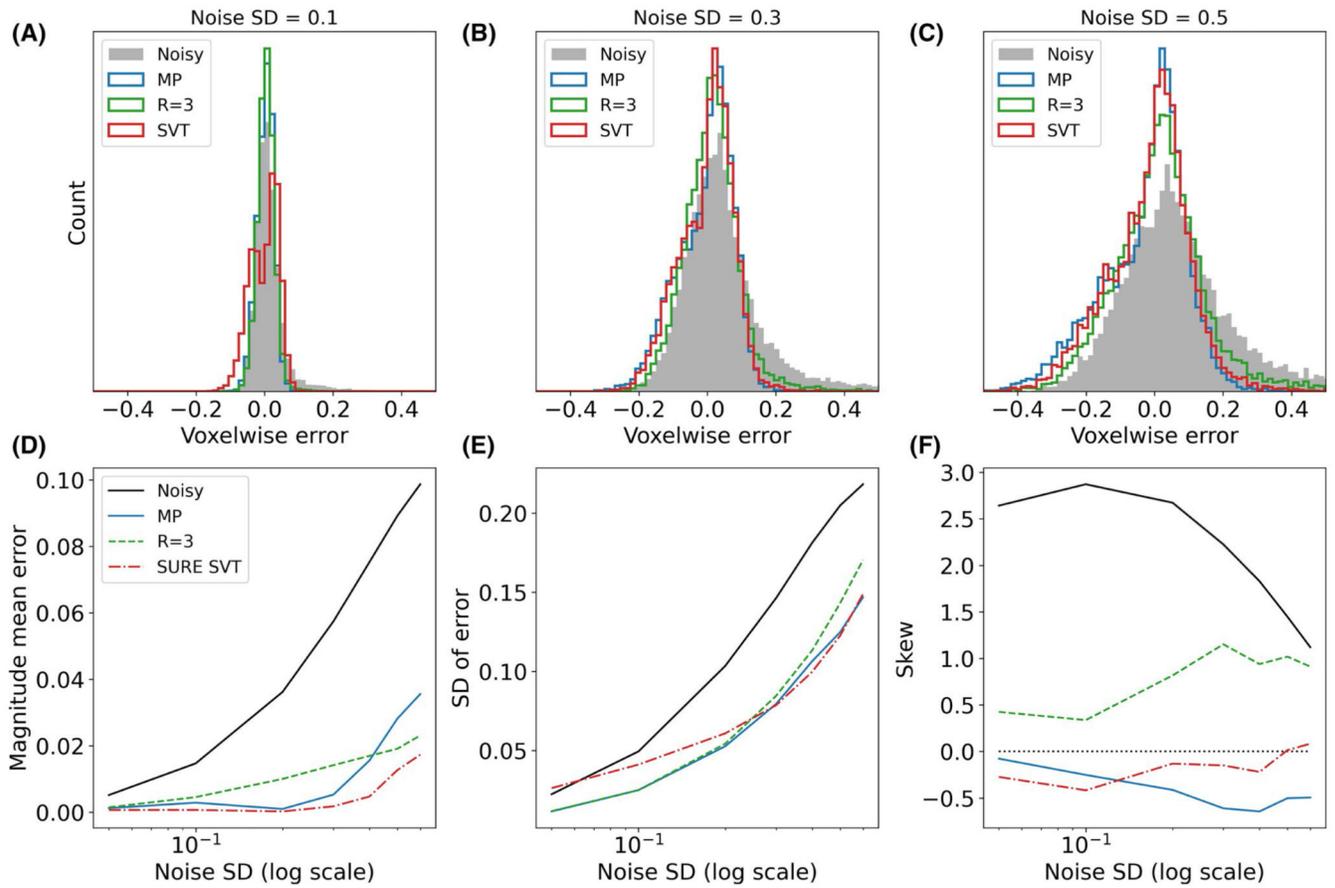
Figure 3.
(A-C) Voxel-wise error distributions for 3 methods at 3 different noise levels. (D-F)
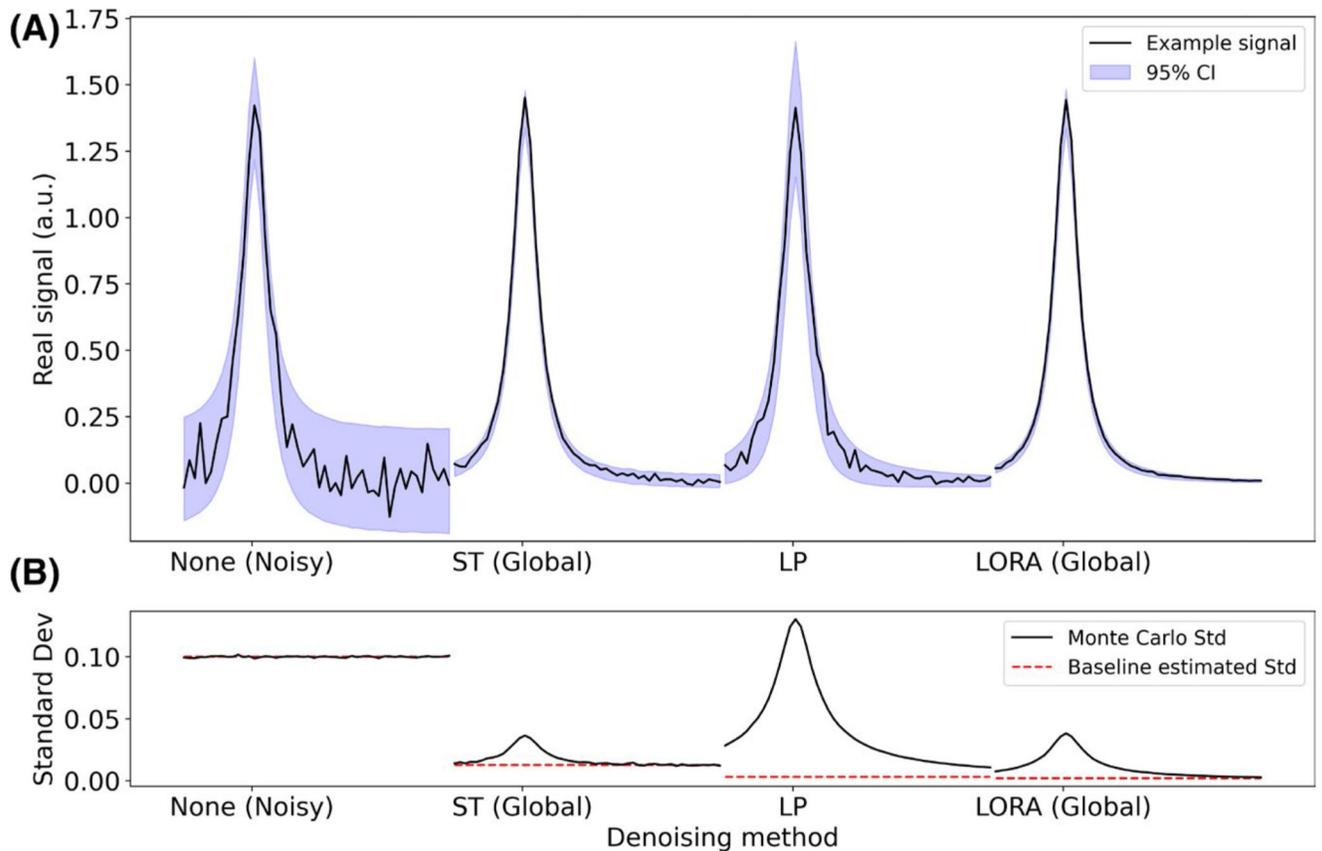Magnitude mean error, SD, and skew of the distributions formed by the voxel-wise error

Figure 4.
(A) Single peak after application of different low-rank denoising methods (black line) and Monte Carlo 95% confidence intervals (shaded). (B) Monte Carlo SD of the denoised data (black) and the noise level estimated from a signal-free region of baseline (red). Although the original noisy data (left) has uniform and high variance the denoised data shows very inhomogeneous variance, with higher values in areas with signal present. Data shown has original noise SD of 0.1. Only a limited frequency range is shown in each panel
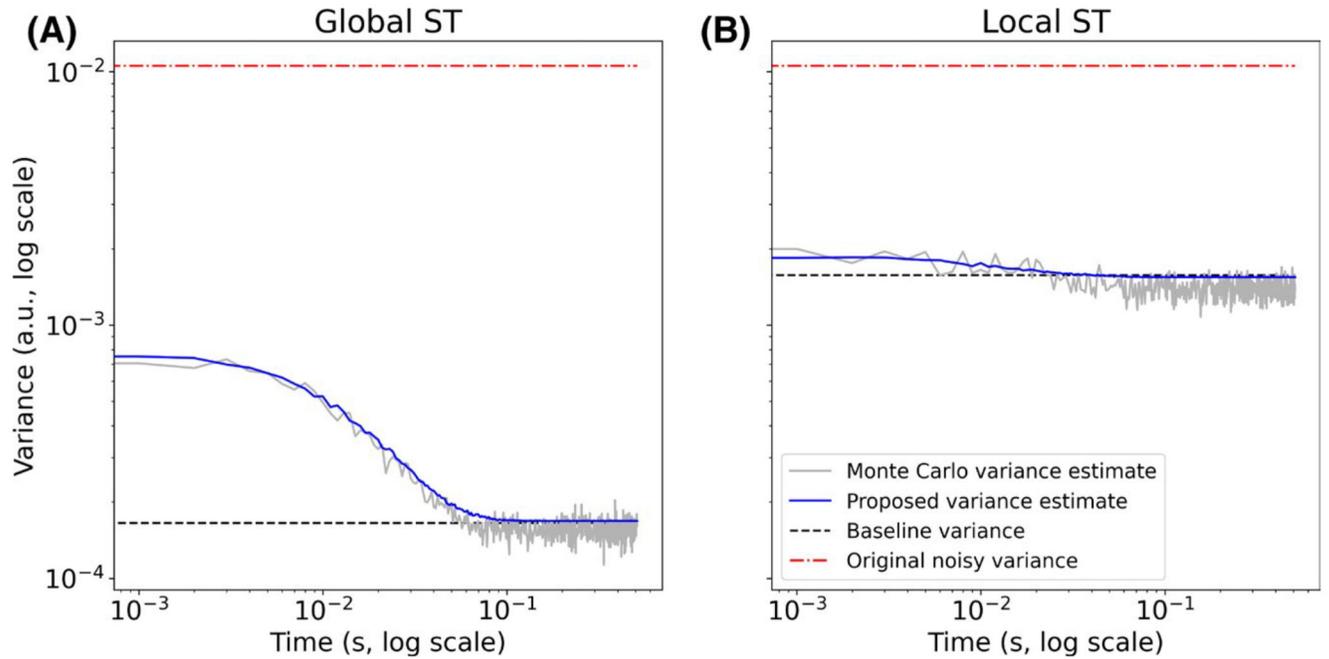
Figure 5.
The variance estimated using the proposed method (blue) after global ST (A) and local ST (B) compared to the Monte Carlo variance estimate (grey). Shown is the noise variance before denoising (red) and that which would be estimated from the denoised signal-free baseline (black)
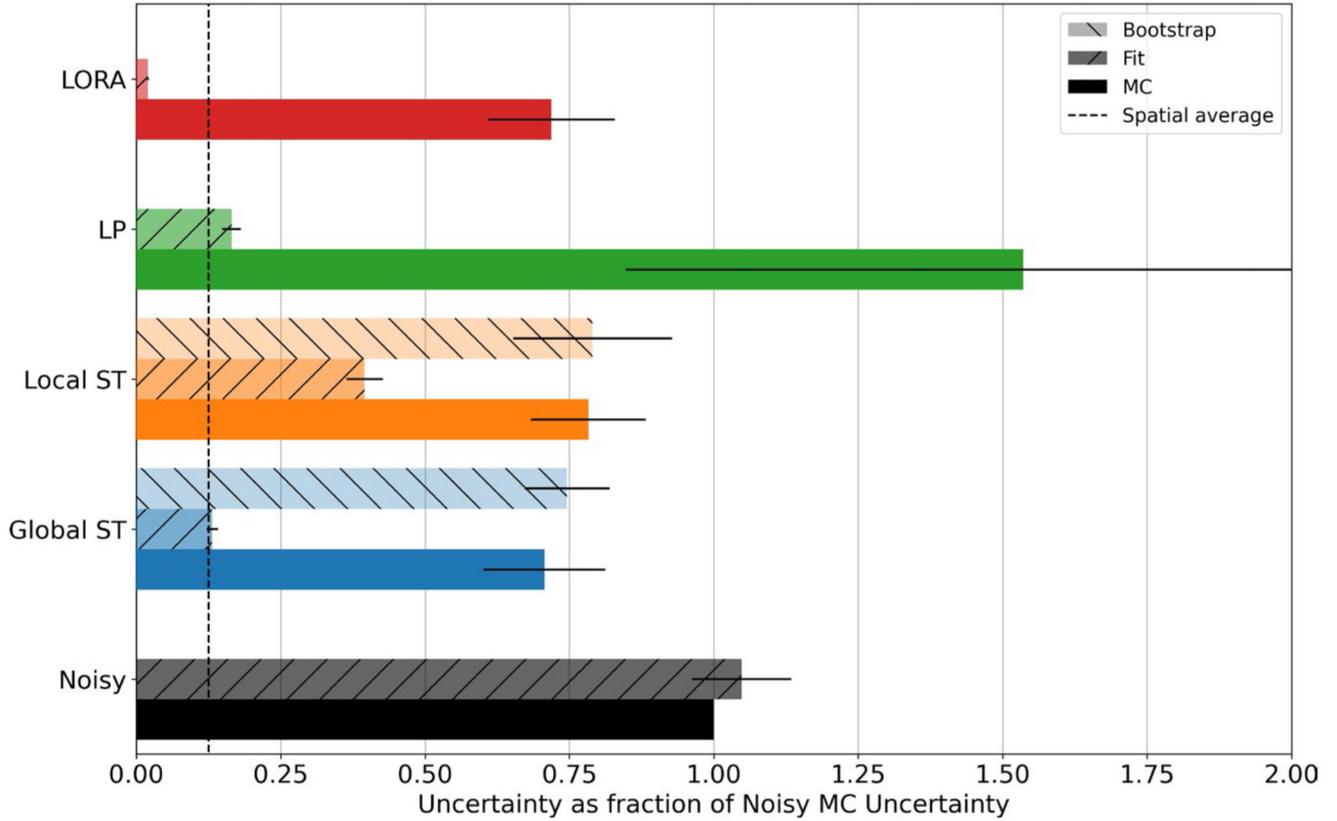
Figure 6.

Single peak fitted amplitude uncertainty (SD) relative to "Noisy" for each denoising algorithm. Data shows the Monte Carlo estimated actual uncertainty ("MC"), the conventional uncertainty ("Fit"), and the Bootstrap estimated uncertainty ("BS"), as mean and SD across all simulated noise levels. The LP algorithm (green) increases uncertainty, all others decrease uncertainty. The conventionally estimated uncertainty is not accurate and for the global ST case is equal to that which would be estimated for the spatial average of all voxels. The proposed bootstrapping method provides an accurate estimate of the uncertainty for both the local and global ST algorithms
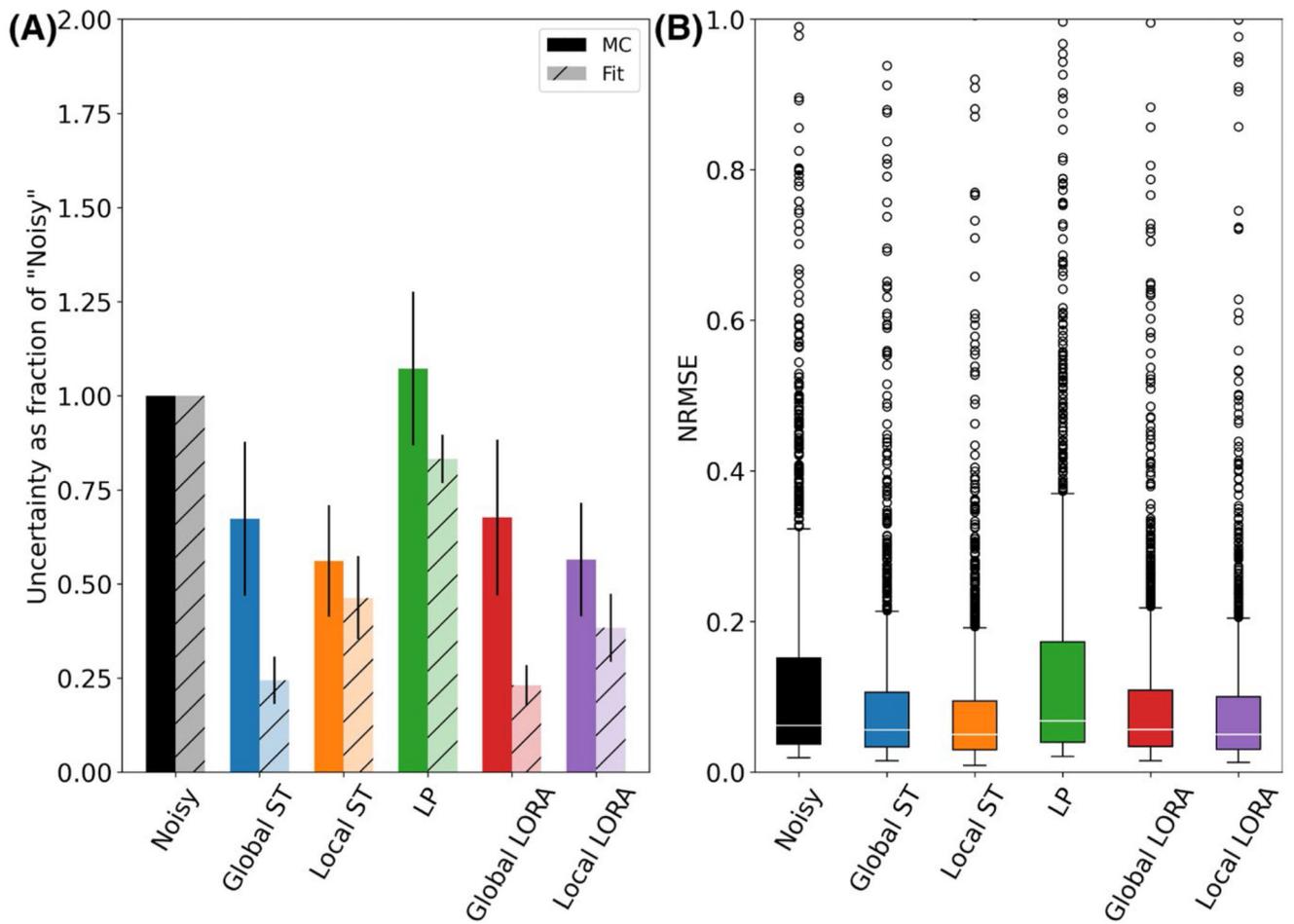
Figure 7.
(A) Mean uncertainty of the fitted concentrations of the "high signal" metabolites expressed as a ratio to the original noisy data. Actual uncertainty measured by Monte Carlo simulation is compared with that conventionally estimated by the FSL-MRS fitting algorithm. (B) Normalized RMSE of the noisy and denoised data comparing the fitted concentrations of the "high signal" metabolites to that of the noiseless synthetic data
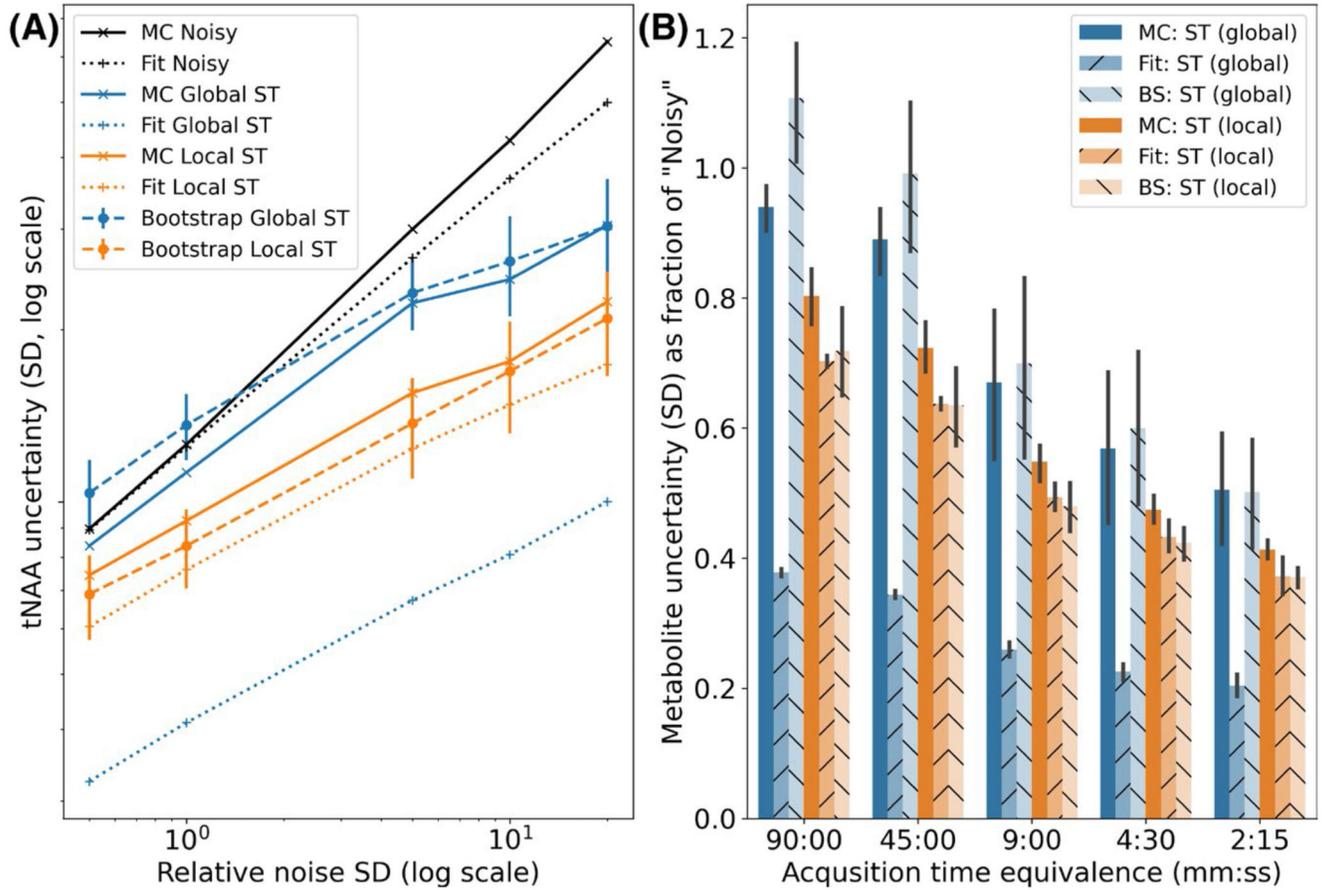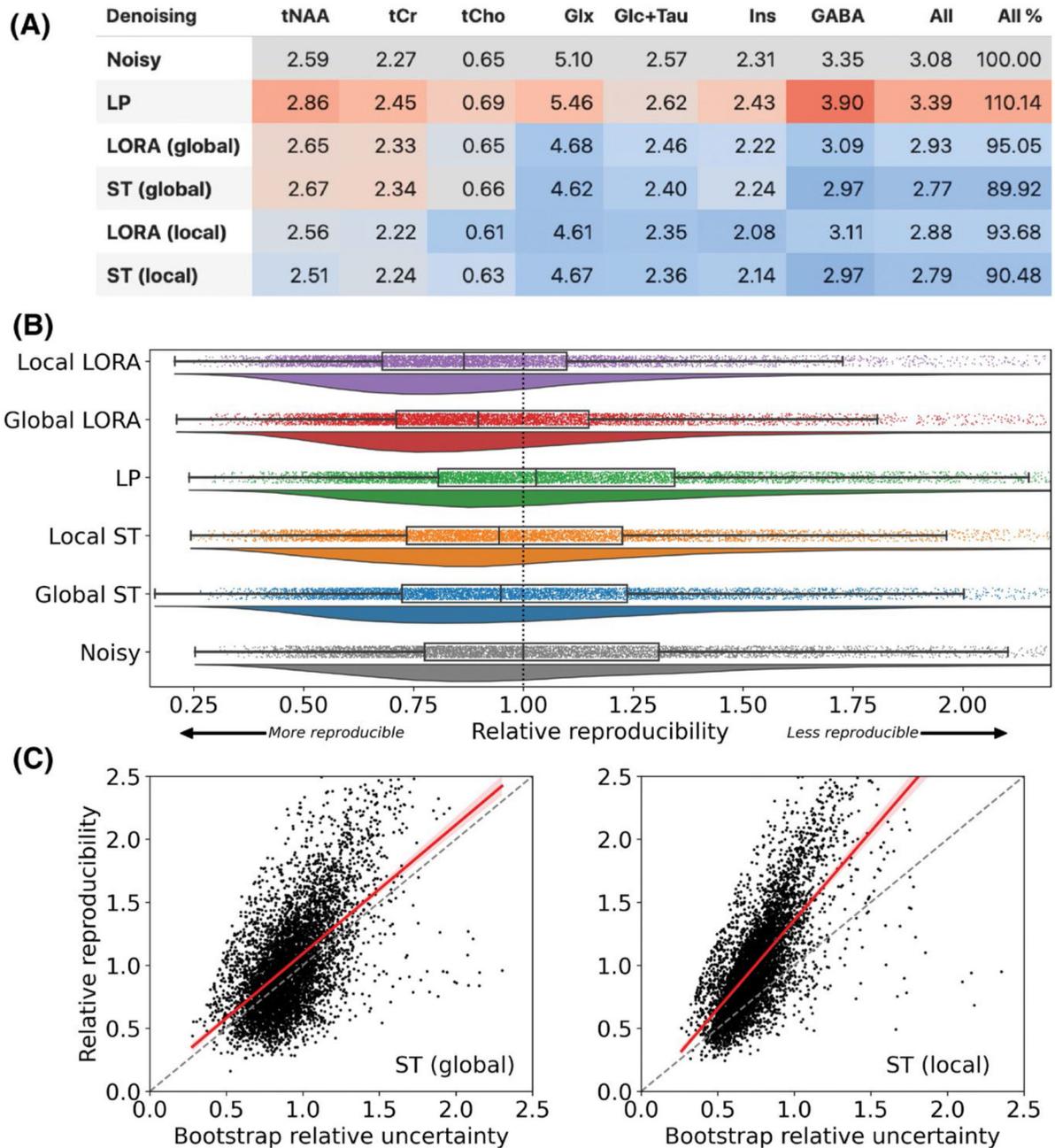
Figure 8.
Estimated uncertainty at different noise levels by Monte Carlo simulation, FSL-MRS fitting, and bootstrap fitting for a single combined resonance (A) (NAA+NAAG), and all "high signal" metabolites (B)

(A)

Denoising	tNAA	tCr	tCho	Glx	Glc+Tau	Ins	GABA	All	All %
Noisy	2.59	2.27	0.65	5.10	2.57	2.31	3.35	3.08	100.00
LP	2.86	2.45	0.69	5.46	2.62	2.43	3.90	3.39	110.14
LORA (global)	2.65	2.33	0.65	4.68	2.46	2.22	3.09	2.93	95.05
ST (global)	2.67	2.34	0.66	4.62	2.40	2.24	2.97	2.77	89.92
LORA (local)	2.56	2.22	0.61	4.61	2.35	2.08	3.11	2.88	93.68
ST (local)	2.51	2.24	0.63	4.67	2.36	2.14	2.97	2.79	90.48

Figure 9.

(A) RMSE for high signal metabolites in denoised 4.5-min scans compared to the 45-min equivalent average. Blue shading indicates values lower than the noisy baseline; red indicates values higher. (B) Rainfall plots of the relative reproducibility (SD of the ten 4.5-min scans) for the "high signal" metabolites. Relative reproducibility is the voxel-wise SD of the ten 4.5-min scans measured relative to the "Noisy" median. Lower values represent more

reproducible voxels. (C) Correlation of bootstrap measured relative uncertainty with relative reproducibility for the ST global (left) and ST local (right) denoised data