Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Simultaneous adaptive smoothing of relaxometry and quantitative magnetization transfer mapping

Siawoosh Mohammadi[1], Chiara D'Alonzo[2], Lars Ruthotto[3], Jörg Polzehl[2], Isabel Ellerbrock[1], Martina F. Callaghan[4], Nikolaus Weiskopf[5], Karsten Tabelow[2]

submitted: September 27, 2017

[1] Universitätsklinikum Hamburg-Eppendorf
Institut für Systemische Neurowissenschaften
Martinistr. 52
20246 Hamburg
Germany
E-Mail: s.mohammadi@uke.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: chiara.dalonzo@wias-berlin.de
joerg.polzehl@wias-berlin.de
karsten.tabelow@wias-berlin.de

[3] Emory University
Department of Mathematics and Computer Science
Atlanta, GA
USA
E-Mail: lruthotto@emory.edu

[4] Wellcome Trust Centre for Neuroimaging
UCL Institute of Neurology
12 Queen Square
London WC1N 3BG
United Kingdom

[5] Max Planck Institute for Human Cognitive and Brain Sciences
Department of Neurophysics
Stephanstr. 1A
04103 Leipzig
Germany

No. 2432

Berlin 2017

2010 *Mathematics Subject Classification.* 62P10, 62G05.

Key words and phrases. Quantitative MRI, Multi-Paramter Mapping, structural adaptive smoothing.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Simultaneous adaptive smoothing of relaxometry and quantitative magnetization transfer mapping

Siawoosh Mohammadi, Chiara D'Alonzo, Lars Ruthotto, Jörg Polzehl, Isabel Ellerbrock, Martina F. Callaghan, Nikolaus Weiskopf, Karsten Tabelow

Abstract

Attempts for in-vivo histology require a high spatial resolution that comes with the price of a decreased signal-to-noise ratio. We present a novel iterative and multi-scale smoothing method for quantitative Magnetic Resonance Imaging (MRI) data that yield proton density, apparent transverse and longitudinal relaxation, and magnetization transfer maps. The method is based on the propagation-separation approach. The adaptivity of the procedure avoids the inherent bias from blurring subtle features in the calculated maps that is common for non-adaptive smoothing approaches. The characteristics of the methods were evaluated on a high-resolution data set (500 μm isotropic) from a single subject and quantified on data from a multi-subject study. The results show that the adaptive method is able to increase the signal-to-noise ratio in the calculated quantitative maps while largely avoiding the bias that is otherwise introduced by spatially blurring values across tissue borders. As a consequence, it preserves the intensity contrast between white and gray matter and the thin cortical ribbon.

1 Introduction

In contrast to classical weighted imaging, e.g., with T1-, T2-, or PD-weighting, that acquires intensity values in arbitrary units, quantitative MRI (qMRI) measures absolute physical parameters, e.g., relaxation times in relaxometry (Koenig, Brown, and Ugolini 1993) or an estimate of water diffusivity in diffusion MRI (Le Bihan 2003). If corrected for instrumental bias (Lutti et al. 2010) its standardized nature facilitates comparison across sites, time points (Weiskopf et al. 2013) and participants in multi-centre trials. Multi-Parameter Mapping (Weiskopf et al. 2013; Lutti et al. 2014) is a framework for qMRI that simultaneously measures the proton density (PD), the longitudinal relaxation rate (R_1), the apparent transverse relaxation rate ($R_2^\star$) and, optionally, the magnetization transfer saturation (MT). Those metrics have been shown to be sensitive to microstructure by means of, e.g., iron or myelin content (Callaghan et al. 2014), and can be used in biophysical models (Mohammadi et al. 2015) to estimate microscopic tissue features using in-vivo MRI, e.g., the aggregated g-ratio of myelinated axons (Stikov et al. 2015).

However, the spatial scale gap complicates the estimation of microscopic tissue features from MRI: While MRI typically works on a millimeter scale, microscopic tissue features, e.g., the g-ratio of a myelinated axon are at the micron scale. Although it is not foreseeable that in-vivo MRI will ever achieve a spatial resolution of microns, it is believed that qMRI data at submillimeter resolution will facilitate improved biophysical modelling (Roebroeck et al. 2008; Weiskopf et al. 2015). One important limiting factor for increasing the spatial resolution of qMRI towards submillimeter resolution is the fact that the MR signal-to-noise ratio (SNR) inherently drops with the voxel volume (Edelstein et al. 1986). Besides expensive hardware improvements, e.g., ultra-high field MRI (Lüsebrink et al. 2017),

DOI 10.20347/WIAS.PREPRINT.2432 Berlin 2017

and SNR-efficient MRI sequences (Xu et al. 2013; Setsompop et al. 2017), denoising methods can be used to increase the SNR. While non-adaptive denoising methods are efficient in reducing the variation in images, they do not preserve tissue boundaries and inherently decrease the nominal resolution. This is because non-adaptive denoising uses a global smoothing strategy not adapted to the specific local data properties and hence leads to a local bias in the smoothed MR image signal values. For qMRI, bias is particularly problematic, since the MR signal is physically interpreted and a bias in the MR signal will convert to artefactual quantitative parameter maps. However, the acquired images contain spatial structure that can be utilized to reliably improve the SNR while avoiding this bias by using adaptive denoising methods. Those adapt the smoothing to the local properties of the data at hand and thus achieve variance reduction, i.e., an improvement of the SNR, but with a largely reduced bias.

Adaptive denoising has been used frequently in other MR imaging modalities like diffusion MRI. The specific methods are based on different methodologies, like anisotropic diffusion (Ding, Gore, and Anderson 2005; Parker et al. 2000; Xu et al. 2010; Duits and Franken 2011), non-local means (Manjón et al. 2009), penalization techniques (McGraw et al. 2009; Haldar et al. 2013), wavelet filtering (Lohmann et al. 2010), model-based methods (Fletcher 2004; Tabelow et al. 2008), the propagation-separation approach (Becker et al. 2012; Becker et al. 2014), or other local techniques (Aja-Fernández et al. 2008; Tristán-Vega and Aja-Fernández 2010). Depending on the method and data adaptive smoothing is performed on each image volume separately or using the combined information from all available raw data.

To our knowledge, adaptive denoising has never been used for increasing the SNR in relaxometry. Here, we introduce an adaptive denoising strategy based on the propagation-separation approach (Polzehl and Spokoiny 2006), specifically designed for the FLASH-based MPM model (Helms, Dathe, and Dechent 2008; Helms et al. 2008; Weiskopf et al. 2013; Lutti et al. 2014). We follow the strategy of estimating a sequence of local adaptive smoothing kernels in scale-space at each voxel from all MPM metric maps simultaneously. To this end, we

- derive a novel analytical relation between the FLASH signal and the MPM metrics R_1, PD, $R_2^\star$, and MT,
- apply the propagation-separation approach to adaptively improve the SNR in the MPM metrics,
- and compare the efficiency, the bias, and the reduction in scan-rescan variability of adaptive to non-adaptive noise reduction.

2 Theory

2.1 Model of the signal in FLASH sequences

The signal obtained in a perfectly spoiled FLASH sequence is given by the Ernst equation (Ernst and Anderson 1966). Specifically, we denote the echo time by TE, the repetition time by TR, the flip angle by α and model the signal intensity S in the acquisition by (Helms, Dathe, and Dechent 2008)

$$S = A \cdot \sin\alpha \cdot \frac{1 - e^{-R_1 \cdot TR}}{1 - \cos\alpha \cdot e^{-R_1 \cdot TR}} \cdot e^{-R_2^\star \cdot TE}. \tag{1}$$

A denotes the maximum signal amplitude at $TE = 0$ and is related to the proton density $A = c \cdot PD$, with c a spatially varying factor related to the detection sensitivity (Weiskopf et al. 2013).

$R_1 = 1/T_1$ and $R_2^\star = 1/T_2^\star$ are the longitudinal and transverse relaxometry rates, respectively. In order to maintain a clear derivation of the formulas used in this paper we introduce the abbreviation $E_1 = e^{-R_1 \cdot TR}$ and write Eq. (1) slightly more compactly as:

$$S = A \cdot \sin\alpha \cdot \frac{1 - E_1}{1 - \cos\alpha \cdot E_1} \cdot e^{-R_2^\star \cdot TE}. \tag{2}$$

For a dual-excitation FLASH sequence the acquired signal S can be written (Helms et al. 2008) as

$$S = A \cdot \sin\alpha_1 \cdot \frac{1 - e^{-R_1 \cdot TR} - (1 - \cos\alpha_2) \cdot (1 - e^{-R_1 \cdot TR_1}) \cdot e^{-R_1 \cdot TR_2}}{1 - \cos\alpha_1 \cdot \cos\alpha_2 \cdot e^{-R_1 \cdot TR}} \cdot e^{-R_2^\star \cdot TE}$$

where $TR = TR_1 + TR_2$ is the total repetition time and α_1 and α_2 are the two flip angles of the two excitations. For the MT (magnetization transfer) FLASH experiment the second excitation is replaced by the MT pulse (Helms et al. 2008). Using the abbreviation $\delta = 1 - \cos\alpha_2$ that describes the MT saturation we obtain the signal intensity

$$S = A \cdot \sin\alpha_1 \cdot \frac{1 - e^{-R_1 \cdot TR} - \delta \cdot (e^{-R_1 \cdot TR_2} - e^{-R_1 \cdot TR})}{1 - \cos\alpha_1 \cdot (1 - \delta) \cdot e^{-R_1 \cdot TR}} \cdot e^{-R_2^\star \cdot TE} \tag{3}$$

for MTw acquisitions. Introducing another abbreviation $E_2 = e^{-R_1 \cdot TR_2}$ we will use a more compact version of Eq. (3), too:

$$S = A \cdot \sin\alpha_1 \cdot \frac{1 - E_1 - \delta \cdot (E_2 - E_1)}{1 - \cos\alpha_1 \cdot (1 - \delta) \cdot E_1} \cdot e^{-R_2^\star \cdot TE} \tag{4}$$

From these signal equations (1) and (3) one can estimate maps of R_1, $R_2^\star$, A, and δ by regression from a suitable set of multi-echo proton density weighted (PDw), T_1-weighted (T1w) and MTw acquisitions, i.e., for specific choices of α, TR, and the echo times TE, cf. Weiskopf et al. 2013. Here, we will follow a different approach for map estimation.

2.2 Re-parameterization of the signal model by ESTATICS

The ESTATICS model was introduced in Weiskopf et al. 2014 for estimating the apparent transverse relaxation time $R_2^\star$ from a multi-parameter mapping (MPM) multi-echo sequence typically involving in practice 6 to 8 echos for each of the PDw, T1w, and MTw acquisitions. Specifically, the model assumes that the transverse signal decay is independent of the contrast weighting:

$$S = (S_{T1} \cdot I_{T1} + S_{PD} \cdot I_{PD} + S_{MT} \cdot I_{MT}) \cdot e^{-R_2^\star \cdot TE} \tag{5}$$

with indicator variables I_{T1}, I_{PD}, and I_{MT} for the different acquisitions. Comparing this model equation with the signal equations (1) and (3) the (non-linear) terms involving R_1, A, and δ are replaced by S_{T1}, S_{MT}, and S_{PD}:

$$S_{T1} = A \cdot \sin\alpha_{T1} \cdot \frac{1 - E_1^{(T1)}}{1 - \cos\alpha_{T1} \cdot E_1^{(T1)}} \tag{6}$$

$$S_{PD} = A \cdot \sin\alpha_{PD} \cdot \frac{1 - E_1^{(PD)}}{1 - \cos\alpha_{PD} \cdot E_1^{(PD)}} \tag{7}$$

$$S_{MT} = A \cdot \sin\alpha_{MT} \cdot \frac{1 - E_1^{(MT)} - \delta \cdot \left(E_2^{(MT)} - E_1^{(MT)}\right)}{1 - \cos\alpha_{MT} \cdot (1 - \delta) \cdot E_1^{(MT)}} \tag{8}$$

In Appendix A we explicitly derive how R_1, A, and δ can be calculated analytically from S_{T1}, S_{PD}, and S_{MT} if $E_1^{(T1)} = E_1^{(PD)}$, i.e., if the repetition time TR for the T1w and PDw acquisitions are identical, an assumption that can be easily fulfilled for MPM sequences. Then, Eq. (5) can be considered as a re-parameterization of the model (1) and (3) and first used for the determination of estimates $\hat{S}_{T1}^{(0)}$, $\hat{S}_{PD}^{(0)}$, $\hat{S}_{MT}^{(0)}$, and $\hat{R}_2^{\star(0)}$ and second, utilizing those, the calculation of estimates $\hat{R}_1$, $\hat{A}$, and $\hat{\delta}$ via Eqs. (21), (22), and (23) in Appendix A, and obviously $\hat{R}_2^{\star} \equiv \hat{R}_2^{\star(0)}$. The meaning of the additional superscript (0) will be clarified in the next section.

2.3 Adaptive smoothing the estimated parameter maps

In this section, we outline an adaptive noise reduction procedure for the estimated parameter maps $\hat{S}_{T1}^{(0)}$, $\hat{S}_{PD}^{(0)}$, $\hat{S}_{MT}^{(0)}$, and $\hat{R}_2^{\star(0)}$ that can be used to reduce the noise variability of the maps without blurring the edges of the anticipated fine structures. This is in contrast to the non-adaptive Gaussian smoothing, that is, e.g., applied in functional MRI to increase the sensitivity of the signal detection. The method is based on the Propagation-Separation approach introduced in Polzehl and Spokoiny 2006. It makes use of the fact that the ESTATICS re-parametrization Eq. (5) of the signal model Eqs. (1) and (3) has a low parameter-induced non-linearity (Bates and Watts 1980) and leads to approximate Gaussianity of the estimates

$$\vec{S}^{(0)} := \left(\hat{S}_{T1}^{(0)}, \hat{S}_{PD}^{(0)}, \hat{S}_{MT}^{(0)}, \hat{R}_2^{\star(0)}\right)^{\top}.$$

The proposed adaptive smoothing is an iterative procedure employing kernel estimates (Fan and Gijbels 1996) with a) an increasing sequence of bandwidths $h_{k=1,\ldots,k^\star}$, and b) locally varying kernels by means of adaptive weighting schemes. Doing so, the procedure then automatically avoids blurring at borders of a homogeneous intensity region in the maps generally observed for kernel estimators like the Gaussian filter while achieving almost optimal noise reduction in homogeneous regions (Polzehl and Spokoiny 2006). The details of the procedure are explained in Appendix B.

The resulting (adaptively) smoothed ESTATICS parameter maps after $k^\star$ iterations

$$\vec{S}^{(k^\star)} := \left(\hat{S}_{T1}^{(k^\star)}, \hat{S}_{PD}^{(k^\star)}, \hat{S}_{MT}^{(k^\star)}, \hat{R}_2^{\star(k^\star)}\right)^{\top}$$

can then be used instead of the initial estimates $\vec{S}^{(0)}$ from the ESTATICS model to determine (implicitly) smoothed R_1, A, and δ-maps through the analytic formulae in Eqs (21), (22), and (23) in Appendix A and obviously $R_2^{\star} \equiv \hat{R}_2^{\star(k^\star)}$. If $k^\star = 0$, i.e., the initial estimates $\vec{S}^{(0)}$ are used, no smoothing is applied to the maps.

3 Methods

3.1 Image acquisition and subjects

We acquired an ultra-high resolution dataset at an isotropic resolution of 500 μm from a single subject as well as data from a group of 12 subjects at an isotropic resolution of 800 μm.

3.1.1 Data acquisition I: Ultra-high resolution dataset (500μm)

Multi-parameter mapping (MPM). One healthy subject was scanned on a 3T Tim Trio (Siemens Healthcare, Erlangen Germany) using the RF body coil to transmit and a 32 channel RF coil to receive. Calibration data were acquired to correct for inhomogeneities in the RF transmit field followed by the acquisition of spoiled multi-echo 3D fast low angle shot (FLASH) acquisitions with predominantly proton density (PD, TR/flip angle = 32ms/6°) and T1 (32ms/30°) weighting. Data were acquired with occipital coverage at an **isotropic resolution of 500μm**. Eight equidistant gradient echoes were acquired with alternating readout polarity with echo times ranging from 4.5 to 21.86ms in steps of 2.48ms. Three repetitions of the data were acquired. To accelerate the data acquisition, partially parallel imaging using the GRAPPA algorithm was employed in the phase-encoded direction, using forty integrated reference lines. A partial Fourier scheme was employed in the phase-encoded left-right direction using a factor of 6/8. Total scan time was about 45 minutes.

3.1.2 Data acquisition II: Group study (800 μm)

Sample. The study sample consisted of 12 healthy volunteers (mean 25.6 $\pm$ 2.2 years, range 22-30 years, 8 female). All participants were recruited locally at the University Medical Center Hamburg-Eppendorf. They were systematically screened to ensure they were free of any lifetime history of neurological or psychiatric illness, provided written informed consent before being included in the study and received remuneration for participation. The study complied with the principles of the Declaration of Helsinki and was approved by the local ethics committee (Ärztekammer Hamburg). Parts of the data used here were used as well in another study (under revision, (Ellerbrock and Mohammadi 2017)).

Multi-parameter mapping (MPM). All MRI sessions were performed on a whole-body 3T Tim TRIO scanner (Siemens, Erlangen, Germany) using a 32-channel radio-frequency (RF) head coil for receive and body coil for transmission. Foam padding was used to stabilize participants' head and minimize motion. An extensive whole-brain qMRI protocol was acquired consisting of MPM based on multi-echo 3D fast low-angle shot (FLASH): First, rapid calibration data were acquired to correct for inhomogeneities in the RF transmit field (Lutti et al. 2010; Lutti et al. 2012). Second, highly SNR-efficient spoiled multi-echo 3D FLASH data were acquired with predominantly PD-, T1- or MT- weighting according to the MPM protocol (Weiskopf et al. 2013). The flip angle was 6° for the PD- and MT-weighted volumes and 21° for the T1-weighted acquisition. MT-weighting was achieved through the application of a Gaussian RF pulse 2 kHz off resonance with 4 ms duration and a nominal flip angle of 220°. To ensure whole-brain coverage at an isotropic resolution of 800 μm, data were acquired using a field-of-view (FoV) of 256 mm head- foot, 224 mm anterior-posterior (AP) and 180 mm right-left (RL). Gradient echoes were acquired with alternating readout gradient polarity at equidistant echo times ranging from 2.34 to 18.44 ms in steps of 2.30 ms using a readout bandwidth of 488 Hz/pixel. For the PD- and T1-weighted acquisitions 8 echoes were acquired. For the MT-weighted acquisition only 6 echoes were acquired in order to maintain a repetition time (TR) of 25 ms for all FLASH volumes. To accelerate the data acquisition, partially parallel imaging using the GRAPPA algorithm was employed in each phase-encoded direction (AP and RL) with forty integrated reference lines and a speed up factor of two. Total scanning time of the MPM protocol was ~25 min. For all participants, MPM acquisition were conducted twice within one week (6-8 days interval) to test reproducibility of measures.

3.2 Spatial alignment, data preprocessing and analysis

3.2.1 Software

Data analyses were performed using MATLAB (The MathWorks, Natick, MA, USA, version R2014b), including customized MATLAB tools, FORTRAN code for the computationally expensive parts of the smoothing method, and Statistical Parametric Mapping (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK). The Gauss-Newton optimization code for the estimation of the ESTATICS parameter was implemented in Matlab. All preprocessing steps were deployed identically for both sets of 800μm resolution images acquired in the two sessions. We checked the validity of the code by comparing the results to those of a corresponding implementation in a package for the R software environment for statistical computing and graphics (R Core Team 2017).

The computation time for a single dataset including parameter estimation, smoothing and application of all corrections was approximately 10 minutes on a current laptop computer with i7 processor and 16GB of memory when restricting the calculation to a brain mask.

3.2.2 Spatial alignment

For the scan-rescan analysis all data acquired per subject from acquisition II were registered across the two time points, using a rigid-body transformation (`spm_realign`). First, to register data from the second to the first time point, skull-stripped and thresholded MT maps were used for each time point (thresholds: $MT > 0$ & $MT < 3\%$). The estimated rigid-body transformation parameters were applied to all calculated maps to before evaluating the quantitative metrics, see below. To ensure that the MPM data at the first and second time points had similar variability, a sinc-interpolation of 9th order was used to resample the data. Then, the deformation field that maps the data into MNI space was estimated using the DARTEL SPM software (Ashburner 2007). To this end, individual mean MT maps (across time points) were produced and segmented into gray and white matter probability maps (Ashburner and Friston 2005). Then, these MT white matter and gray matter segments were used as input in DARTEL to estimate the deformation fields using an existing template from a previous study (Mohammadi et al. 2015). The estimated DARTEL fields were applied to the regions of interests (ROIs) defined in the next paragraph. The MPM data remains in subject-space ensuring a low spatial correlation required by the adaptive smoothing method.

3.2.3 Regions of interest (ROI) analyses

Average values of MPM metrics were calculated within specific ROIs in subject space. The subject-specific ROIs were obtained by applying the backwards transformation estimated from the DARTEL-derived deformation field on atlas-based ROIs in gray and white matter: for each hemisphere the gray matter ROIs consisted of 56 regions extracted from the Harvard-Oxford atlas (Frazier et al. 2005; Makris et al. 2006; Desikan et al. 2006; Goldstein et al. 2007) and the white matter ROIs consisted of 13 regions extracted from the SPM Anatomy toolbox (Eickhoff et al. 2005). Each atlas ROI was thresholded at > 50 %, and additionally masked by the individual white or gray matter segments in native space (thresholded at probability < 0.5 as well). In total this yielded 102 gray matter and 10 white matter ROIs with sufficient numbers of voxel for the quantitative analysis.

3.2.4 Estimation of ESTATICS parameters and calculation of quantitative maps

For a given voxel, let $S^1_{\mathrm{obs}}, S^2_{\mathrm{obs}}, \ldots, S^n_{\mathrm{obs}}$ be given measurements for the different multi-echo sequences obtained at echo times $TE_1, TE_2 \ldots TE_n$. Given the measurements we aim at estimating the parameters $S_{T1}, S_{PD}, S_{MT}, R_2^\star$ of the ESTATICS model (5) by solving the non-negative least-squares problem

$$\min_{S_{T1},S_{PD},S_{MT},R_2^\star \geq 0} \frac{1}{2} \sum_{k=1}^{n} \|\vec{S}(S_{T1}, S_{PD}, S_{MT}, R_2^\star; TE_k) - \vec{S}^k_{\mathrm{obs}}\|^2. \tag{9}$$

Here, the non-negativity constraints are imposed to enforce plausible solutions, i.e., non-negative intensities that decay with TE; see (5). The optimization problem (9) is low-dimensional since only four parameters need to be estimated, however, one problem instance is required for each voxel. Since all instances are independent of one another the process is accelerated by performing computations in parallel.

To deal with the non-negativity constraints, we use the projected Gauss-Newton method presented in (Haber 2015). The key idea is to solve the problem approximately by performing a Gauss-Newton step on the inactive components (i.e., where entries are strictly positive) and a projected steepest descent step on the active components. A projected Armijo line search is performed to ensure the objective function decreases.

The adaptive smoothing procedure requires an estimate for the local covariance matrix Σ of the estimates. We determined Σ from the Jacobian and the residuals of the ESTATICS model. As the residual variance can be expected to be a spatially smooth and slowly varying function, we stabilized the estimates of Σ by a local mean with a small bandwidth. This reduced the local variability of the estimate $\hat{\Sigma}$.

3.2.5 Choice of parameters for the adaptive smoothing method

From $\vec{S}^{(0)}$ we calculated the unsmoothed maps of $R_2^\star$, R_1, A, and δ using the analytic formulae in Eqs (21), (22), and (23) from Appendix A. Furthermore, we used $\vec{S}^{(0)}$ and $\hat{\Sigma}$ as input for the adaptive smoothing method introduced in this paper, see Section 2.3 and Appendix B. From the resulting smoothed ESTATICS maps $\vec{S}^{(k^\star)}$ we calculated (implicitly) smoothed maps of $R_2^\star$, R_1, A, and δ using the analytic formulae from Appendix A.

The adaptive smoothing method depends on a number of parameters. The two kernel functions K_{loc} and K_{st} were chosen to be computationally cheap, i.e., the Epanechnikov kernel function for K_{loc} and a Plateau kernel function for K_{st}, cf. with the adaptive smoothing method msPOAS for diffusion MRI data (Becker et al. 2014).

The adaptation bandwidth λ is the main parameter of the smoothing procedure. Specifically, $\lambda = 0$ will leave the input data unchanged, while $\lambda = \infty$ results in a (classical) non-adaptive kernel estimate with kernel K_{loc} and bandwidth $h_{k^\star}$. For optimal adaptation, λ can be chosen by a propagation condition by simulation (Becker and Mathé 2013), which here yielded $\lambda = 12$, see also the discussion section.

The sequence h_k of bandwidths is chosen such that the variance reduction for a non-adaptive smoothing from each iteration step is 25%. The specific value is not very important, cf. Li et al. 2012, it has turned out to be a good compromise between sufficient increase of variance reduction between steps and careful increase of h_k in order to obtain sufficiently neat coverage of the scale space.

The number of iteration steps $k^\star$ determines the final bandwidth $h_{k^\star}$ and thus the maximum smoothness within homogeneous regions. We used $k^\star = 12$ ($h_{k^\star} = 1.69$) with dataset II for a compromise between sufficient variance reduction on the one hand and computational costs and a sufficiently smooth result on the other hand. With the ultra-high resolution dataset I we used $k^\star = 16$ ($h_{k^\star} = 2.33$) to compensate for the lower SNR. See the discussion section for further elaboration of the rationale behind a good parameter choice.

3.2.6 Correction for RF-inhomogeneities and imperfect spoiling

The local flip angle deviates from its nominal value due to inhomogeneities of the transmitted RF-field (Lutti et al. 2010); this can be locally corrected using a correction field $f_T(x)$ from consecutive pairs of spin-echo/stimulated spin-echo images (Lutti et al. 2010; Lutti et al. 2012)

$$\alpha(x) = f_T(x) \cdot \alpha \tag{10}$$

before calculating R_1, A, and δ from $\hat{S}^{(0)}_{T1}$, $\hat{S}^{(0)}_{MT}$, and $\hat{S}^{(0)}_{PD}$ using the analytic formulas given in Appendix A.

The description of the FLASH signal by the Ernst equation is only valid for perfectly spoiled sequencers with no signal due to residual transverse coherences. In practice, we require an additional correction for imperfect spoiling when estimating R_1 (or T_1) in addition to the one related to the local flip angle correction above, see Preibisch and Deichmann 2009. We formulate the (equivalent) correction for R_1 here:

$$R_1^c = \frac{R_1^{uc}}{P_a(f_T) \cdot R_1^{uc} + P_b(f_T)} \tag{11}$$

where $P_a(f_T)$ and $P_b(f_T)$ are quadratic functions in f_T with coefficients that depend on the specific sequence settings (timings, RF and gradient spoiling) and can be determined by simulation (Preibisch and Deichmann 2009; Yarnykh 2012; Weiskopf et al. 2013). The correction has been applied to all data from acquisition II.

Finally, denote α_{sat} the off-resonance Gaussian shaped RF pulse of the MTw sequence. It is expected that the magnetization transfer δ has a α_{sat}^2-dependence and cancels the bias due to the excitation flip angle bias and R_1 bias. However, the transmitted RF-field $f_T(x)$ leads to a local correction of α_{sat}, too, and thus to a residual bias field for the estimated δ. In Helms 2015 the corresponding correction factor for δ was found to be

$$\frac{1 - 0.4}{(1 - 0.4 \cdot f_T) \cdot f_T^2} \tag{12}$$

with the typically used value $\alpha_{sat} = 220^\circ$ in MPM sequences (Draganski et al. 2011; Weiskopf et al. 2013).

3.2.7 On the determination of the proton density (PD)

The receive field bias on the A maps is estimated similarly to the UNICORT approach (Weiskopf et al. 2011). Accurate bias-field correction requires the use of a brain mask and therefore image segmentation, which is based on the MT map due to its higher contrast in basal ganglia regions (Helms et al. 2009). The bias-corrected A map was converted into a PD map by calibrating the mean PD in white matter to 69% (Tofts 2005).

3.2.8 Analysis

We analyzed the properties of the proposed method using averaged metrics from a ROI analysis of dataset II. Doing so, the influence of the anatomical variance is largely reduced.

Spatial variability. To assess the spatial variability after denoising, the relative coefficient of variation (CoV) is calculated for each MPM metric $q = R_2^\star, MT, R_1, PD$ and for each of the different data processing approaches $m = orig, ad, nad, ref$ ($orig$ - original data from first time point, ad - adaptively smoothed data, nad - non-adaptively smoothed data, and ref - average across the two time points, which serves as reference noise reduction). To this end, the standard deviation was calculated over all voxels within each ROI $j = 1, \ldots, N_R$, for each subject $i = 1, \ldots, N$, and for each time point $t = 1, 2$ and averaged across these separately for gray and white matter:

$$msd(q, m) = \frac{1}{2} \sum_{t=1}^{2} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N_R} \sum_{j=1}^{N_R} sd_{ROI}[q(j, i, t, m)]. \tag{13}$$

Finally, the reduction of the spatial variability was assessed by the relative change of the metric $msd(q, m)$ with respect to $msd(q, orig)$ calculated in percent:

$$rsd(q, m) = 100 \times \frac{msd(q, orig) - msd(q, m)}{msd(q, orig)} \tag{14}$$

N_R is the number of ROIs ($N_R = 10$ for white matter, and $N_R = 102$ for gray matter), N the number of subjects ($N = 12$). This metric is used in Fig. 4.

There are two quantities contributing to $msd(q, m)$: variability of q due to random measurement errors and variation of q within each ROI (anatomic variability) due to microstructure. An efficient denoising algorithm should reduce (random) variability of q but leave anatomic variability unchanged. In case of high anatomic variability such a procedure would show only small changes in $rvar$.

Temporal variability. To assess the temporal variability for each MPM metric $q = R_2^\star, MT, R_1, PD$ we averaged the relative change of q between the two time points over all voxel $k = 1, \ldots N_V(j)$ within each ROI $j = 1, \ldots, N_R$, and for each subject $i = 1, \ldots, N$ assuming that the metric within the ROI is summarized by the mean:

$$raq(q, m) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N_R} \sum_{j=1}^{N_R} \frac{1}{N_V(j)} \sum_{k=1}^{N_V(j)} \frac{2 \cdot |q(j, i, t_1, m, k) - q(j, i, t_2, m, k)|}{\mathrm{mean}_{ROI}[q(j, i, t_1, m)] + \mathrm{mean}_{ROI}[q(j, i, t_2, m)]}, \tag{15}$$

This metric is used in Fig. 5.

Relative bias. To estimate the relative bias $rbias_q$ for each MPM metric $q = R_2^\star, MT, R_1, PD$, first the averaged value was calculated within each ROI $j = 1, \ldots, N_R$. Then, relative differences with respect to the value using the original data ($m = orig$) and finally its average across time, ROI and subject was calculated:

$$rbias(q, m) = \frac{1}{2} \sum_{t=1}^{2} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N_R} \sum_{j=1}^{N_R} \frac{\mathrm{mean}_{ROI}[q(j, i, t, orig)] - \mathrm{mean}_{ROI}[q(j, i, t, m)]}{\mathrm{mean}_{ROI}[q(j, i, t, orig)]}, \tag{16}$$

This metric is used in Fig. 6.

4 Results

4.1 Data at 500 μm resolution

This dataset features an extra-ordinary high resolution at a very low SNR. Qualitative assessment of the the MPM metrics after adaptive denoising a single data acquisition in comparison with the calculated MPM metrics from all three repeated measurements reveals the potential of the smoothing method to achieve significant variance reduction without blurring structural borders in the MPM metrics, see Fig. 2. For comparison we also show the resulting MPM metrics after adaptively smoothing all three acquired data repetitions.

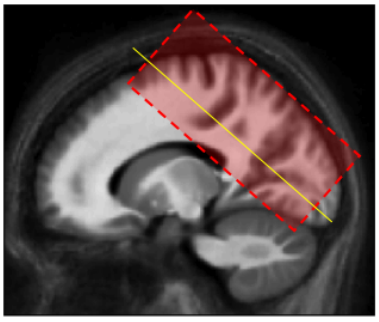
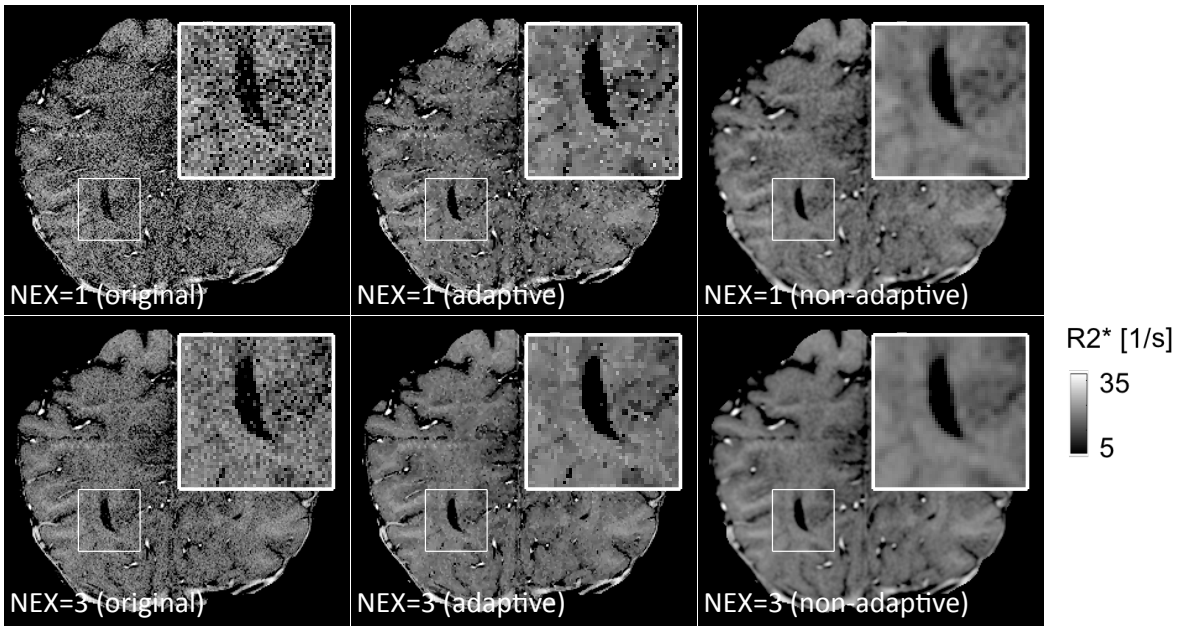
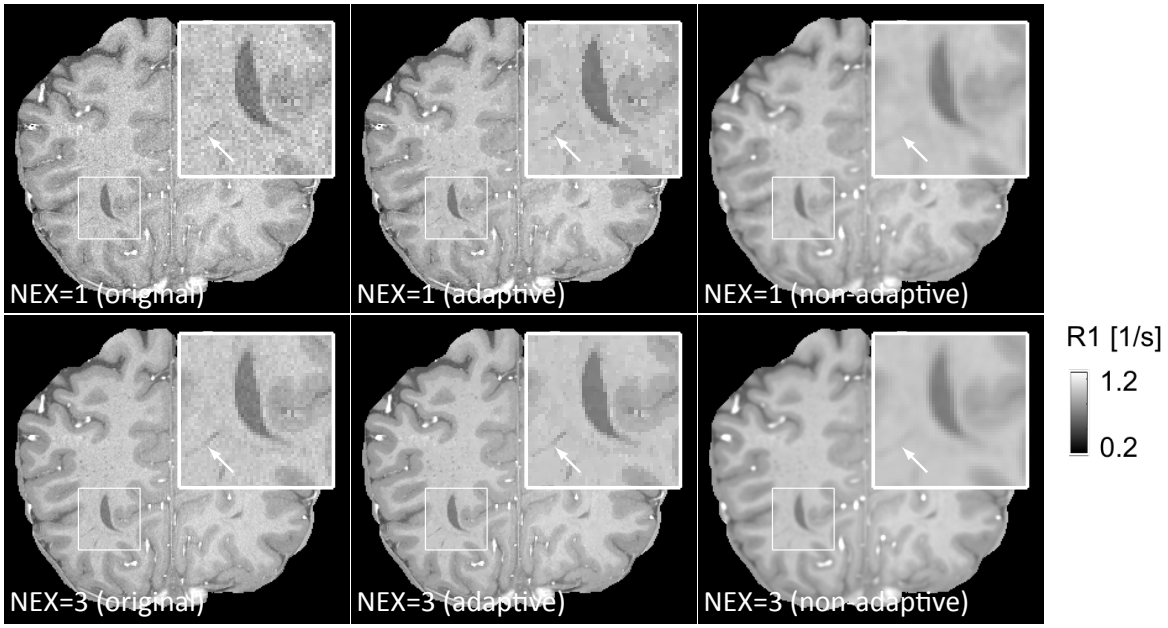


Figure 1 – *Localization of the 500micron dataset within the brain. The axial slice in Figs.2a-c is depicted as a yellow line.*



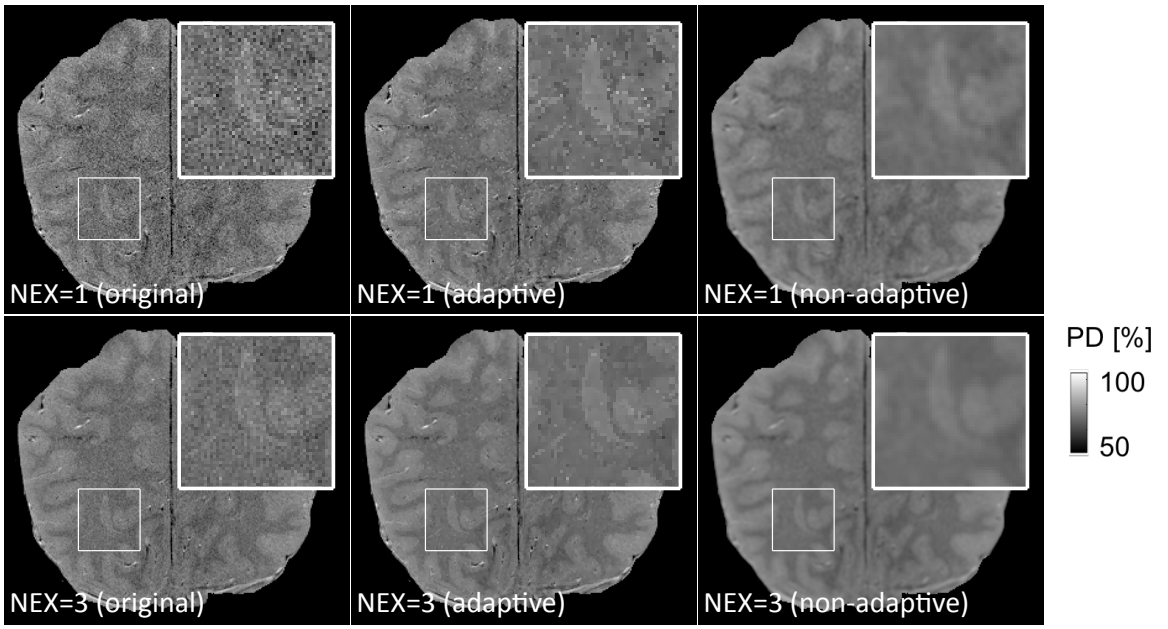
(a) – *We show the apparent transverse relaxation rate ($R_2^\star$) for a single oblique slice using a single repetition of the data (upper), and all three repetitions of the data (lower). From left to right: metric calculated from original data, adaptively smoothed data, and non-adaptively smoothed data. The zoomed region includes white matter pathways towards the occipital lobe that surround the left posterior horn of the lateral ventricles.*

Figure 2 – *Qualitative assessment of denoising on MPM for a 500micron dataset.*



(b) – *As before, but for the longitudinal relaxation rate (R_1). Arrows highlight a vein towards the ventricels which shows up well with adaptive smoothing but is blurred by the non-adaptive method.*

Figure 2 – *cont.*



(c) – *As before, but for the proton density (PD).*

Figure 2 – *cont.*

4.2 Data at 800 μm resolution

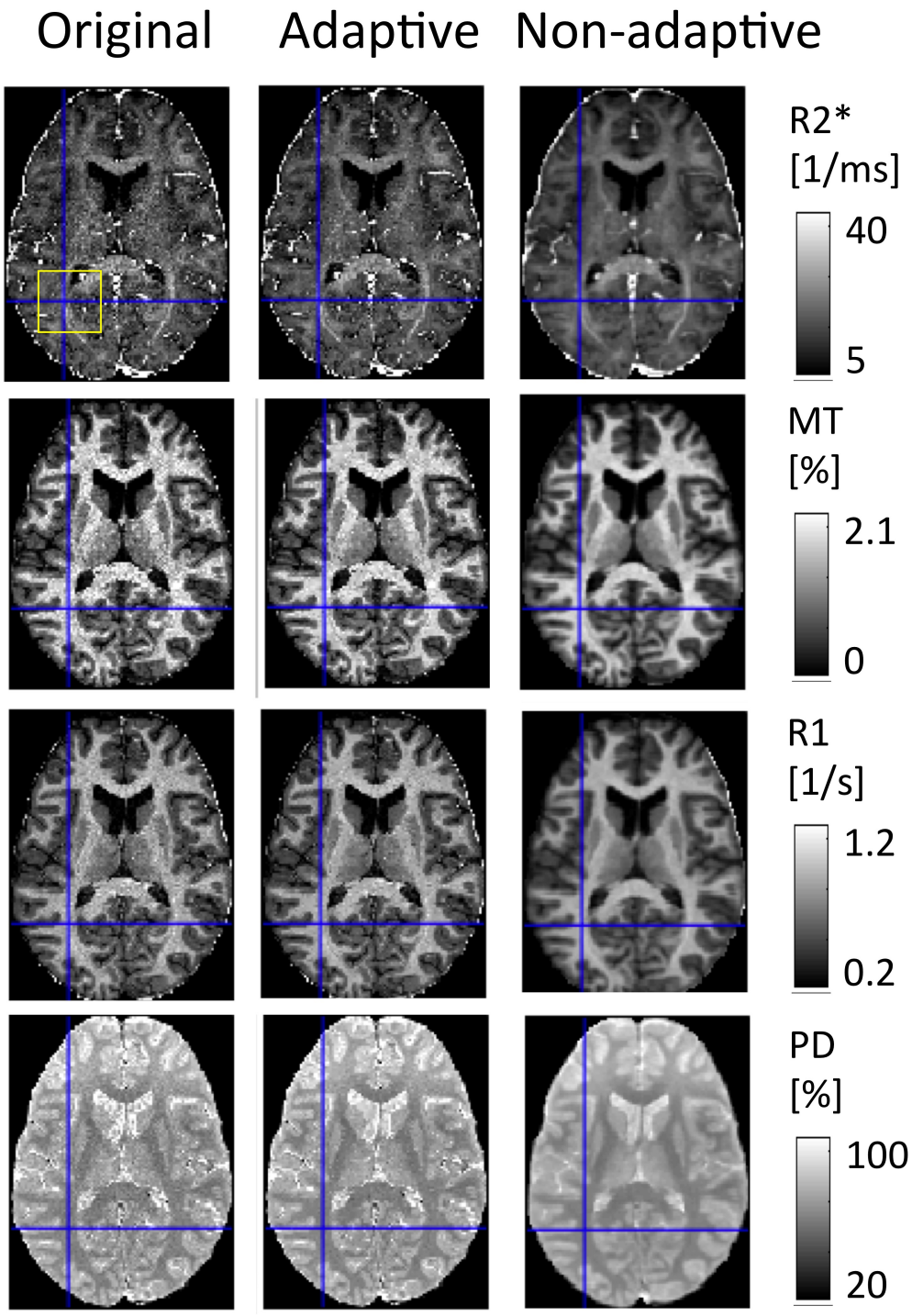
Here, we quantitatively analyze a dataset at a resolution that is feasible for group studies. To this end, we provide assessment of metrics to show the improvements by the proposed method for statistical analysis.

Qualitative assessment. Qualitative assessment of MPM metrics after different processing approaches (i.e., using original, adaptive, non-adaptive denoised data) demonstrated that the adaptive denoising reduces noise and keeps the contrast at tissue boundaries whereas non-adaptive denoising leads to blurring (see Figs. 2 and 3). Moreover, adaptive noise reduction preserves boundaries within a tissue class. Two examples are depicted: (1) in Figure 2b a vein towards the ventricels is highlighted, which is clearly visible in the original and the adaptive denoised image but becomes invisible in the non-adaptive denoised image, (2) in Figure 3b (top left) the border between two adjacent fibers, the optic radiation (OR) and the splenium of the corpus callosum (forceps major, FM), remains clearly visible after adaptive denoising but becomes less sharp after non-adaptive denoising.

Reduction of spatial variability. Relative reduction in variability in terms of the metric $rvar$ is in the order of $20 - 50\%$ for non-adaptive smoothing with the utilized bandwidth $h_{k^\star} = 1.63$ at $k^\star = 12$, see Fig. 4. For a spatially fully homogeneous and uncorrelated situation the non-adaptive procedure with the utilized kernel function would lead to a noise reduction of 74%. The fact that we observed less reduction in practice is due to anatomic variability within ROI's. The results for non-adaptive smoothing indicate, e.g., that anatomic variability is larger in GM compared to WM largest in grey matter MT maps. In general, non-adaptive smoothing reduces both noise variance, as well as to a smaller extend anatomic variability. In contrast, the adaptive procedure effectively restricts smoothing to local vicinities with limited anatomic variability. Noice variance reduction is therefore smaller, while anatomic variability is almost left untouched. This results in smaller values of $rvar$ that reflect reduction in noise standard deviation relative to the joint random and structural effect. Results are provided in Fig. 4. Furthermore, it depended on the MPM metric, e.g., for adaptive denoising the effectiveness of variance reduction was up to a factor of 3 larger in the $R_2^\star$ as compared to the R_1 metric (for gray matter).

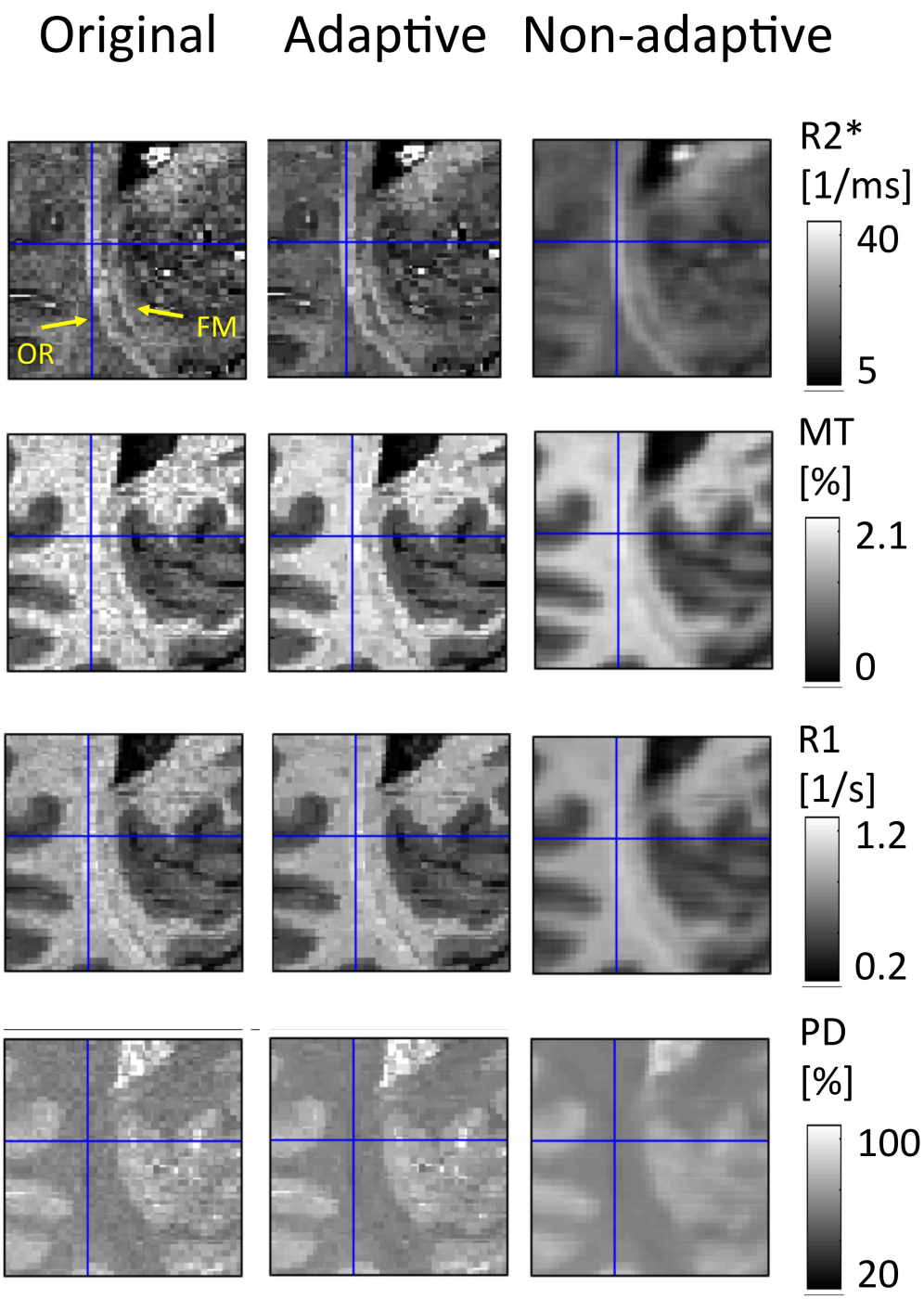
Reduction of temporal variability for longitudinal analysis. The scan-rescan variability after denoising was reduced (Fig. 5). In general, the reduction was larger after non-adaptive denoising than after adaptive.

Reduction of relative bias. The bias after denoising was negligible for the adaptive approach ($\ll 1\%$) but relevant for the non-adaptive approach (-1 and 12%) (Figure 6). It varied in directionality and magnitude between MPM metrics. It was negative for $R_2^\star$ in gray matter and PD in white matter, but positive elsewhere. It was largest in MT.



(a) – *Transverse slice across the whole brain. The MPM metrics include: apparent transverse relaxation rate ($R_2^\star$), magnetization transfer rate (MT), longitudinal relaxation rate (R_1), and proton density (PD). The yellow box in the top left is magnified in (b).*

Figure 3 – *Qualitative assessment of denoising on MPM metrics for subject 3.*



(b) – *Magnification of brain region includes fibers of the optic radiation (OR) and the splenium of the corpus callosum (forceps major, FM) - see arrows in top left. Non-adaptive denoising reduces noise most efficiently but at the same time can reduce effective resolution, e.g. the boundary between the OR and FM tract is blurred out. Adaptive denoising removes noise and keeps the effective resolution at the same time (boundary between adjacent tracts is preserved).*

Figure 3 – *cont.*

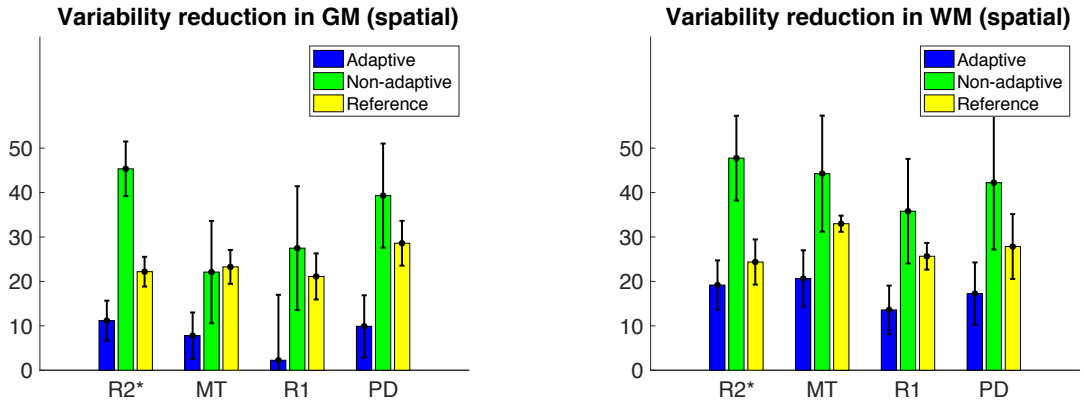


Figure 4 – *Relative reduction of the variability in the processed MPM metrics as compared to the original maps, depicted for gray (left) and white (right) matter. The processed MPM metrics underwent: adaptive denoising (blue), non-adaptive smoothing (green), and averaging of the two acquired time points (yellow), where the latter served as reference data. Values reflect a reduction in noise and anatomic variability. Adaptive denoising shows a smaller effect since it limits noise reduction to anatomically homogeneous vicinities and is supposed to not affect anatomic variability.*

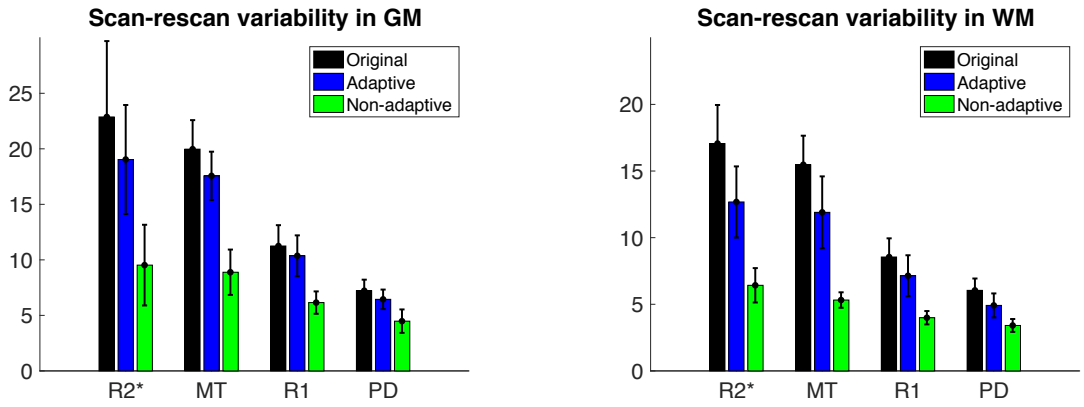


Figure 5 – *Scan-rescan variability (in percent) with respect to the average value for each MPM metric using different processing approaches, depicted in gray (left) and white (right) matter. The processing approaches included: using the original data (blue), adaptive denoising (green), and non-adaptive denoising (yellow). The scan-resan variability was reduced after adaptive denoising, and even more, for the same reason as for the spatial variability after non-adaptive denoising. Error bars indicate variability of the reduction over subjects.*

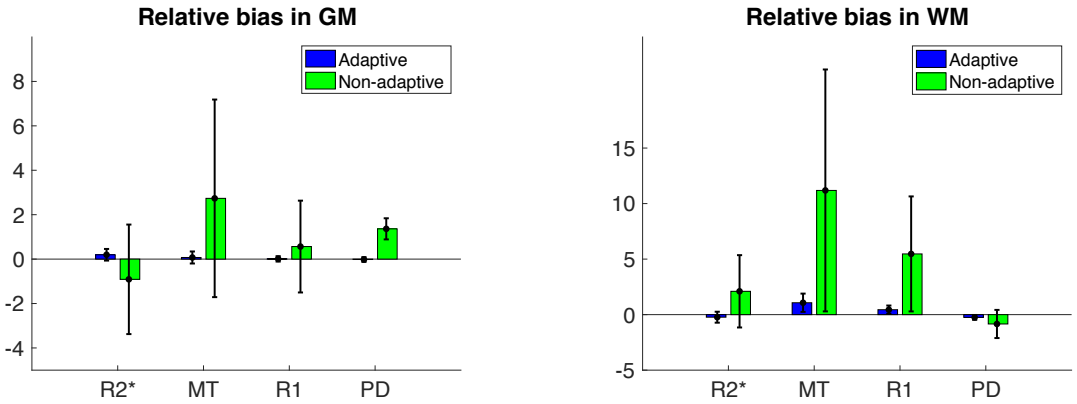


Figure 6 – *Relative bias (in percent) in the processed MPM metrics as compared to the original MPM metrics, depicted for gray (left) and white (right) matter. The processed MPM metrics underwent: adaptive denoising (blue) and non-adaptive denoising (green). The adaptively smoothed data showed a smaller bias (<1%) than the non-adaptively denoised data (up to 12%). Error bars indicate variability of the reduction over subjects.*

5 Discussion

In this paper we developed the (to our best knowledge) first adaptive smoothing method that simultaneously denoises all four qMRI maps obtained by the Multi-Parameter Mapping (MPM) framework: the proton density (PD), the longitudinal relaxation rate (R_1), the apparent transverse relaxation rate ($R_2^\star$) and, optionally, the magnetization transfer saturation (MT). Moreover, we derived a novel analytical relation between the FLASH signal from the multi-echo sequences and the MPM metrics R_1, PD, MT, and $R_2^\star$. Besides higher precision, the analytic relation directly reveals an experimental constraint on the acquisition protocol, namely that the TRs should be kept the same for all three contrasts. Our main findings are that this adaptive denoising method reduces spatial and temporal variability (up to 20% for the chosen bandwidth $h_{k^\star}$) while marginally biasing the quantitative MPM metrics (less than 1%).

Adaptive densoising in quantitative MRI. Today, adaptive denoising methods in anatomical imaging are either applied to single weighted anatomical images to improve the gray-to-white matter contrast for voxel-based morphometry studies analyses (Yi et al. 2015) or in diffusion MRI, see the papers referred to in the introduction, whereas in quantitative relaxometry or magnetization transfer imaging denoising approaches are not commonly used. One reason is that denoising can bias the quantitative outcome measures, if the smoothing kernel mixes tissue compartments with different MRI properties (Draganski et al. 2011). The simultaneous denoising of all four MPM metrics that is used in our approach can help to reduce this kind of bias due to the following reasons. First, the precision of the structure-adaptive smoothing kernel benefits from the joint SNR of the four MPM maps. Second, and even more importantly, homogeneous regions within tissue classes (e.g. due to similar microstructural composition of fiber pathways (Kleinnijenhuis et al. 2012)) can be better delineated because the information of four qMRI maps with, in parts, complementary sensitivity to microstructure (Callaghan et al. 2014; Callaghan et al. 2015), is combined in this approach. This notion was supported by our findings. We found that our adaptive denoising method produced less than 1% bias, whereas the bias introduced by non-adaptive denoising was up to 12% (Fig. 6). As a consequence our method leaves, e.g., the gray-white matter contrast almost unchanged, whereas the contrast would be systematically reduced by non-adaptive denoising.

Spatial analysis. Sensitivity to cross-sectional voxel-by-voxel differences will be affected by spatial variability due to noise. Spatial variability can be either due to microstructural differences in the MPM metrics that vary between anatomical regions or due to noise. Ideally adaptive denoising should only reduce the latter kind of variability. To reduce the influence of anatomical variability, we assessed the variability within atlas-defined region of interests (ROIs) in the gray and white matter, where the anatomical variability was assumed to be reduced because of similar microscopic tissue composition. We found that non-adaptive smoothing leads to a reduction in spatial variability of up to $20 - 50\%$ compared to a theoretical 74% in case of structural homogeneity and the utilized kernel and bandwidth. This indicates the presence of structure and the reduction of both noise and anatomic variability by non-adaptive smoothing.

Adaptive denoising is restricted to anatomically homogeneous regions due to the presence of the second kernel function K_{st} in the definition of weights $w_{ij}^{(k)}$, at a given bandwidth $h_{k^\star}$. Variance reduction for the adaptive method is therefore less than for non-adaptive smoothing employing the same bandwidth. Anatomic variability is almost untouched by construction. Both effects lead to an efficient denoising and a much smaller value in $rvar$ for the adaptive procedure.

Longitudinal analysis. Sensitivity to longitudinal voxel-by-voxel differences will be affected by temporal variability due to noise. Here, we assessed the temporal variability of MPM metrics. We found that denoising reduces temporal variability as well (Fig. 5). In accordance to our findings for spatial variability, also temporal variability was more strongly reduced when using non-adaptive denoising than adaptive denoising. We also observed that adaptive denoising worked more efficiently in white than in gray matter. While the main source of spatial variability thermal noise and thus should be efficiently reduced by denoising methods, it was, on the first glance, surprising to find reduction of temporal noise due to denoising, because temporal noise is more associated with instrumental noise, e.g., change in the receive-field due to repositioning or temporal variability in the MR systems' hardware. Our hypothesis for this finding is that the quantitative MPM data were corrected for systematic bias due to instrumentally induced temporal variation, thus main source of temporal variability was random noise or residual misalignment between the two time points. Both of which can be reduced by spatial denoising.

Mathematical considerations The procedure is based on the propagation separation approach which a) defines for each voxel an adaptive weighting scheme $w_{ij}^{(k)}$ by local comparison of the vector of ESTATICS parameters taking the covariance estimate into account and b) repeating this definition in an iterative inspection of the scale space at a sequence of bandwidths h_k.

It is worth noting that in the absence of an MTw sequence in the data, the method can still be applied with only minor changes: The S_{MT} related term in the ESTATICS model Eq. (5) is then missing, the smoothing procedure is only performed on $\hat{S}_{T1}^{(0)}$, $\hat{S}_{PD}^{(0)}$, and $\hat{R}_2^{\star(0)}$. This has been used in this paper for the 500 μm data, where no MTw data was available. Furthermore, the workflow of the method is designed such that setting $k = 0$ implies that no smoothing is applied to the ESTATICS parameters.

Note also that the non-adaptive smoothing of the estimated covariance maps $\hat{\Sigma}$ we used in this paper may contribute to the residual bias of the adaptive smoothing method, as it might blur discontinuities in the component maps of $\hat{\Sigma}$. However, the benefit of this approach is a largely reduced local variability of those estimates which can be justified by their overall smoothness.

In contrast to former approaches (Helms et al. 2008) the analytic derivation in Appendix A allows for the calculation of the MPM metrics from the data using the Ernst equation (1) without making linear approximations of the involved terms for small flip angle α and short TR.

Moreover, the estimation of the ESTATICS model parameters $\hat{S}_{T1}$, $\hat{S}_{PD}$, $\hat{S}_{MT}$, and $\hat{R}_2^{\star}$ before analytically calculating the MPM metrics has further advantages over direct estimation of the metrics from the Ernst equation: the positivity constraints on the parameters can be easily incorporated, the objective function for the optimization used to estimate the parameters is computationally cheaper, and the ESTATICS parametrization has a low parameter induced non-linearity (which can be easily checked using the methods from (Bates and Watts 1980)) such that the parameter estimates are approximately Gaussian and a suitable estimate of the covariance matrix can be obtained.

Methodological considerations and outlook. Obviously, the sequential nature of the data acquisition poses potential problems for the smoothing method in this paper. If, e.g., motion artifacts are not sufficiently corrected by registration of the image volumes, the estimation of the ESTATICS parameters may be biased. While this is not specific to the method in this paper but a general feature of any sequential data modeling, it can here potentially lead to a large value in the statistical penalty in the adaptive weights and hence less smoothing than desirable. In this multi-echo FLASH acquisition there are two different times scales at which motion artifacts can manifest in the data: (a) the intra-volume

time lag, (b) the inter-volume time lag between the different contrasts in the order of several minutes. In the MPM framework, the latter kind of motion is reduced by performing rigid body transformations between reference images from each echo train (e.g. the root-mean-square across the echoes).

The smoothing method outlined here, has a number of parameters, the influence of which we now discuss. λ is the main parameter of the procedure and determines the amount of adaptivity: If $\lambda = \infty$ the weighting scheme $w_{ij}^{(k)}$ is identical to a non-adaptive one determined by a bandwidth h_k in the kernel function K_{loc}. Then, the procedure is identical to a non-adaptive kernel estimator. On the other hand, for $\lambda = 0$ the weighting scheme will always include only the central voxel in the weighted average. Thus, the data will not be changed at all by the procedure. A good choice for λ lies between these extremes: The strategy to choose the adaptation parameter λ independent from the data is, e.g., described in Becker and Mathé 2013. Other parameters of the procedure like the kernel functions K_{loc} and K_{st} have only minor impact on the smoothing result, cf. Becker et al. 2014. The same applies to the specific choice of the sequence of bandwidth, cf. Li et al. 2012.

The usage of the assumption of the procedure, i.e., a local homogeneity structure of the ESTATICS parameters, results in the appearance of a step function in the metric maps for very large $k^\star$. However, in our experience the usage of an intermediate number $k^\star = 12 - 16$ of iteration steps like the one used in this paper avoids this effect in favor of a smooth result in homogeneous regions (Becker and Mathé 2013).

The deviation of the signal distribution from a Gaussian at very low SNR poses the same additional bias problem for the estimates of the ESTATICS model and the quantitative metrics R_1, PD, MT, and $R_2^\star$ by the non-linear estimation problem in (9) as has been observed, e.g., for parameters of diffusion models in dMRI (Polzehl and Tabelow 2016). The consideration of an alternative quasi-likelihood estimation for the parameters is capable of reducing this bias also in MPM acquisitions, see Tabelow et al. 2017 for a basic outline.

The workflow with a) the estimation of the ESTATICS parameter maps, b) smoothing of these, and c) the calculation of the MPM metric from the (smoothed) parameter maps is suited for the application of any smoothing procedure for $\hat{S}_{T1}^{(0)}$, $\hat{S}_{PD}^{(0)}$, $\hat{S}_{MT}^{(0)}$, and $\hat{R}_2^{\star(0)}$ based on different approaches than propagation-separation, e.g., anisotropic diffusion, non-local means, or others, see the list of adaptive smoothing methods for diffusion MRI data in the introduction. Since our method developed here is the first adaptive smoothing approach that is applied MPM data, this will be part of future research.

We also note, that instead of using the adaptive weighting schemes to smooth the ESTATICS parameter maps, we could apply the weights directly to the original volumes and then re-estimate the maps. This procedure would benefit from the dimension reduction to determine the adaptive weights and still act on the original data. Furthermore this would also address the bias in estimating $R_2^\star$ that is induced by the non-linear part of the ESTATICS model. However, this would require much more computing resources and we do not expect a large difference here in view of the approximate linearity of the ESTATICS model.

6 Conclusion

Mathematical methods for adaptively utilizing the local structure of the data at hand alongside with hardware improvements and optimized sequences may pave the way for higher resolutions in qMRI by substantially increasing the SNR. Here, we introduced a new noise reduction method for qMRI data acquired in the MPM framework. The procedure does not mask the microstructure observable at the chosen resolution by blurring the tissue borders. It uses the available information from all acquired

image volumes of the MPM data for the definition of its locally adaptive smoothing kernel. The method is computationally efficient and can readily be applied. It is made freely available as a toolbox for SPM or an R package on request.

7 Acknowledgments

SM received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 658589.

IE was supported by a Forschungsförderungsfond der Medizinischen Fakultät (FFM) Postdoctoral Fellowship from the University Medical Center Hamburg-Eppendorf. This project was also supported by the ERA-Net Neuron Cofund under the BMBF No 01EW1711A and the Deutsche Forschungsgemeinschaft (DFG) Grant MO 2397/4-1.

NW has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 616905.

The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 0915/Z/10/Z.

LR was supported in part by Emory University's Research Council and the US National Science foundation under award DMS 1522599.

Appendices

A Analytic formulas for R_1, A, and δ

Using the additional abbreviation $E_1 = e^{-R_1 \cdot TR}$ under the assumption, that TR is constant over all T1w and PDw volumes, we immediately get from Eqs. (1) and (5):

$$\begin{aligned} S^0_{T1} &= A \cdot \sin\alpha_{T1}\, \frac{1-E_1}{1-\cos\alpha_{T1}\cdot E_1} \\ S^0_{PD} &= A \cdot \sin\alpha_{PD}\, \frac{1-E_1}{1-\cos\alpha_{PD}\cdot E_1}, \end{aligned} \tag{17}$$

where we denote the corresponding flip angles by α_{T1} and α_{PD}. Simple re-ordering of terms yields:

$$S^0_{T1} - S^0_{T1}\cdot\cos\alpha_{T1}\cdot E_1 = A\cdot\sin\alpha_{T1} - A\cdot\sin\alpha_{T1}\cdot E_1 \tag{18}$$

$$S^0_{PD} - S^0_{PD}\cdot\cos\alpha_{PD}\cdot E_1 = A\cdot\sin\alpha_{PD} - A\cdot\sin\alpha_{PD}\cdot E_1. \tag{19}$$

From Eq. (19) we find by multiplication with $\frac{\sin\alpha_{T1}}{\sin\alpha_{PD}}$

$$S^0_{PD}\frac{\sin\alpha_{T1}}{\sin\alpha_{PD}} - S^0_{PD}\cdot\frac{\sin\alpha_{T1}}{\sin\alpha_{PD}}\cdot\cos\alpha_{PD}\cdot E_1 = A\cdot\sin\alpha_{T1} - A\cdot\sin\alpha_{T1}\cdot E_1,$$

where the right hand side can be replaced inserting Eq. (18):

$$S^0_{PD}\frac{\sin\alpha_{T1}}{\sin\alpha_{PD}} - S^0_{PD}\cdot\frac{\sin\alpha_{T1}}{\sin\alpha_{PD}}\cdot\cos\alpha_{PD}\cdot E_1 = S^0_{T1} - S^0_{T1}\cdot\cos\alpha_{T1}\cdot E_1.$$

Finally, evaluation of E_1 yields

$$E_1 = \frac{S^0_{T1} - S^0_{PD} \cdot \frac{\sin \alpha_{T1}}{\sin \alpha_{PD}}}{S^0_{T1} \cdot \cos \alpha_{T1} - S^0_{PD} \cdot \frac{\sin \alpha_{T1}}{\sin \alpha_{PD}} \cdot \cos \alpha_{PD}}, \tag{20}$$

from which we obtain

$$R_1 = -\ln E_1 / TR. \tag{21}$$

Inserting the result for E_1 into Eq. (17) yields

$$A = \frac{(1 - \cos \alpha_{T1} \cdot E_1) \cdot S^0_{T1}}{\sin \alpha_{T1} \cdot (1 - E_1)}. \tag{22}$$

Thus, given estimates for S^0_{T1} and S^0_{PD} the maps for R_1 and A can be simply calculated analytically.

If additionally MTw are acquired, we can also determine the MT saturation parameter δ. Using the abbreviation $E_2 = e^{-R_1 \cdot TR_2} = E_1^{TR_2/TR}$ we obtain from Eq. (3)

$$S^0_{MT} = A \cdot \sin \alpha_{MT} \, \frac{1 - E_1 - \delta \cdot (E_2 - E_1)}{1 - \cos \alpha_{MT} \cdot (1 - \delta) \cdot E_1}$$

and transposed to δ:

$$\delta = 1 - \cos \alpha_2 = 1 - \frac{S^0_{MT} - A \cdot \sin \alpha_{MT} \cdot (1 - E_2)}{S^0_{MT} \cdot \cos \alpha_{MT} \cdot E_1 + A \cdot \sin \alpha_{MT} \cdot (E_2 - E_1)}. \tag{23}$$

B Adaptive smoothing procedure

We denote the original estimates from the ESTATICS model by

$$\vec{S}^{(0)} = (\hat{S}^{(0)}_{T1}, \hat{S}^{(0)}_{PD}, \hat{S}^{(0)}_{MT}, \hat{R}^{\star(0)}_2)^\top.$$

We further assume that a local estimate $\hat{\Sigma}$ for the covariance of the parameters is available.

We define a sequence of bandwidths h_k for $k = 1, \ldots, k^\star$ for all $k^\star$ steps of the iterative procedure. Let K_{loc} and K_{st} be some suitable kernel functions (Fan and Gijbels 1996) on the unit interval. Finally, let λ denote the adaptation bandwidth, that controls the amount of adaptivity of the procedure.

For each iteration step $k = 1, \ldots, k^\star$ we define adaptive weights

$$w^{(k)}_{ij} = K_{loc}\left(|i - j|^2 / h_k^2\right) \cdot K_{st}\left(s^{(k)}_{ij} / \lambda\right) \tag{24}$$

for all pairwise voxel locations i and j with the statistical penalty $s^{(k)}_{ij}$ defined as

$$s^{(k)}_{ij} = N^{(k-1)}_i \cdot \left(\vec{S}^{(k-1)}_i - \vec{S}^{(k-1)}_j\right)^\top \hat{\Sigma}_i^{-1} \left(\vec{S}^{(k-1)}_i - \vec{S}^{(k-1)}_j\right) \tag{25}$$

based on the estimates $\vec{S}^{(k-1)}_i$ and $\vec{S}^{(k-1)}_j$ and the sum of weights $N^{(k-1)}_i = \sum_j w^{(k-1)}_{ij}$ from the previous step. For $k = 1$ use $N^{(0)}_i \equiv 1$. We finally calculate new estimates $\vec{S}^{(k)}$

$$\vec{S}^{(k)}_i = \sum_j w^{(k)}_{ij} \vec{S}^{(0)} / \sum_j w^{(k)}_{ij} \tag{26}$$

as a weighted average of the data using the local adaptive weights $w_{ij}^{(k)}$ and stop the iteration at $k = k^\star$ to obtain the final smoothed maps

$$\vec{S}^{(k^\star)} = \left(\hat{S}_{T1}^{(k^\star)}, \hat{S}_{PD}^{(k^\star)}, \hat{S}_{MT}^{(k^\star)}, \hat{R}_2^{\star(k^\star)}\right)^\top.$$

Data: ESTATICS parameters $S_{T1}^{(0)}, S_{PD}^{(0)}, S_{MT}^{(0)}, R_2^{\star(0)}$ and covariance matrix estimate $\hat{\Sigma}$
Initialization: $\vec{S}^{(0)} = (S_{T1}^{(0)}, S_{PD}^{(0)}, S_{MT}^{(0)}, R_2^{\star(0)})^\top$, choose sequence of bandwidths $h_1 < \cdots < h_{k^\star}$, choose number $k^\star$ of iterations, choose kernel functions K_{loc}, K_{st}, choose adaptation parameter λ, set $N_i^{(0)} \equiv 1$, set $k = 1$;
while $k \leq k^\star$ **do**
  $w_{ij}^{(k)} = K_{loc}\left(|i-j|^2/h_k^2\right) K_{st}\left(s_{ij}^{(k)}/\lambda\right)$ for all pairs of voxel i, j;
  with $s_{ij}^{(k)} = N_i^{(k-1)} \left(\vec{S}_i^{(k-1)} - \vec{S}_j^{(k-1)}\right)^\top \Sigma^{-1} \left(\vec{S}_i^{(k-1)} - \vec{S}_j^{(k-1)}\right)$;
  calculate $\vec{S}_i^{(k)} = \sum_j w_{ij}^{(k)} \vec{S}^{(0)} / \sum_j w_{ij}^{(k)}$;
  calculate $N_i^{(k)} = \sum_j w_{ij}^{(k)}$;
end
Result: Denoised maps $\vec{S}^{(k^\star)} = \left(S_{T1}^{(k^\star)}, S_{PD}^{(k^\star)}, S_{MT}^{(k^\star)}, R_2^{\star(k^\star)}\right)^\top$ of ESTATICS parameters

Algorithm 1: Summary of the adaptive smoothing developed in this paper.

References

Aja-Fernández, S., M. Niethammer, M. Kubicki, M. E. Shenton, and C.-F. Westin (2008). „Restoration of DWI Data Using a Rician LMMSE Estimator“. *IEEE Trans Med Imaging* 27, pp. 1389–1403.

Ashburner, J. (2007). „A fast diffeomorphic image registration algorithm“. *NeuroImage* 38.1, pp. 95–113.

Ashburner, J. and K. J. Friston (2005). „Unified segmentation“. *NeuroImage* 26.3, pp. 839–851.

Bates, D. M. and D. G. Watts (1980). „Relative Curvature Measures of Nonlinearity“. *Journal of the Royal Statistical Society. Series B* 42.1, pp. 1–25.

Becker, S. M. A., K. Tabelow, H. U. Voss, A. Anwander, R. M. Heidemann, and J. Polzehl (2012). „Position-orientation adaptive smoothing of diffusion weighted magnetic resonance data (POAS).“ eng. *Med Image Anal* 16.6, pp. 1142–1155.

Becker, S., K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl (2014). „Adaptive smoothing of multi-shell diffusion-weighted magnetic resonance data by msPOAS“. *NeuroImage* 95.1809, pp. 90–105.

Becker, S. M. A. and P. Mathé (2013). „A different perspective on the Propagation-Separation Approach“. *Electron J Statist* 7, pp. 2702–2736.

Callaghan, M. F., P. Freund, B. Draganski, E. Anderson, M. Cappelletti, R. Chowdhury, J. Diedrichsen, T. H. B. Fitzgerald, P. Smittenaar, G. Helms, A. Lutti, and N. Weiskopf (2014). „Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging.“ *Neurobiol Aging* 35.8, pp. 1862–1872.

Callaghan, M. F., G. Helms, A. Lutti, S. Mohammadi, and N. Weiskopf (2015). „A general linear relaxometry model of R1 using imaging data“. *Magn Reson Med* 73.3, pp. 1309–1314.

R Core Team (2017). R*: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org`.

Desikan, R. S., F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, et al. (2006). „An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest“. *NeuroImage* 31.3, pp. 968–980.

Ding, Z., J. Gore, and A. Anderson (2005). „Reduction of Noise in Diffusion Tensor Images Using Anisotropic Smoothing“. *Magn Reson Med* 53.2, pp. 485–490.

Draganski, B., J. Ashburner, C. Hutton, F. Kherif, R. Frackowiak, G. Helms, and N. Weiskopf (2011). „Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ)“. *NeuroImage* 55.4, pp. 1423–1434.

Duits, R. and E. Franken (2011). „Left-Invariant Diffusions on the Space of Positions and Orientations and their Application to Crossing-Preserving Smoothing of HARDI images“. *Int J Comput Vis* 92.3, pp. 231–264.

Edelstein, W., G. Glover, C. Hardy, and R. Redington (1986). „The Intrinsic Signal-To-Noise Ratio in NMR Imaging“. *Magn Reson Med* 3, pp. 604–618.

Eickhoff, S. B., K. E. Stephan, H. Mohlberg, C. Grefkes, G. R. Fink, K. Amunts, and K. Zilles (2005). „A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data“. *NeuroImage* 25.4, pp. 1325–1335.

Ellerbrock, I. and S. Mohammadi (2017). „Four in vivo g-ratio weighted imaging methods: comparability and repeatability at the group level“. submitted to Human Brain Mapping.

Ernst, R. R. and W. A. Anderson (1966). „Application of Fourier Transform Spectroscopy to Magnetic Resonance“. *Rev Sci Instrum* 37, pp. 93–102.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and its Applications*. Chapman & Hall, London.

Fletcher, P. T. (2004). „Statistical Variability in Nonlinear Spaces: Application to Shape Analysis and DT-MRI“. Doctoral dissertation. University of North Carolina at Chapel Hill.

Frazier, J. A., S. Chiu, J. L. Breeze, N. Makris, N. Lange, D. N. Kennedy, M. R. Herbert, E. K. Bent, V. K. Koneru, M. E. Dieterich, et al. (2005). „Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder“. *Am J Psychiatry* 162.7, pp. 1256–1265.

Goldstein, J. M., L. J. Seidman, N. Makris, T. Ahern, L. M. O'Brien, V. S. Caviness, D. N. Kennedy, S. V. Faraone, and M. T. Tsuang (2007). „Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability“. *Biol Psychiatry* 61.8, pp. 935–945.

Haber, E. (2015). *Computational methods in geophysical electromagnetics*. Mathematics in Industry (Philadelphia). Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Haldar, J. P., V. J. Wedeen, M. Nezamzadeh, G. Dai, M. W. Weiner, N. Schuff, and Z.-P. Liang (2013). „Improved diffusion imaging through SNR-enhancing joint reconstruction“. *Magn Reson Med* 69.1, pp. 277–289.

Helms, G. (2015). „Correction for residual effects of B1+ inhomogeniety on MT saturation in FLASH-based multi-parameter mapping of the brain“. In: *Proc Intl Soc Mag Reson Med*. Vol. 23.

Helms, G., H. Dathe, and P. Dechent (2008). „Quantitative FLASH MRI at 3T Using a Rational Approximation of the Ernst Equation“. *Magn Reson Med* 59, pp. 667–672.

Helms, G., H. Dathe, K. Kallenberg, and P. Dechent (2008). „High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI.“ *Magn Reson Med* 60.6, pp. 1396–1407.

Helms, G., B. Draganski, R. Frackowiak, J. Ashburner, and N. Weiskopf (2009). „Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps“. *NeuroImage* 47.1, pp. 194 –198.

Kleinnijenhuis, M., M. Barth, D. C. Alexander, A.-M.v. C. van Walsum, and D. G. Norris (2012). „Structure Tensor Informed Fiber Tractography (STIFT) by combining gradient echo MRI and diffusion weighted imaging“. *NeuroImage* 59.4, pp. 3941–3954.

Koenig, S. H., R. D. Brown, and R. Ugolini (1993). „A unified view of relaxation in protein solutions and tissue, including hydration and magnetization transfer“. *Magn Reson Med* 29.1, pp. 77–83.

Le Bihan, D. (2003). „Looking into the functional architecture of the brain with diffusion MRI.“ *Nat Rev Neurosci* 4.6, pp. 469–480.

Li, Y., J. H. Gilmore, J. Wang, M. Styner, W. Lin, and H. Zhu (2012). „TwinMARM: Two-Stage Multiscale Adaptive Regression Methods for Twin Neuroimaging Data“. *IEEE Trans Med Imaging* 31.5, pp. 1100–1112.

Lohmann, G., S. Bohn, K. Müller, R. Trampel, and R. Turner (2010). „Image restoration and spatial resolution in 7-tesla magnetic resonance imaging.“ eng. *Magn Reson Med* 64.1, pp. 15–22.

Lüsebrink, F., A. Sciarra, H. Mattern, R. Yakupov, and O. Speck (2017). „Erratum: T1-weighted in vivo human whole brain MRI dataset with an ultrahigh isotropic resolution of 250 μm“. *Sci Data* 4.

Lutti, A., C. Hutton, J. Finsterbusch, G. Helms, and N. Weiskopf (2010). „Optimization and validation of methods for mapping of the radiofrequency transmit field at 3T“. *Magn Reson Med* 64, pp. 229–238.

Lutti, A., J. Stadler, O. Josephs, C. Windischberger, O. Speck, J. Bernarding, C. Hutton, and N. Weiskopf (2012). „Robust and fast whole brain mapping of the RF transmit field B1 at 7T“. *PLoS ONE* 7.e32379.

Lutti, A., F. Dick, M. I. Sereno, and N. Weiskopf (2014). „Using high-resolution quantitative mapping of R1 as an index of cortical myelination.“ *NeuroImage* 93 Pt 2, pp. 176–188.

Makris, N., J. M. Goldstein, D. Kennedy, S. M. Hodge, V. S. Caviness, S. V. Faraone, M. T. Tsuang, and L. J. Seidman (2006). „Decreased volume of left and total anterior insular lobule in schizophrenia“. *Schizophr Res* 83.2, pp. 155–171.

Manjón, J. V., N. A. Thacker, J. J. Lull, G. Garcia-Martí, L. Martí-Bonmatí, and M. Robles (2009). „Multicomponent MR image denoising“. *Int J Biomed Imaging* 2009, p. 756897.

McGraw, T., B. Vemuri, E. Özarslan, Y. Chen, and T. Mareci (2009). „Variational denoising of diffusion weighted MRI“. *Inverse Probl Imaging* 3, pp. 625–648.

Mohammadi, S., D. Carey, F. Dick, J. Diedrichsen, M. I. Sereno, M. Reisert, M. F. Callaghan, and N. Weiskopf (2015). „Whole-brain in-vivo measurements of the axonal g-ratio in a group of 37 healthy volunteers“. *Front Neurosc* 9, p. 441.

Parker, G., J. Schnabel, M. Symms, D. Werring, and G. Barker (2000). „Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging“. *J Magn Reson Imaging* 11, pp. 702–710.

Polzehl, J. and V. Spokoiny (2006). „Propagation-separation approach for local likelihood estimation“. *Probab Theory Relat Fields* 135.3, pp. 335–362.

Polzehl, J. and K. Tabelow (2016). „Low SNR in diffusion MRI models“. *J Amer Statist Assoc* 111.516, pp. 1480–1490.

Preibisch, C. and R. Deichmann (2009). „Influence of RF spoiling on the stability and accuracy of T1 mapping based on spoiled FLASH with varying flip angles.“ *Magn Reson Med* 61.1, pp. 125–135.

Roebroeck, A., R. Galuske, E. Formisano, O. Chiry, H. Bratzke, I. Ronen, D.-s. Kim, and R. Goebel (2008). „High-resolution diffusion tensor imaging and tractography of the human optic chiasm at 9.4 T“. *NeuroImage* 39.1, pp. 157–168.

Setsompop, K., Q. Fan, J. Stockmann, B. Bilgic, S. Huang, S. F. Cauley, A. Nummenmaa, F. Wang, Y. Rathi, T. Witzel, et al. (2017). „High-resolution in vivo diffusion imaging of the human brain with generalized slice dithered enhanced resolution: Simultaneous multislice (gSlider-SMS)“. *Magn Reson Med*. URL: https://dx.doi.org/10.1002/mrm.26653.

Stikov, N., J. S. Campbell, T. Stroh, M. Lavelée, S. Frey, J. Novek, S. Nuara, M.-K. Ho, B. J. Bedell, R. F. Dougherty, et al. (2015). „In vivo histology of the myelin g-ratio with magnetic resonance imaging“. *NeuroImage* 118, pp. 397–405.

Tabelow, K., J. Polzehl, V. Spokoiny, and H. U. Voss (2008). „Diffusion Tensor Imaging: Structural adaptive smoothing“. *NeuroImage* 39, pp. 1763–1773.

Tabelow, K., C. D'Alonzo, L. Ruthotto, M. F. Callaghan, N. Weiskopf, J. Polzehl, and S. Mohammadi (2017). „Removing the estimation bias due to the noise floor in multi-parameter maps“. In: *Proc Intl Soc Mag Reson Med*. Vol. 25.

Tofts, P. (2005). *Quantitative MRI of the brain: Measuring changes caused by disease*. John Wiley & Sons.

Tristán-Vega, A. and S. Aja-Fernández (2010). „DWI filtering using joint information for DTI and HARDI“. *Med Image Anal* 14.2, pp. 205–218.

Weiskopf, N., A. Lutti, G. Helms, M. Novak, J. Ashburner, and C. Hutton (2011). „Unified segmentation based correction of R1 brain maps for RF transmit field inhomogeneities (UNICORT)“. *NeuroImage* 54.3, pp. 2116 –2124.

Weiskopf, N., J. Suckling, G. Williams, M. M. Correia, B. Inkster, R. Tait, C. Ooi, E. T. Bullmore, and A. Lutti (2013). „Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: a multi-center validation.“ *Front Neurosci* 7, p. 95.

Weiskopf, N., M. F. Callaghan, O. Josephs, A. Lutti, and S. Mohammadi (2014). „Estimating the apparent transverse relaxation time (R2(*)) from images with different contrasts (ESTATICS) reduces motion artifacts.“ *Front Neurosci* 8, p. 278.

Weiskopf, N., S. Mohammadi, A. Lutti, and M. F. Callaghan (2015). „Advances in MRI-based computational neuroanatomy: from morphometry to in-vivo histology“. *Curr Opin Neurol* 28.4, pp. 313–322.

Xu, J., S. Moeller, E. J. Auerbach, J. Strupp, S. M. Smith, D. A. Feinberg, E. Yacoub, and K. Uğurbil (2013). „Evaluation of slice accelerations using multiband echo planar imaging at 3T“. *NeuroImage* 83, pp. 991–1001.

Xu, Q., A. W. Anderson, J. C. Gore, and Z. Ding (2010). „Efficient anisotropic filtering of diffusion tensor images“. *Magn Reson Imaging* 28.2, pp. 200 –211.

Yarnykh, V. L. (2012). „Fast macromolecular proton fraction mapping from a single off-resonance magnetization transfer measurement“. *Magn Reson Med* 68.1, pp. 166–178.

Yi, D., Y. M. Choe, M. S. Byun, B. K. Sohn, E. H. Seo, J. Han, J. Park, J. I. Woo, and D. Y. Lee (2015). „Differences in functional brain connectivity alterations associated with cerebral amyloid deposition in amnestic mild cognitive impairment“. *Front Aging Neurosci* 7, p. 15.