Diffusion MRI noise mapping using random matrix theory

Jelle Veraart[1,2,*], **Els Fieremans**[1], and **Dmitry S. Novikov**[1]

[1] Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine New York, NY, USA

[2] iMinds-Vision Lab, Department of Physics, University of Antwerp, Antwerp, Belgium

Abstract

Purpose—To estimate the spatially varying noise map using a redundant magnitude MR series.

Methods—We exploit redundancy in non-Gaussian multi-directional diffusion MRI data by identifying its noise-only principal components, based on the theory of noisy covariance matrices. The bulk of PCA eigenvalues, arising due to noise, is described by the universal Marchenko-Pastur distribution, parameterized by the noise level. This allows us to estimate noise level in a local neighborhood based on the singular value decomposition of a matrix combining neighborhood voxels and diffusion directions.

Results—We present a model-independent local noise mapping method capable of estimating noise level down to about 1% error. In contrast to current state-of-the art techniques, the resultant noise maps do not show artifactual anatomical features that often reflect physiological noise, the presence of sharp edges, or a lack of adequate a priori knowledge of the expected form of MR signal.

Conclusions—Simulations and experiments show that typical diffusion MRI data exhibit sufficient redundancy that enables accurate, precise, and robust estimation of the local noise level by interpreting the PCA eigenspectrum in terms of the Marchenko-Pastur distribution.

Keywords

PCA; Marchenko-Pastur; Rician; diffusion MRI; noise

INTRODUCTION

The transition from "weighted" MR images to full-scale microstructural modeling is setting a much higher bar on estimating signal parameters in an unbiased way. This is especially relevant for diffusion MRI (dMRI) (1, 2), which holds unrivaled promise for non-invasive quantification of microstructural parameters. Unfortunately, the non-Gaussian nature of noise in magnitude MR data generally biases the diffusion parameters of interest (3). The correction for this noise-induced bias relies on an independent unbiased estimate for the noise map (4).

[*] Corresponding author: Jelle Veraart, Center for Biomedical Imaging, Department of Radiology, New York University Langone Medical Center, 660 First Avenue, New York, NY 10016, USA, Jelle.Veraart@nyumc.org, ☎ +1 212 263 4818, Twitter: @jelle_veraart.

Accurate, precise, and robust noise estimation has remained a challenging problem. The overarching reason is the non-Gaussian, typically non-central χ (nc-χ) noise statistics, that couples an unknown noise-free signal nonlinearly with the noise in the original complex-valued data from each coil element. This coupling is especially important at low signal-to-noise ratio (SNR), typical of dMRI acquisitions. Further complications arise from low spatial resolution, involuntary motion, and generally spatially varying noise pattern due to the use of parallel imaging (5).

So far, only a few noise estimation methods are able to deal with the spatially varying nature of noise (4, 6–11). The existing techniques depend on (i) the assumption of homogeneous signal intensities in a local neighborhood (10, 11); (ii) diffusion model assumptions (9); (iii) repeated measurements that are often not available (6–8); or (iv) some way of decomposing images into "low frequency" (signal) and "high frequency" (noise), e.g. using wavelets (4, 12, 13), which tends to overestimate the noise level, since sharp edges contribute to the high frequency sub-band. Physiological noise, image misalignment and model inaccuracies, all tend to bias the diffusion model and repetition-based methods (4, 6). The obvious manifestation of this bias is the artifactual anatomical structure visible in the estimated noise maps.

Here we present a model-independent local noise estimation method based on the random matrix theory (RMT) results for noisy covariance matrices. RMT was pioneered by Eugene Wigner and Freeman Dyson in the 1950s-1960s (14, 15) to describe statistics of energy levels in large nuclei and further applied in a wide range of fields. Our method is free of the above limitations and artifacts. Simulations and experiments show it outperforms existing methods in terms of precision and accuracy by an order of magnitude, reducing error in local noise estimation down to about 1% in routine dMRI acquisitions.

Technically, we employ principal component analysis (PCA) coupled with RMT, to exploit the redundancy in multi-directional dMRI data. In the limit of a large covariance matrix size, the noise contribution to the histogram of PCA eigenvalues becomes *deterministic*, and is given by the universal Marchenko-Pastur (MP) distribution (16) parametrized by the non-Gaussian noise variance σ^2. This universality allows us to determine the noise level. Hereafter, we will refer to the proposed method as MP-PCA.

Noise level estimation using RMT has previously been explored in the context of finance (17), wireless communication systems (18), cardiac MRI (19), DCE-MRI (20), and fMRI (21). Here we demonstrate for the first time the applicability of this approach in the context of dMRI. In particular, we show with simulations and experiments that typical dMRI data exhibit sufficient *redundancy* in order to identify the noise bulk from the PCA eigenspectrum using RMT. Next, we make the necessary step from Gaussian to nc-χ distributions in order to resolve noise bias in quantitive MR modalities. For that, we study the manifestation of the nc-χ bias in the PCA eigenspectrum. We demonstrate that by applying the inversion technique (22), the underlying Gaussian noise level σ_c (i.e. the noise standard deviation in the coils) that parameterizes the non-central χ distribution can accurately be determined from σ and the reconstructed signal obtained by nullifying the noise bulk.

THEORY

Noise in magnitude MR data

Phased array coil technology has seen significant developments in the last decades. The introduction of coil systems with a large number of channels, along with new parallel imaging techniques has resulted in significant improvements in scan times or SNR. The resulting magnitude MR data distribution, however, depends on the reconstruction method that is used to combine the complex signals from all independent channels. In general, magnitude MR data is approximately nc-χ distributed with $2L$ degrees of freedom (23). Here, L corresponds to the effective number of coils and depends on both the hardware and the applied reconstruction technique (24, 25).

The nc-χ distributed magnitude signal m is described by following probability density function:

$$p_L\left(m|\nu,\sigma_c,L\right)=\frac{m^L}{\sigma_c^2}\nu^{1-L}\quad \exp\left(-\frac{\nu^2+m^2}{2\sigma_c^2}\right)I_{L-1}\left(\frac{\nu m}{\sigma_c^2}\right),$$ [1]

with ν the noise-free magnitude signal and I_{L-1} the $(L-1)^{\text{th}}$-order modified Bessel function of the first kind (26). For $L=1$, the distribution reduces to the Rice distribution. The expectation value μ of the nc-χ distributed variable m is

$$\mu=\beta_{L\,1}F_1\left(-\frac{1}{2},L,-\theta^2\right)\sigma_c,$$ [2]

with $\theta=\nu/\sigma_c$ the SNR, $_1F_1$ the confluent hypergeometric function, and

$$\beta_L=\sqrt{2}\frac{\Gamma\left(L+\frac{1}{2}\right)}{\Gamma\left(L\right)}=\sqrt{\frac{\pi}{2}}\frac{(2L-1)!!}{2^{L-1}(L-1)!}.$$ [3]

The variance of m equals

$$\sigma^2=2L\sigma_c^2+\nu^2-\mu^2.$$ [4]

Since the noise bias of the measured signal increases with L, it is beneficial to choose the image reconstruction of the complex coil images that minimizes that parameter. Indeed, reconstructions algorithms such as SENSE, adaptive combine (27) or the recently proposed SENSE1 (28) algorithm always result in $L=1$, i.e. a Rician data distribution, even if parallel imaging has been applied, and therefore, they are superior to sum-of-squares reconstructions from a statistical point-of-view.

Principal component analysis (PCA)

Let X be an $M \times N$ real-valued measurement matrix, with rows representing M voxels within a sliding window, and columns representing N weighted magnitude MR measurements in each voxel. In what follows, we will take voxels from a square or a cubic patch of a dMRI data set; however, one could also group voxels based on other principles, e.g. within a particular region of interest or tissue type. What will be important is that the matrix X is sufficiently large, i.e. $R \equiv \min\{M, N\} \gg 1$.

We expect (as confirmed a posteriori), that the dMRI data is *redundant*. Namely, due to correlations between many diffusion weighted (DW) signals brought by a generally unknown underlying diffusion process, and/or because the dMRI signal can be efficiently represented by a small number of parameters, X can be approximately viewed as a linear combination of a few, $P << R$, linearly independent sources, or *principal components*. In other words, the high-dimensional measurement is effectively reduced to a P-dimensional hyperplane. PCA (29, 30) finds the basis of this hyperplane, the principal components, as eigenvectors of the $M \times M$ sample covariance matrix

$$\sum = \frac{1}{N} X X^T = U \Lambda^2 U^T. \qquad [5]$$

In Equation [5], U is an orthonormal matrix whose columns are the principal components, and Λ^2 is an $M \times M$ diagonal matrix of eigenvalues $\lambda_1 \cdots \lambda_M$. For numerical precision, U and Λ are generally computed via the singular value decomposition of $X = \sqrt{N} U \Lambda V^{\mathbf{T}}$.

We will focus on the eigenvalues, which we sort in the descending order, $\lambda_1 \geq \lambda_2 \geq \ldots$. In the noise-free *population* case, corresponding to fixed M and $N \to \infty$, Σ will have exactly P nonzero eigenvalues.

In reality, we are faced with a *sample* covariance matrix with $N \sim M$, built from noisy data. Noise propagates through PCA to the eigenvalues and eigenvectors, making all R eigenvalues nonzero, so that rank $\Sigma = R$. Remarkably, while noise in each measurement is random, its contribution to the histogram of covariance matrix eigenvalues becomes deterministic in the limit $R \gg P$ (16).

PCA of a completely random covariance matrix

To understand the role played by noise, we first consider the spectrum of Equation [5] when the measurement is pure white noise, $P = 0$. If X has uncorrelated independent identically distributed (iid) Gaussian entries with variance σ^2 and zero mean, the nonzero eigenvalues of Σ are distributed according to the following universal "quarter-circle" law (Figure 1):

$$p(\lambda | \sigma, \gamma) = \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\gamma\lambda\sigma^2}, \lambda_- \leq \lambda \leq \lambda_+ \qquad [6]$$

and $p(\lambda) \equiv 0$ otherwise, where $\lambda_{\pm}\sigma^2(1 \pm \sqrt{\gamma})^2$ with $\gamma = M/N$ the matrix size ratio. Equation [6] has been derived by Vladimir Marchenko and Leonid Pastur in 1967 (16), and is analogous to the celebrated Wigner's semicircle law for the eigenvalue distribution of random Hamiltonians. For $\gamma < 1$, $p(\lambda)$ is normalized to $\int_{\lambda_-}^{\lambda_+} pd\lambda = 1$, whereas, for $\gamma > 1$, $p(\lambda)$ is normalized to $1/\gamma$, with the $(1 - 1/\gamma)\delta(\lambda)$ density at zero. The corrections to Equation [6] for finite R decrease fast with R, as $\sim 1/R$.

We will use the fact that the spectrum resulting from Equation [6] has lower and upper bounds $\lambda_{\pm}$ that depend on the noise level σ and matrix size ratio γ. Physically, the MP distribution arises due to eigenvalue repulsion around the population density $\delta(\lambda - \sigma^2)$ (Equation [6] for $\gamma \rightarrow 0$). The repulsion between levels $\lambda_{i,j}$ with the force $\sim 1/|\lambda_i - \lambda_j|$ maps finding $p(\lambda)$ onto an equilibrium profile in an electrostatic problem (15), fixing the bounds $\lambda_{\pm}$ and the spectral shape with square root singularities resulting in Equation [6].

PCA of noisy redundant data

The universality of the MP distribution for the noisy part of PCA still applies if the data X includes P signal sources (significant components), as long as $P << R$ (31). Intuitively, this is guaranteed by the electrostatic analogy (15) — a few charges, representing the significant eigenvalues, cannot noticeably distort the bulk charge density. Figure 1 shows a diffusion tensor based example for $P = 7$ (1 non-DW signal + 6 DTI elements), $N = 256$, $M = 512$, where the shape of the MP "sea" of eigenvalues is very well described by Equation [6].

Depending on the noise level, some of the significant eigenvalues can fall into the MP bulk (31) and become invisible (indistinguishable from noise). Hence, from now on we denote P the number of eigenvalues *above the MP edge* λ_+. We will develop an iterative noise mapping procedure based on identifying the MP bulk of $\tilde{N} = R - P$ eigenvalues.

METHODS

Noise level estimation

The combination of local PCA with RMT can now be used to estimate the noise level locally. Distribution fitting, that is, minimizing the error between $p(\lambda|\sigma, \tilde{\gamma})$, i.e. the MP distribution of nonzero eigenvalues, and the histogram of the lowest $\tilde{N}$ eigenvalues of Σ by varying the noise level σ and matrix size ratio $\tilde{\gamma} \equiv M/\tilde{N}$, yields an estimate of the parameters of interest.

An heuristic search algorithm is used to solve this estimation problem. The general idea is to estimate σ and $\tilde{N}$ simultaneously by changing $\tilde{N}$ and estimating the corresponding σ iteratively until $\tilde{N}$ equals the number of eigenvalues ranged by λ_- and λ_+, which are determined by both $\tilde{N}$ and σ. More specifically, the algorithm is as follows (a schematic overview of the algorithm is shown in Figure 2): (i) At the ith iteration, we have the number $\tilde{N}_i = R - P_i$ of MP (noise-only) eigenvalues, i.e. $\left[\lambda_{P_i+1}, \cdots, \lambda_R\right]$. Its initial guess $\tilde{N}_1 = R - P_1$ is related to the physically expected number P_1 of "observable" model parameters; we take $P_1 = 6$ for $b \leq 1$ ms/μm^2 DW data and $P_1 = 15$ for higher b, cf. the

number of spherical harmonic coefficients in a 2nd and 4th order basis, respectively. Overall, we found the dependency of the algorithm to the initial P_1 to be negligible. (ii) The histogram of the smallest nonzero $\tilde{N}_i$ eigenvalues is created by binning $\left[\lambda_{P_i+1}, \cdots, \lambda_R\right]$ into n_b bins with center locations $\left[\bar{\lambda}_1, \cdots, \bar{\lambda}_{n_b}\right]$ and corresponding bin counts $\rho=[\rho_1, \cdots, \rho_{n_b}]$. The computation of the number of bins of the histogram, n_b, is data-driven:

$$n_b = \arg\ \min_n \frac{2\bar{\rho} - v}{h^2\left(n, \tilde{N}\right)}. \qquad [7]$$

with $h\left(n, \tilde{N}\right) = \left|\lambda_R - \lambda_{P_i+1}\right|/n$ being the bin width. Here,

$\bar{\rho} = \frac{1}{n}\sum_{i=1}^n \rho_i$ and $v = \frac{1}{n}\sum_{i=1}^n \left(\rho_i - \bar{\rho}\right)^2$ the mean and biased variance of the bin counts, respectively (32).

(**iii**) A rough estimate of the noise level is computed as $\tilde{\sigma}_i = \sqrt{\lambda_R}/\left|1 - \sqrt{\tilde{\gamma}_i}\right|$ with $\tilde{\gamma}_i = M/\tilde{N}_i$ and λ_R the lowest nonzero eigenvalue and a proxy for λ_-. Note that we prefer to approximate σ_i by means of an estimate of λ_- instead of λ_+ as its dependency on $\tilde{N}$ is expected to be weaker. Indeed, conceptually λ_{Pi+1} can serve as an estimation for λ_+. However, λ_{Pi}, λ_{Pi+1}, and λ_{Pi-1} might practically differ by an order of magnitude (the right edge of the MP "sea" fluctuates according to the Tracy-Widom distribution (31)), whereas λ_R is nearly independent of the search variable P_i.

(**iv**) An extensive grid search, center around $\tilde{\sigma}_i$, is done to determine the optimal σ_i. We minimize the weighted squared error between the histogram of $\tilde{N}_i$ eigenvalues and the distribution (Equation [6]) with given $\tilde{\gamma}_i$:

$$\sigma_i = \arg\ \min_\sigma \sum_{n=1}^{n_b} W_n \left(\rho_n - \varsigma p\left(\bar{\lambda}_n | \sigma, \tilde{\gamma}_i\right)\right), \qquad [8]$$

with $\varsigma = h\left(n_b, \tilde{N}_i\right)\tilde{N}_i$ a scaling factor to convert probabilities into counts. The inverse of those predicted counts are chosen as weights, i.e. $W_n = \left(\varsigma p\left(\bar{\lambda}_n | \sigma, \tilde{\gamma}_i\right)\right)^{-1}$, to compensate for the strong non-uniformity of the distribution when $M \approx \tilde{N}$, due to the $\lambda^{-1/2}$ divergency of $p\left(\lambda\right)|_{\tilde{\gamma}-1}$ as $\lambda \to 0$.

(**v**) We then compute the right MP edge $\lambda_{+,i} = \sigma_i^2(1 + \sqrt{\tilde{\gamma}_i})^2$ and compute the number P_{i+1} of observable parameters, as the number of eigenvalues exceeding $\lambda_{+,i}$.

(**vi**) We iterate (**i**) to (**v**) until $P_{i+1} = P_i$.

The universality of MP law requires both M, the number of voxels, and $\tilde{N}$, the number of noiseonly components, to largely exceed 1. Obviously, $\tilde{N}$ depends on the chosen diffusion protocol and the degree of redundancy (which depends both on M and on the strength of diffusion weighting). The freedom to choose $M \gg 1$ enables finding a compromise between spatial resolution and the redundancy. The use of a sliding window (e.g. $M = [11 \times 11]$ or $M = [5 \times 5 \times 5]$) will yield the 2d or 3d noise map, respectively. We hereby assume the true noise map can be considered to be fairly constant within those local neighborhoods. Because of that assumption, we here opt to analyze the native diffusion-weighted measurement matrix X instead of $\log(X)$ to avoid strong signal dependency of the noise level induced by the log transform.

Correction for non-central χ statistics

The standard deviation

$$\sigma = \xi(\theta) \leq 1 \quad [9]$$

of a nc-χ distributed variable with L degrees of freedom is always smaller than the standard deviation σ_c of the Gaussian noise in the individual complex-valued coil images (22). The factor

$$\xi(\theta) = 2L + \theta^2 - \beta_L^2 \cdot \left[{}_1F_1\left(-1/2, L, -\theta^2/2\right) \right]^2 \quad [10]$$

is a function of SNR θ. Because the SNR is not known a priori, the correction $\xi(\theta)$ needs to be estimated iteratively until convergence, initiated by both the mean, $\mu(\theta)$, and variance, $\sigma^2(\theta)$, of the magnitude signal, cf. Koay's inversion (22).

We use P significant eigenvalues (i.e. the ones exceeding the right MP edge λ_+) to estimate the signal mean in each voxel and diffusion gradient direction, $X = \sqrt{N} U \Lambda V^{\mathbf{T}}$, where $\hat{\Lambda}$ is obtained from the SVD-derived Λ by nullifying all but the first P eigenvalues. The remaining eigenvalues can further be corrected, cf. (31). We then use Koay's inversion for each element $\hat{X}_{i,j}$ separately.

Overall, we expect the noise map to be slightly *underestimated* because some of the noise corrupts the P significant eigenvalues and is not accounted in the distribution fitting. We estimate this bias at the level of $1/R$.

Aja-Fernández et al. showed that in case of of parallel imaging, the actual magnitude of MR signal may deviate from a pure nc-χ distribution to an unknown degree (23, 25). This happens if the individual coil images are combined using the sum-of-squares while noise correlations exist between the different coil images. Obviously, if an accurate analytical expression for the actual data distribution is lacking, accurate bias correction cannot be performed. However, this potential source of bias roots in noise modeling and the unknown propagation of the coil correlations into the data distribution, and is not inherent to the use of

MP-PCA. For more information regarding this related, though independent problem, we refer to (4, 23, 25).

Previous noise mapping approaches

We will compare our method with following state-of-the-art techniques for noise estimation:

Repetitions—In the simplest case, σ is computed as the standard deviation of a set of repeated measurements. If the mean is calculated as well, Koay's inversion can be applied to correct for nc-χ bias. Physiological noise may interfere with the estimation of the noise level. To overcome that potential source of inaccuracy, one can replace the mean and the standard deviation by their robust counterparts, i.e. median and median absolute deviation (MAD) estimators, respectively (6).

Residuals—Weighted least squares fitting of an adequately chosen signal representation, e.g. DTI, to the data will result in residuals with statistics directly related to the noise variance σ^2 (33). An estimate of the expectation value of the directional DW signals is found by reconstructing the signal based on the estimated noise parameters, employing Koay's inversion. Note that this approach is equally accurate as the simultaneous estimation of signal and noise parameters by means of log-likelihood functions (34, 35) if the data is Rician distributed with a minimal SNR of 2 (33, 36).

Wavelets—The wavelet decomposition provides a means to separate low from high frequency information. Assuming the high frequency components contains residual signal that originated in high image gradients in addition to noise, Coupé et al suggested the MAD of the absolute value of the high frequency components as a global estimate of σ (12). Correction for nc-χ statistics is again possible. A local dMRI-specific variant of their approach was presented by Veraart et al (4).

Data

As the overarching issue with noise mapping is our inherent lack of noise-free ground truth, below we employ both the simulated data based on an atlas (37), and a human dMRI measurement with large oversampling for a range of diffusion weightings.

Simulated dMRI data—We simulated whole brain DW data, derived from a hybrid diffusion atlas that was created with 10 healthy subjects (37). The choice for that atlas roots in its high precision and a high order parameterization of the underlying diffusion signal by spherical harmonics, allowing us to employ it as a noise-free "ground truth". We resampled the $b = 1$ ms/μm^2-shell of the atlas data on 90 isotropically distributed gradient directions (g) to serve as gold standard for the antipodal symmetric signal

$S(g) = \sum_{l=0,2,4} \sum_{m=-l}^{l} c_m^l Y_l^m(g)$ with c_m^l the real harmonic series coefficients and Y_l^m an orthonormal basis for real functions of the unit sphere with order l and phase factor m (38). After adding complex Gaussian noise with standard deviation σ_c, the magnitude images were computed to generate Rician distributed DW data. In total, 4 noise realizations were

made: spatially uniform noise level with average SNR$_{b=0}$ = 20 and 40 within a brain mask, and spatially varying noise level with average SNR$_{b=0}$ = 20 and 40.

We applied our proposed MP-PCA method and the wavelet-based method to estimate the noise map on each complete simulated data set, i.e. N = 90 gradient directions, as well as on subsets of N = 30 and 60 gradient directions. Subsets were chosen in such a way that each subset was composed out of isotropically distributed directions. For MP-PCA, we additionally varied the sliding window size ($M = [n \times n \times n]$ with $n = \{5, 7, 9\}$). For a quantitative comparison of the different techniques, we use the relative error, i.e. a voxelwise measure that is calculated as the difference between a measurement and the local ground truth value, divided by the ground truth value. The performance of the residual-based method would be misleadingly high due to the lack of physiological noise and motion in the simulation set-up and, as such, not shown here.

Experiment—A 30 year-old healthy volunteer underwent imaging on a Siemens Skyra (3T) MR scanner (Siemens AG, Siemens Medical Solutions, Erlangen, Germany) after obtaining informed consent, using a 20-channel receiver head coil. The body coil was used for transmission. An EPI-DW sequence was used to acquire the dMRI data. Diffusion weighting was applied along 90 isotropically distributed gradient directions with b = 1, 2, and 3 ms/μm^2. Same gradient direction were applied for each b. Additionally, 4×23 additional images were acquired: 23 images without diffusion sensitization and 3×23 DW images with diffusion sensitization (b = 1, 2, and 3 ms/μm^2) along a single gradient direction. Other imaging parameters were kept constant throughout the data acquisition sequences: TR/TE : 5000/106, ms, matrix: 88×88, slice thickness: 2.5 mm, slices: 50, parallel imaging factor: GRAPPA with acceleration factor 2, reconstructed using the adaptive combine algorithm (L = 1), multiband acceleration with factor 2.

Noise maps were estimated from b-value dependent subsets of the measured data, by employing (i) robust noise map estimator using repeated measurements; $23 \times \{b = 0, 1, 2, 3\}$ ms/μm^2 (6), (ii) statistics of the residuals after a diffusion tensor imaging (DTI) fit using the weighted linear least squares estimator; subsets were $5 \times \{b = 0\} + 90 \times \{b = 0, 1, 2, 3\}$ ms/μm^2 (33, 36), (iii) the local wavelet-based approach; subsets were $90 \times \{b = 1, 2, 3\}$ ms/μm^2 (4), and (iv) our proposed MP-PCA with a $[7 \times 7 \times 7]$ sliding window; $90 \times \{b = 1, 2, 3\}$ ms/μm^2. Additionally, we applied our proposed MP-PCA method on subsets of b_i consisting of 30 and 60 isotropically distributed gradient directions, respectively.

RESULTS

nc-χ distributed random matrices

Figure 3 shows the average noise level estimated using MP-PCA from 500 independent realization of a 125×90 matrix X of which each entry X_{ij} is a sample for a nc-χ distribution with ν = SNR and σ_c = 1. The noise level is shown as function of the SNR for L=1, 8, 16, and 32, before ($\hat{\sigma}$; red) and after ($\hat{\sigma}_c$; green) correction for the nc-χ bias, respectively. The standard deviation of the respective nc-χ distributed variables, σ, are shown to demonstrate that in case of nc-χ distributed data, the proposed framework returns an accurate estimate of σ instead of σ_c. As shown in Figure 3, the signal $\hat{\mu}$ reconstructed from the principal

components after nullifying all but the first P eigenvalues matches the expectation value μ instead of the respective noise-free underlying signal ν. In this simulation, P was always determined to be 1 by our algorithm, even in case of SNR=0, which indicates that the nc-χ signal bias leaks out of the noise bulk and contributes to *signal* component(s). The observation that we can estimate both σ and μ justifies the use of Koay's inversion technique (22) to compute estimates of the Gaussian noise level σ_c regardless of L. Hence, without loss of generality, we limit rest of the manuscript to $L = 1$. It can be observed in Figure 3 that for very low SNR, $\hat{\sigma}_c$ is slightly underestimated. However, the underestimation is inherent to the inversion technique, not to the MP-PCA estimation of σ or μ, which is shown to be accurate.

Simulation data

Figure 4 and Figure 5 show the noise maps estimated from Rician distributed simulated data with spatially uniform and spatially varying noise levels, respectively, using (i) our proposed local MP-PCA technique and (ii) the local wavelet-based noise map estimator (4). The spatially varying noise map is shown in Figure 5 (labeled REF) for reference. A visual inspection of the maps shows us that the wavelet approach tends to overestimate the noise, especially near sharp edges in the DW images. As a result, the underlying image is reflected in the estimated noise maps, e.g. CSF. However, this bias depends on the SNR. The lower the SNR, the better the decoupling of noise and high frequency edge information by a wavelet decomposition. Nonetheless, regardless the SNR, our proposed technique, MP-PCA, shows about 99% accuracy and precision in the local estimation of the noise level, even for low N.

In Figure 6 we show the histograms of the relative errors in the estimated noise levels within the brain mask. These quantitative comparison includes the results for the global noise estimator based on the wavelet decomposition, proposed by Coupé et al. (12) in case of the spatially uniform noise map. For the global wavelet estimator, a 12.3% overestimation of the noise level is observed for SNR=40, whereas, for SNR=20, the bias reduces to 3.7%.

In Table I, we listed the mean relative error (MRE) and the standard deviation of the relative errors (SRE), both calculated within the brain mask. The global wavelet approach resulted in a single noise level for each DW image. The corresponding MRE and SRE were computed from that set of values. The MRE indicates the high accuracy of MP-PCA, regardless of the SNR. In the worst case, we observed a MRE of 2.9%, but in most cases the absolute value of the MRE < 1%. For the local wavelet based approaches, the performance in terms of MRE ranged between 5% and 20%. Similarly, MP-PCA is better in terms of SRE, which probes the precision. In general, SRE decreases with increasing M and N.

Increasing the window size might lower the performance in case of spatially varying noise. Indeed, the selection of the sliding window size is influenced by two opposing effects. First, the histograms for the spatially uniform noise suggest that increasing the window size increases the precision of the estimator due to increase in the random matrix size. However, for the spatially varying noise case, we observed a slight bias in the local noise estimation due to the reduced *resolution* of large window sizes. This is visible in both the maps and histograms. We find [7×7×7] the better trade-off in between precision and accuracy in our experiments.

Experiment

Figure 7 shows the noise maps estimated from b-value dependent subsets of the measured data for the different estimators. The following qualitative observations can be made. First, despite the use of robust statistics, noise map estimation based on repeated measurements are prone to artifacts related to physiological noise, CSF pulsation in particular. The use of high b-value largely minimized that effect. Hence, the map estimated from the 23 repetitions with $b = 3$ ms/μm^2, i.e. the highest b-value in our study, serves as the most reliable proxy for the ground truth, hereafter termed *bronze standard*. Second, the noise maps based on residuals after a diffusion model fit are only as accurate as the diffusion model. Unsurprisingly, the accuracy decreases with the increase in b-value, especially noticeable in the major white matter structures. Moreover, although not overly represented in this data set, this technique might suffer from subject motion and physiological noise. Fig. 8 in ref. (4) shows for example how bad it can get in clinical data. Third, wavelet-based approaches suffer from an incomplete separation of noise and signal by a wavelet transformation at high SNR. Therefore, at low b-value, anatomical structure can be observed in the estimated noise maps. Although the effect reduces with b, it does not disappear in our data. Finally, MP-PCA shows a smooth, artifact-free noise map that is more consistent with the bronze standard than other methods for all b-values.

Figure 7 shows scatter plots in which the *best* result of (ii), (iii), and (iv) are shown voxel-by-voxel against our bronze standard (i.e. $b = 3$ ms/μm^2). It can be observed that MP-PCA shows the highest correspondence. Moreover, that technique has the least b-value dependency and, as such, is the most generally applicable one. The high correspondence can also be quantified by the correlation coefficient. All MP-PCA maps correlate well with the bronze standard, correlation coefficients are all 0.90, whereas for the residual-based method, the coefficients range between 1.81 and 0.83, and for the wavelet-based technique, they range between 0.41 and 0.88.

Figure 8 shows the dependence of MP-PCA on the number N of DW images by selecting isotropically distributed subsets with $N = 30$, 60, and 90. Furthermore, the performance of MP-PCA is evaluated as function of the sliding window size. In line with previous experiments, MP-PCA results are very consistent, regardless of N and sliding window size. In general, the precision tends to increase with N and the window size. However, theoretically, the performance is eventually expected to decrease with (a) increasing the window size (potential heteroscedasticity of the underlying noise level) or (b) reducing N (redundancy loss when N approaches P).

A final demonstration of the high accuracy of MP-PCA in comparison to other approaches is given in Figure 9. The DW signals in the CSF voxels are suppressed down to $e^{-3 \times 3} \approx 10^{-4}$ level at $b = 3$ ms/μm^2 and, as such, can be assumed to be *noise-only*. Hence, the average of those signals equals the noise floor, i.e. the minimally measurable signal. Since our data is Rician distributed, the noise floor is given by $\sqrt{\pi/2}\sigma_c$. After normalization of the DW images with the spatially varying noise maps of σ_c, obtained from the different approaches, we averaged the signals within the CSF. The mean and standard deviation over all the DW images present for each b-value are shown in Figure 9. One can readily observe that the MP-

PCA normalized DW signals yield the best correspondence to the normalized noise floor, $\sqrt{\pi/2},$ at $b=3\mathrm{ms}/\mu\mathrm{m}^2$. Normalization of the DW signals was needed to cope with the spatially varying nature of the noise maps.

DISCUSSION

We presented an accurate, precise and robust method for the local noise estimation. The method exploits the redundancy in multi-directional diffusion MRI data by identifying the noise-only principal components using the knowledge that the corresponding eigenvalues are described by the universal Marchenko-Pastur distribution, parameterized by the noise level.

Previous noise mapping methods all come with serious limitations. Repeated measurements are often not available due to scan time constraints. Misalignment of data limits the accuracy of all methods other than the wavelet-based approaches. Model inaccuracies due to physiological effects such as CSF pulsation will distort the statistics of the residuals or the accuracy of likelihood functions, and the accuracy of wavelet-based approaches strongly depends on the presence of residual signal in the high frequency sub-band. We demonstrated all those limitations in clinically representative experiments while showing that our MP-PCA technique is much less sensitive to the above complications inherent to a dMRI data set.

Here, we further elaborate on the strengths, limitations, and potential applications of the method.

Redundancy requirement

The Marchenko-Pastur distribution is an asymptotic universal law that holds in the large matrix size limit. In this regard, one must consider the size of the $\left[M \times \tilde{N} \right]$ noise-only matrix that we obtain by decomposing the $[M \times N]$ measurement matrix in its principal components. Hence, the noise part $\tilde{N}=N-P$ will only be large if the measurements are highly redundant, $P \ll N$. First, small P can be achieved by limiting the diffusion-weighting. Indeed, the complexity of the underlying, generally non-Gaussian diffusion process, and as such P, tends to increase with the bvalue. Second, P increases with the size M of the sliding window due to the potential heterogeneity in noise level within the set of voxels. Simulations showed that the performance of our approach might become lower for large M (see Figure 6). Hence, the beauty of our proposed technique is that *it performs best when applied locally*. However, one rather chooses $M > N$ to avoid the situation when M becomes the dominating factor in the deviation from the Marchenko-Pastur law, which is of order $O(1/\min\{M, N\})$.

Using the above guidelines, we demonstrated high accuracy of the noise estimator even for DW data sets limited to 30 gradient directions. Therefore we showcased the wide applicability of this theoretical framework.

We would like to emphasize that redundancy is not tied to dMRI, the modality of our interest throughout this work. Indeed, the proposed technique can be applied in other modalities

including, but not limited to, redundant time series, such as DCE-MRI (20) and fMRI. Neither is application of this technique limited to a single DW shell in q-space.

Ground truth

The main limitation of our study is the lack of ground truth in our experimental part. Due to obvious reasons one cannot measure noise-free signal. Regarding the ground truth for the noise, unfortunately, one cannot measure exactly such a reference map either. On the one hand, the thermal noise level depends on the load of the scanner (39). Hence, an air-only acquisition will not provide the required information. On the other hand, although switching off the excitation pulses might mimic such an acquisition, image reconstructions often require the signal, i.e. autocalibration signal (ACS) lines, for accurate calculation of the coil sensitivity maps that determine the spatially varying nature of the noise map (40).

Given all these limitations, we assessed the accuracy of the proposed technique. First, we tried to create the most reliable proxy ("bronze standard") for the ground truth by acquiring multiple repetitions of the same DW image. Despite the use of a robust estimator (6), we observed that physiological noise and misalignments between repeated measurements tend to bias the noise estimation. The effect was significantly lower for higher b-values. Second, we used the knowledge that CSF is fully suppressed at $b = 3$ ms/μm^2, to evaluate the accuracy of the methods. Indeed, within the CSF the average signal should match the noise floor, which is parameterized by the noise level. As shown in Figure 9, the best correspondence was achieved for MP-PCA.

Applications

Noise bias—To remedy the systematic bias of diffusion parameters that is inherent to the use of least squares estimators in case of nc-χ distributed data, one may prefer using more advanced estimators that explicitly account for the actual MR data distribution (4, 35). Those distributions are parameterized by the noise level. Hence, prior knowledge of that parameter enables an improved precision and accuracy of diffusion model parameter estimators.

Outlier detection—Signal variability in DW images roots in both thermal noise and spatially and temporally varying artifacts such as subject motion and cardiac pulsation. Those spatially and temporally varying artifacts, which are referred to as physiological noise, generally cannot be modeled. Robust parameter estimation techniques such as RESTORE (41, 42) have been introduced to identify the signal outliers and minimize their effect on the the diffusion model parameters. However, the thermal noise level is required in most of the robust estimators in order to identify those outliers with some user-defined probability.

Gibbs ringing—Gibbs ringing has recently been recognized as a confound in diffusion MR analyses (43–45). Moreover, we analytically derived its significant quantitative impact on the diffusion signal and model parameters (45). To remedy the strong bias on diffusion metrics due to the Gibbs ringing artifact, we suggested to extrapolate the k-space beyond the measured part (46) by means of second order total generalized variation minimization (45, 47). These types of regularized optimization problems generally come with a regularization

term that controls the data fidelity. Optimal choice of this regularization term is based on the thermal noise level (48).

Denoising

By applying PCA to redundant data and preserving only the significant principal components, the noise can be strongly reduced. Manjon et al. (49) recently proposed image denoising in the context of dMRI by using a sliding window wherein the local PCA transformation matrix was estimated. Denoising is then achieved by canceling all principal components with eigenvalues below some threshold; the key uncertainty in this approach is choosing this threshold. The authors rely on empirically defined thresholds, e.g. $5.29\sigma^2$ for their particular M and N. Here, we suggest using an objective threshold $\lambda_+ = \sigma^2 \left(1 + \sqrt{\tilde{\gamma}}\right)^2$ based on the right edge of the Marchenko-Pastur distribution [6]. Hence we foresee that considering λ_+ as a threshold value will further improve and generalize PCA based denoising methods. However, this is beyond the scope of this work.

Nonetheless, the term denoising must be put in right perspective. Noise is not only represented in a limited number of principal components that can be canceled upon request. Indeed, in practice, all principal components, the eigenvalues and eigenvectors, are corrupted by noise to some degree. Some (or all) significant eigenvalues might be indistinguishable from the noise spectrum in case of low SNR and large γ. Hence, denoising techniques can remove some of the principal components describing the noise-free signal. Next, eigenvalues distinct from the noise bulk are biased by a coherent repulsion of the significant eigenvalues by all the pure noise eigenvalues (31). Additionally, the corruption of the eigenvectors by the noise cannot be undone. These elements explain the negative MREs in Table I and put the term denoising in perspective. Consequently, the standard deviation of the reconstruction of the noise-only signal has systematically lower variance than the estimated noise variance (results not shown). Hence, data variability due to noise is only partially canceled by MP-PCA denoising.

CONCLUSION

We used the universal Marchenko-Pastur distribution to identify the noise-only eigenvalues obtained by a local PCA of nc-χ distributed dMRI data. Typical dMRI data has shown to exhibit sufficient redundancy which allows to distinguish between the noise and signal components in a model-independent way, and without the need to empirically set the PCA threshold value. Our proposed local noise estimator outperforms known spatially varying noise estimation methods in terms of accuracy ($\sim 99\%$) and precision, and does not require acquisition of repeated measurements, instead utilizing data redundancy.

ACKNOWLEDGMENTS

References

1. Le Bihan D, Breton E. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. Comptes-Rendus de l'Académie des Sciences. 1985; 93:27–34.

2. Jones, DK. Diffusion MRI: Theory, methods, and applications. Oxford University Press; 2010.

3. Jones DK, Basser PJ. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. Magnetic Resonance in Medicine. 2004; 52:979–993. [PubMed: 15508154]

4. Veraart J, Rajan J, Peeters RR, Leemans A, Sunaert S, Sijbers J. Comprehensive framework for accurate diffusion MRI parameter estimation. Magnetic Resonance in Medicine. 2013; 70:972–984. [PubMed: 23132517]

5. Robson PM, Grant AK, Madhuranthakam AJ, Lattanzi R, Sodickson DK, McKenzie CA. Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions. Magnetic Resonance in Medicine. 2008; 60:895–907. [PubMed: 18816810]

6. Maximov II, Farrher E, Grinberg F, Shah NJ. Spatially variable Rician noise in magnetic resonance imaging. Medical image analysis. 2012; 16:536–548. [PubMed: 22209560]

7. Glenn GR, Tabesh A, Jensen JH. A simple noise correction scheme for diffusional kurtosis imaging. Magnetic resonance imaging. 2015; 33:124–133. [PubMed: 25172990]

8. Landman BA, Bazin PL, Prince JL. Estimation and application of spatially variable noise fields in diffusion tensor imaging. Magnetic resonance imaging. 2009; 27:741–751. [PubMed: 19250784]

9. Landman BA, Bazin PL, Smith SA, Prince JL. Robust estimation of spatially variable noise fields. Magnetic Resonance in Medicine. 2009; 62:500–509. [PubMed: 19526510]

10. Aja-Fernández S, Pieçiak T, Vegas-Sánchez-Ferrero G. Spatially variant noise estimation in MRI: A homomorphic approach. Medical image analysis. 2015; 20:184–197. [PubMed: 25499191]

11. Tabelow K, Voss HU, Polzehl J. Local estimation of the noise level in MRI using structural adaptation. Medical image analysis. 2015; 20:76–86. [PubMed: 25465845]

12. Coupé P, Manjón JV, Gedamu E, Arnold D, Robles M, Collins DL. Robust Rician noise estimation for MR images. Medical image analysis. 2010; 14:483–493. [PubMed: 20417148]

13. Goossens, B.; Pizurica, A.; Philips, W. Wavelet domain image denoising for non-stationary noise and signal-dependent noise; Image Processing, 2006 IEEE International Conference on; IEEE. 2006. p. 1425-1428.

14. Wigner EP. Lower limit for the energy derivative of the scattering phase shift. Physical Review. 1955; 98:145.

15. Dyson FJ. Statistical theory of the energy levels of complex systems. i. Journal of Mathematical Physics. 1962; 3:140–156.

16. Marchenko VA, Pastur LA. Distribution of eigenvalues for some sets of random matrices. Matematicheskii Sbornik. 1967; 114:507–536.

17. Laloux L, Cizeau P, Bouchaud JP, Potters M. Noise dressing of financial correlation matrices. Physical review letters. 1999; 83:1467.

18. Ahmed A, Hu YF, Noras JM. Noise variance estimation for spectrum sensing in cognitive radio networks. AASRI Procedia. 2014; 9:37–43.

19. Ding Y, Chung YC, Simonetti OP. A method to assess spatially variant noise in dynamic MR image series. Magnetic Resonance in Medicine. 2010; 63:782–789. [PubMed: 20187185]

20. Jahani, J.; Johnson, G.; Kiselev, VG.; Novikov, DS. Proceedings 21th Scientific Meeting, International Society for Magnetic Resonance in Medicine. Salt Lake City, USA: 2013. Random matrix theory-based noise reduction for dynamic imaging: Application to DCE-MRI; p. 3073

21. Sengupta AM, Mitra PP. Distributions of singular values for some random matrices. Physical Review E. 1999; 60:3389.

22. Koay CG, Basser PJ. Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. Journal of Magnetic Resonance. 2006; 179:317–322. [PubMed: 16488635]

23. Aja-Fernández S, Tristán-Vega A, Hoge WS. Statistical noise analysis in GRAPPA using a parametrized noncentral Chi approximation model. Magnetic resonance in medicine. 2011; 65:1195–1206. [PubMed: 21413083]

24. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. Journal of Magnetic Resonance Imaging. 2007; 26:375–385. [PubMed: 17622966]

25. Aja-Fernández S, Tristán-Vega A. Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction. Magnetic Resonance in Medicine. 2012; 67:580–585. [PubMed: 21656560]

26. Constantinides CD, Atalar E, McVeigh ER. Signal-to-noise measurements in magnitude images from NMR phased arrays. Magnetic Resonance in Medicine. 1997; 38:852–857. [PubMed: 9358462]

27. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array MR imagery. Magnetic Resonance in Medicine. 2000; 43:682–690. [PubMed: 10800033]

28. Sotiropoulos S, Moeller S, Jbabdi S, Xu J, Andersson J, Auerbach E, Yacoub E, Feinberg D, Setsompop K, Wald L, et al. Effects of image reconstruction on fiber orientation mapping from multichannel diffusion MRI: Reducing the noise floor using SENSE. Magnetic Resonance in Medicine. 2013; 70:1682–1689. [PubMed: 23401137]

29. Pearson K Liii. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901; 2:559–572.

30. Hotelling H. Analysis of a complex of statistical variables into principal components. Journal of educational psychology. 1933; 24:417.

31. Johnstone IM. High dimensional statistical inference and random matrices. arXiv preprint math/0611589. 2006

32. Shimazaki H, Shinomoto S. A method for selecting the bin size of a time histogram. Neural computation. 2007; 19:1503–1527. [PubMed: 17444758]

33. Salvador R, Peña A, Menon DK, Carpenter TA, Pickard JD, Bullmore ET. Formal characterization and extension of the linearized diffusion tensor model. Human brain mapping. 2005; 24:144–155. [PubMed: 15468122]

34. Sijbers J, den Dekker AJ, Scheunders P, Van Dyck D. Maximum-likelihood estimation of Rician distribution parameters. IEEE Trans Med Imaging. 1998; 17:357–361. [PubMed: 9735899]

35. Veraart J, Van Hecke W, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. Magnetic Resonance in Medicine. 2011; 66:678–686. [PubMed: 21416503]

36. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. NeuroImage. 2013; 81:335–346. [PubMed: 23684865]

37. Dhollander, T.; Veraart, J.; Van Hecke, W.; Maes, F.; Sunaert, S.; Sijbers, J.; Suetens, P. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011. Springer; 2011. Feasibility and advantages of diffusion weighted imaging atlas construction in q-space; p. 166-173.

38. Courant R, Hilbert D. Methods of mathematical physics, volume 1. CUP Archive. 1966

39. Hoult D, Lauterbur PC. The sensitivity of the zeugmatographic experiment involving human samples. Journal of Magnetic Resonance (1969). 1979; 34:425–433.

40. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic Resonance in Medicine. 2002; 47:1202–1210. [PubMed: 12111967]

41. Chang LC, Jones DK, Pierpaoli C. RESTORE: robust estimation of tensors by outlier rejection. Magnetic Resonance in Medicine. 2005; 53:1088–1095. [PubMed: 15844157]

42. Collier Q, Veraart J, Jeurissen B, den Dekker AJ, Sijbers J. Iterative reweighted linear least squares for accurate, fast, and robust estimation of diffusion magnetic resonance parameters. Magnetic Resonance in Medicine. 2014

43. Tournier JD, Mori S, Leemans A. Diffusion tensor imaging and beyond. Magnetic Resonance in Medicine. 2011; 65:1532–1556. [PubMed: 21469191]

44. Perrone D, Aelterman J, Pižurica A, Jeurissen B, Philips W, Leemans A. The effect of Gibbs ringing artifacts on measures derived from diffusion MRI. NeuroImage. 2015; 120:441–455. [PubMed: 26142273]

45. Veraart J, Fieremans E, Jelescu IO, Knoll F, Novikov DS. Gibbs ringing in diffusion MRI. Magnetic resonance in medicine. 2015 DOI: 10.1002/mrm.25866.

46. Block KT, Uecker M, Frahm J. Suppression of MRI truncation artifacts using total variation constrained data extrapolation. International journal of biomedical imaging. 2008; 2008

47. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. Magnetic resonance in medicine. 2011; 65:480–491. [PubMed: 21264937]

48. Morozov VA, Nashed Z, Aries A. Methods for solving incorrectly posed problems. Springer. 1984

49. Manjón JV, Coupé P, Concha L, Buades A, Collins DL, Robles M. Diffusion weighted image denoising using overcomplete local PCA. PloS one. 2013; 8:e73021. [PubMed: 24019889]
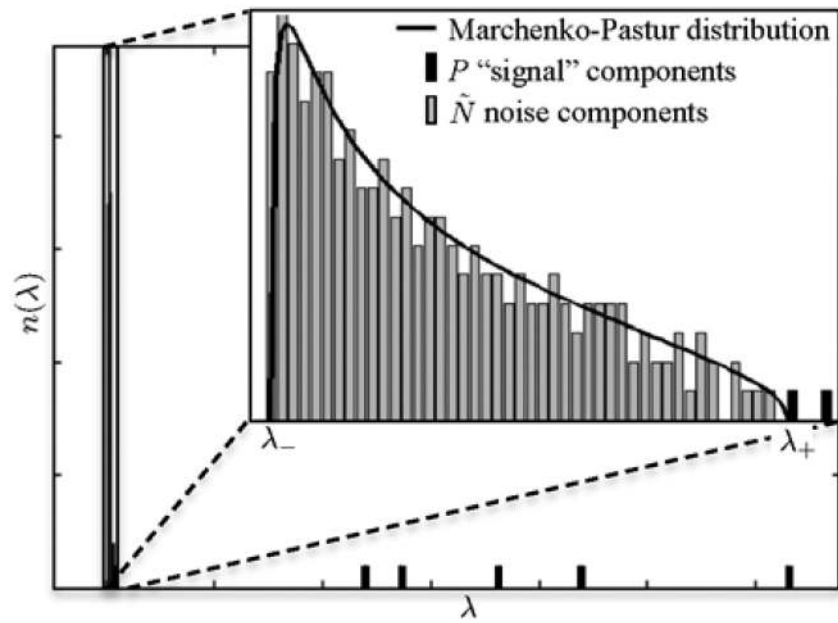
Figure 1.
Eigenvalue spectrum of sample covariance matrix of simulated DW data. MP distribution is superimposed. The P significant eigenvalues are separated from the MP "sea".
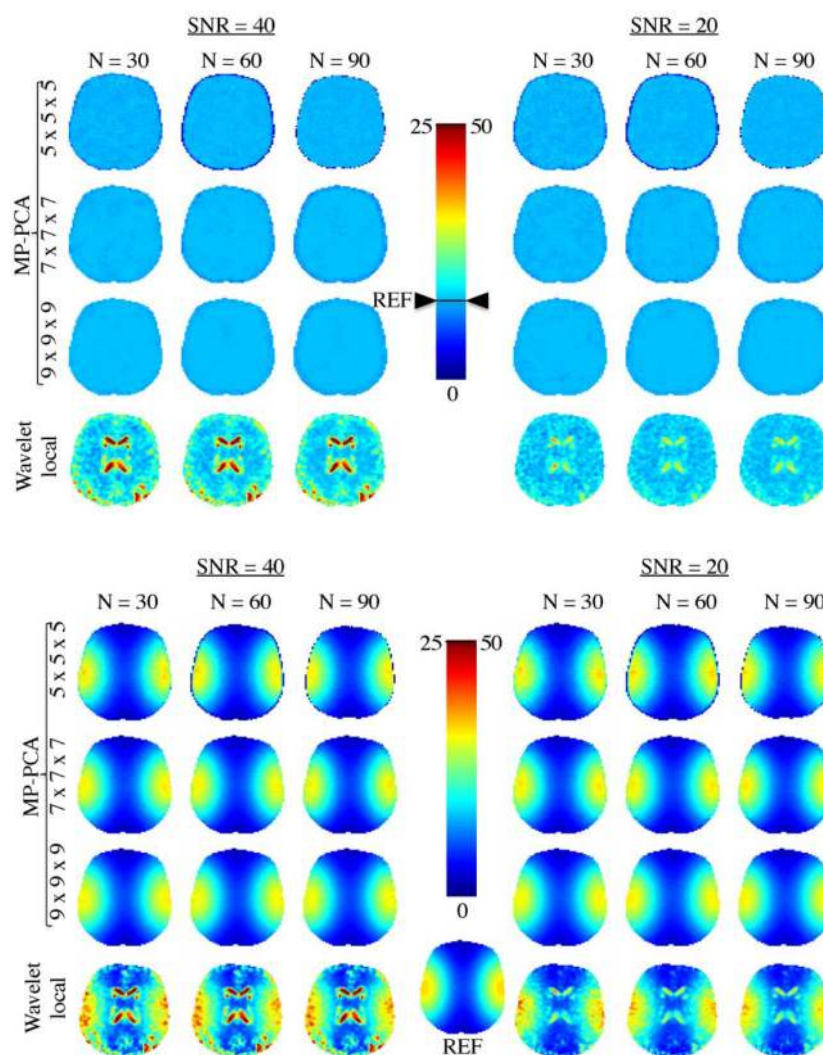
Figure 2.
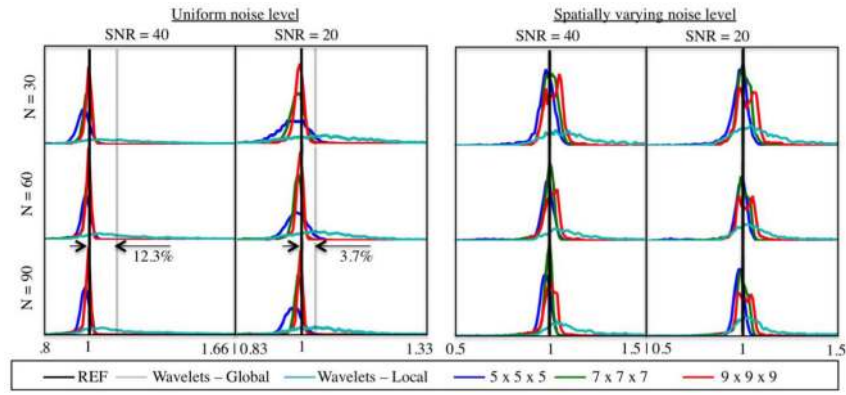Schematic overview of the MP-PCA algorithm described in section "Noise level estimation".

Figure 3.
The bias correction framework relies on the accurate estimation of the expectation value (top row) and standard deviation (bottom row) of the distribution. The theoretical (μ) and the estimated $(\hat{\mu})$ expectation values are shown in the top row in dashed blue and red curves, respectively. The magnitude noise standard deviation $\hat{\sigma}$ and underlying Gaussian noise level $\hat{\sigma}_c$, estimated from a 125×90 matrix X of which each entry X_{ij} is a sample for a nc-χ distribution with noise free signal $\nu = $ SNR and $\sigma_c = 1$, are shown as function of the SNR for L=1, 8, 16, and 32, before and after correction for the nc-χ bias, respectively (bottom row). The solid lines (red and green) are the average over 500 independent noise realizations, whereas the shaded colored areas show the confidence interval bounded by the standard deviation.
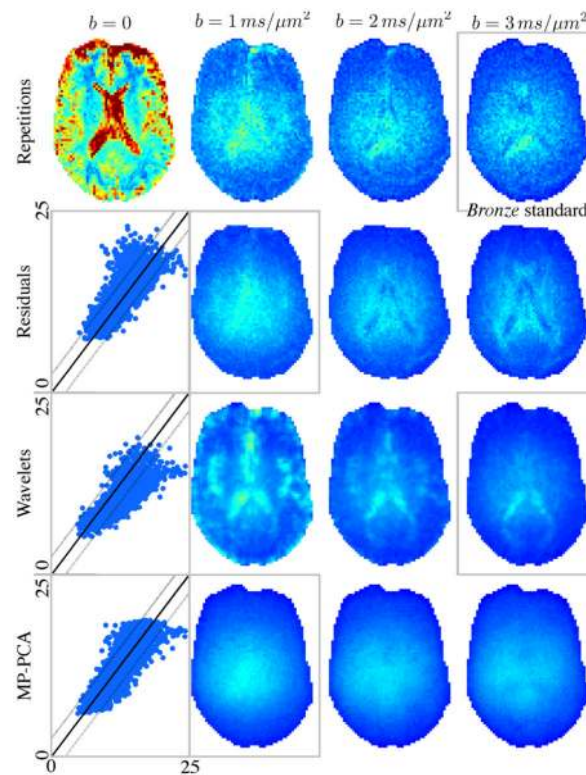
Figure 4.
Noise maps estimated from Rician distributed simulated data with spatially uniform noise level. Our proposed method, MP-PCA, was evaluated as function of SNR, number of DW acquisitions N, and 3d sliding window size M. Comparison with a local wavelet-based noise estimator (4) demonstrates the superior performance of MP-PCA, in terms of both accuracy and precision, regardless the underlying signals and noise-uniformity. The advantages of MP-PCA are obvious already by having no artifactual anatomical features, that are present in the wavelet-based method.
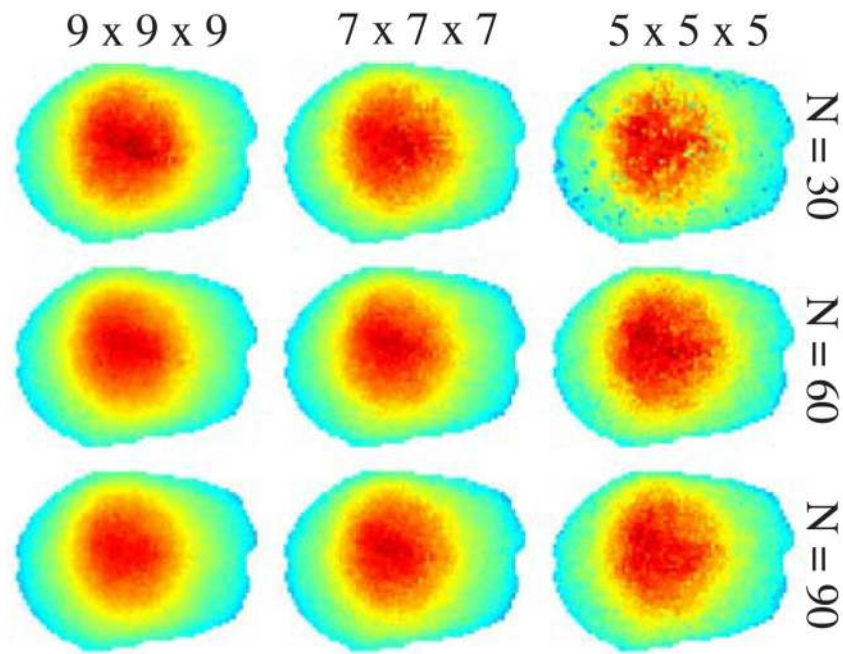
Figure 5.
Noise maps estimated from Rician distributed simulated data with spatially varying noise level. Our proposed method, MP-PCA, was evaluated as function of SNR, number of DW acquisitions N, and 3d sliding window size M. Comparison with a local wavelet-based noise estimator (4) demonstrates the superior performance of MP-PCA, in terms of both accuracy and precision, regardless the underlying signals and noise-uniformity. The advantages of MP-PCA are obvious already by having no artifactual anatomical features, that are present in the wavelet-based method.
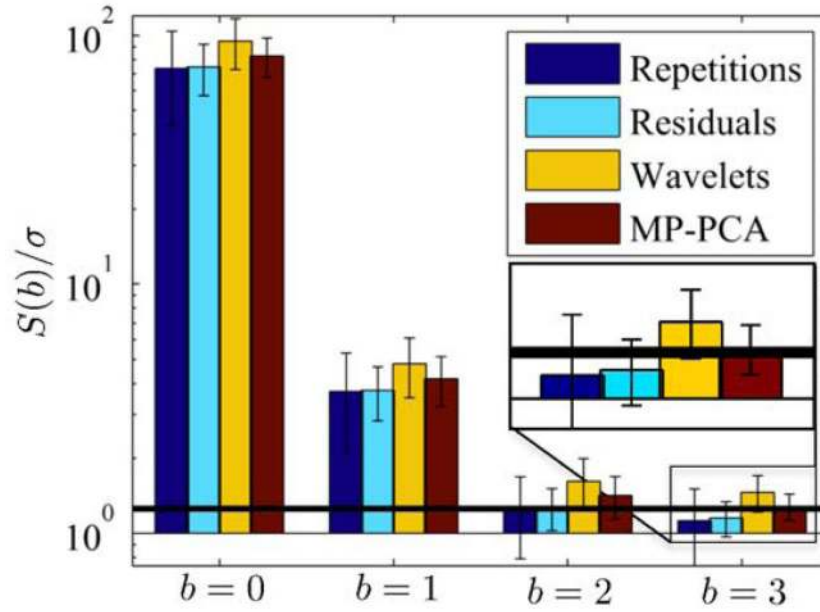
Figure 6.
Histograms of the relative error of the noise maps estimated from Rician distributed simulated data with spatially uniform (left) and varying noise levels (right). The proposed MP-PCA outperforms local and global wavelet-based noise estimators, both in terms of accuracy and precision. The precision of MP-PCA tend to increase with the size of the sliding window. However, the resolution of the largest window becomes insufficient to capture local variability in noise level in case of spatially varying noise, thereby biasing the results (red curve). This points at the existence of an optimal sliding window size M for a given coil array.
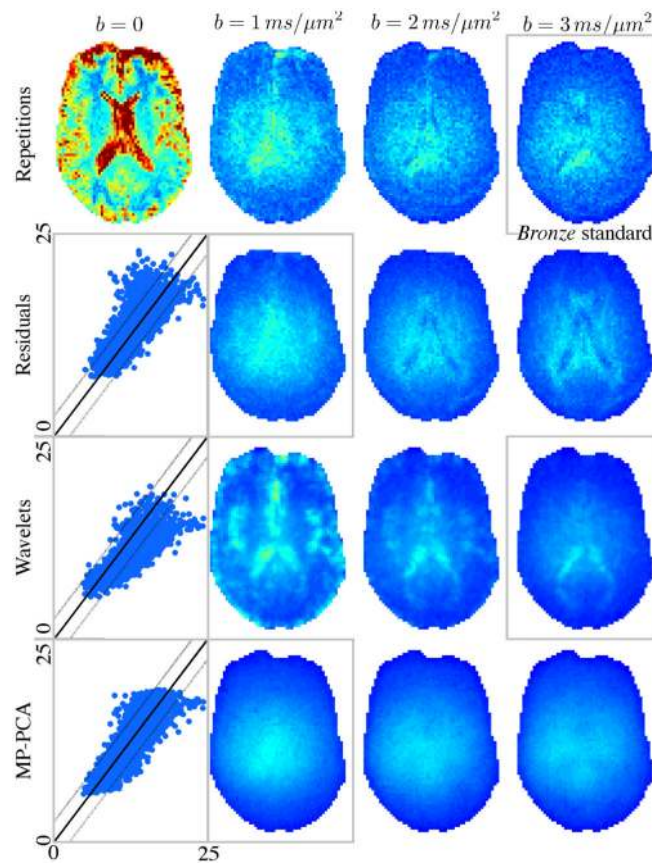
Figure 7.
Experiments show qualitatively and quantitatively (see Table I) that our proposed method (MP-PCA) outperforms the residual and wavelet-based methods if compared to the most reliable proxy for the ground truth. This map was computed from 23 repetitions at the highest *b*-value to suppress CSF-related physiological noise and is referred to as the *bronze standard*. In the scatter plots, the best estimated noise maps (gray boxes) of the corresponding row are shown voxel-by-voxel against our bronze standard (gray dots). The black line of unity and the dashed lines (offset of ±10% of window range to unit line) are shown for reference. The highest correspondence is observed for the MP-PCA method. Indeed all MP-PCA maps correlate well with the bronze standard, correlation coefficients are all 0.90, whereas for the residual-based method, the coefficients range between 0.81 and 0.83, and for the wavelet-based technique, they range between 0.41 and 0.88. The images are ranged by [0 50].
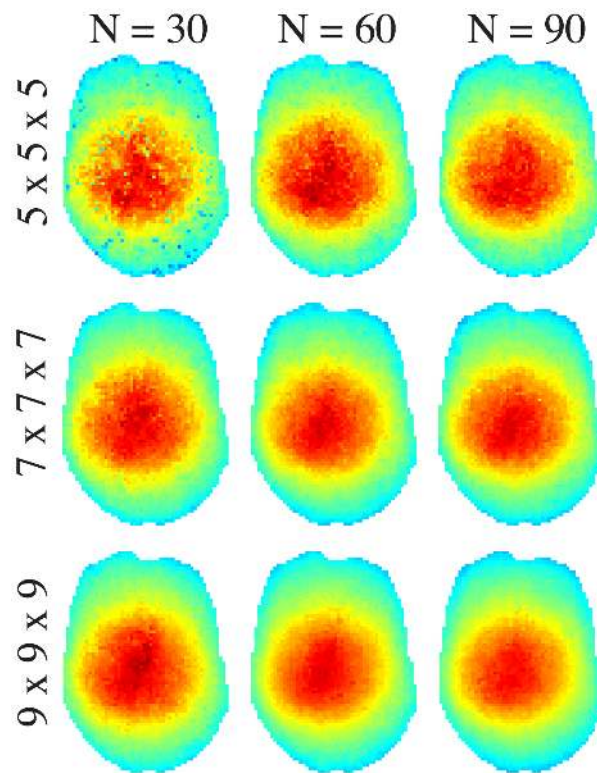
Figure 8.
Weak dependence of MP-PCA noise maps on N and sliding window size is shown by evaluation of subets of $b = 1$ ms/μm^2-shell. The images are ranged by [0 20].
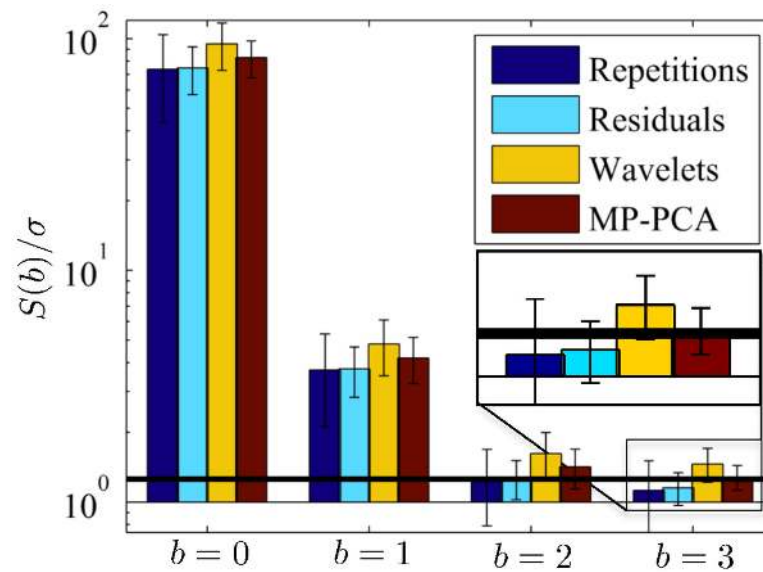
Figure 9.
Average CSF signal as function of the b-value [ms/μm^2]. The DW data was first normalized by the noise maps estimated with the different techniques. The error bars indicate the standard deviation over the different DW images present for each b-value. The black line shows the noise floor, i.e. $\sqrt{\pi/2}$ for the normalized data. The CSF signal is suppressed by > 99.99% at b = 3 ms/μm^2 and can therefore be assumed to be *noise-only*.

Table I

Quantitative comparison of different noise estimation techniques as function of SNR, *N*, and, if applicable, sliding window size. MRE = mean relative error; SRE = standard deviation of the relative error. The MRE for MP-PCA is about 1%, an order of magnitude below MRE for wavelet-based methods.

Spatially uniform

| SNR | N | MP-PCA | | | | | | WAVELET | | | |
| | | 5 × 5 × 5 | | 7 × 7 × 7 | | 9 × 9 × 9 | | local | | global | |
		MRE	SRE	MRE	SRE	MRE	SRE	MRE	SRE	MRE	SRE
20	30	−0.0209	0.0406	−0.0090	0.0188	−0.0030	0.0117	0.0614	0.1081	0.0376	0.0126
	60	−0.0135	0.0308	−0.0084	0.0125	−0.0025	0.0097	0.0537	0.0838	0.0370	0.0134
	90	−0.0179	0.0239	−0.0073	0.0113	−0.0023	0.0083	0.0503	0.0766	0.0366	0.0148
40	30	−0.0227	0.0145	−0.0031	0.0088	0.0035	0.0061	0.1601	0.1187	0.1224	0.0153
	60	−0.0126	0.0113	−0.0014	0.0071	0.0014	0.0052	0.1450	0.1075	0.1230	0.0154
	90	−0.0135	0.0100	−0.0017	0.0065	−0.0007	0.0055	0.1368	0.1014	0.1247	0.0155

Spatially varying

| SNR | N | MP-PCA | | | | | | WAVELET | | | |
| | | 5 × 5 × 5 | | 7 × 7 × 7 | | 9 × 9 × 9 | | local | | global | |
		MRE	SRE	MRE	SRE	MRE	SRE	MRE	SRE	MRE	SRE
20	30	−0.0128	0.0385	0.0075	0.0363	0.0259	0.0569	0.0809	0.1448		
	60	−0.0075	0.0324	0.0021	0.0330	0.0193	0.0499	0.0852	0.1606		
	90	−0.0172	0.0280	−0.0031	0.0301	0.0104	0.0453	0.0754	0.1425		
40	30	−0.0120	0.0393	0.0101	0.0376	0.0291	0.0487	0.1680	0.2534		
	60	−0.0102	0.0328	0.0026	0.0255	0.0166	0.0384	0.2029	0.3155		
	90	−0.0193	0.0302	−0.0040	0.0220	0.0084	0.0364	0.1965	0.2964		