Santiago Aja-Fernández

Selected papers on statistical noise analysis in MRI

May 5, 2015

LPI, Laboratorio de Procesado de Imagen
Universidad de Valladolid

Preface

This book is a compendium of the main contributions that our lab has done in the field of noise analysis in MRI along the last 8 years. Since many different approaches have been proposed, we think that it is important to have a chronological point of view to fully understand the different approaches. This book is the seed of another book in which the journals are rewritten with the same notation to give a deeper unity to the work.

Valladolid, Spain
Feb. 2015
Santiago Aja Fernández

Contents

Part III   Noise estimators for pMRI

Part I
Noise Models

Statistical noise models for Magnetic Resonance Imaging

Santiago Aja-Fernández, Antonio Tristán-Vega*

Abstract Many image processing applications within MRI are grounded on stochastic methods based on the prior knowledge on the statistics of noise. The ubiquitous Gaussian model provides a poor fitting for medium-low SNRs, yielding to the use of Rician statistics: the noise in MRI has been traditionally modeled as a stationary process governed by a Rician distribution with constant noise power at each voxel. Modern MRI systems turn this model questionable, making it necessary to develop into more complex patterns. We aim at comprehensively reviewing the main statistical rationales and formulations for the noise in MRI lately found in the literature. We attend to three different criteria: the first-order, voxel-wise probability law, the possible spatial variability of the parameters of such distribution, and the possible noise interdependences between neighboring voxels. Several applications using statistical methods are overviewed, discussing the implications each of the models has on them. Finally, we explore the applicability of the surveyed models to some MRI protocols commonly used. Whereas many parallel and nonparallel acquisitions like GRAPPA and SENSE may be fitted into one of the existing models, other nonlinear reconstruction procedures are lacking a proper noise characterization.

1 Introduction

Magnetic Resonance (MR) data is known to be affected by several sources of quality deterioration, due to limitations in the hardware, scanning times, movement of patients, or even the motion of molecules in the scanning subject. One source of degradation that affects most of the acquisitions is noise.

The term *noise* in MR can have different meanings depending on the context. For example, it has been applied to degradation sources such as physiological and respiratory distortions in some MR applications and acquisitions schemes [1, 2]. Even acoustic sources (the sound produced by the pulse sequences in the magnet) are sometimes referred to as *noise* [3]. In this paper we will obviate these issues, focusing on the thermal noise introduced during data acquisition. The principal source of thermal noise in most MR scans is the subject or object to be imaged, followed by electronics noise during the acquisition of the signal in the receiver chain. It is produced by the stochastic motion of free electrons in the RF coil, which is a conductor, and by eddy current losses in the patient, which are inductively coupled to the RF coil.

The presence of noise over the acquired MR signal is a problem that affects not only the visual quality of the images, but also may interfere with further processing techniques such as segmentation, registration or fMRI analysis [4, 5, 6]. There are different ways to cope with this degradation, but, due to the random nature of thermal noise, a probabilistic modeling is a proper and powerful solution. The accurate modeling of signal and noise statistics in MR data usually underlies the tools for processing and interpretation within Magnetic Resonance Imaging (MRI).

One of the most direct approaches to cope with acquisition noise in MRI (of course, not the only one) is signal estimation via noise removal. Traditionally, noise filtering techniques in different fields have been based on a well-defined prior statistical model of data, usually a Gaussian model. Noise models in MRI have allowed the natural extension of many well known techniques to cope with features specific of MRI. Many examples can be found in the literature, such as the Conventional Approach (CA) [4], Maximum Likelihood (ML) [7], linear estimators [8, 9, 10], or adapted non-local mean schemes [11, 12, 13]. However, an accurate noise modeling may be useful in MRI not only for filtering purposes, but also for many other processing techniques. For instance, Weighted Least Squares methods to estimate the Diffusion Tensor (DT) have proved to be nearly optimal when the data follows a Rician [14] or a non-central Chi (nc-χ) distribution [15]. Other approaches for the estimation of the DT also assume an underlying Rician model of the data: ML and Maximum a Posteriori (MAP) estimation [16, 17], or sequential techniques for online estimation [18, 19] have been proposed. The use of an appropriated noise model is crucial in all these methods to attain a statistically correct characterization of the underlying signals.

Other methods in MRI processing that benefit from relying on a precise noise distribution model include automatic segmentation of regions [20, 21], compressed sensing for signal reconstruction [22, 23], and fMRI activation and simulation [24, 25, 26]. All in all, many examples in literature have shown the advantage of statistically modeling the specific features of noise for a specific typology of data.

For practical purposes, it has been usually assumed that the noise in the image domain is a zero-mean, spatially uncorrelated Gaussian process, with equal variance in both the real and imaginary parts. In case the data is acquired by several receiving coils, the exact same distribution is assumed for all of them. As a result, in single coil systems the magnitude data in the spatial domain are modeled using a station-

ary Rician distribution [5]. When multiple (independent) coils are considered and the **k**-space is fully sampled, the natural extension of the Rician model yields to a stationary noncentral-χ (nc-χ) distribution [27], whenever the different images are combined using sum of squares, the variance of noise is the same for all coils, and no correlations exist between them.

These two distributions, Rician and nc-χ, have been extensively used in the MR literature whenever a noise model is needed. However, in modern acquisition systems, they may no longer hold as reliable distributions. Interpolations due to non-Cartesian sampling, ghost-correction post-processing for acquisitions schemes such as EPI [28, 29], manufacturer-specific systems for noise and artifacts reduction, or coil uniformity correction techniques will dramatically alter spatial noise characteristics. Thus, in many practical situations, the initial assumptions on the nature of noise do not completely hold. On the other hand, even when these schemes are not used, nowadays MRI systems often collect subsampled versions of the **k**-space to speed-up the acquisitions and palliate phase distortions. In order to correct the aliasing artifacts produced by this subsampling, some reconstruction methods are to be used, the so-called Parallel MRI (pMRI) techniques [30, 31]. The most popular among them, owing to their common use in commercial devices, are Sensitivity Encoding (SENSE) [32] and GeneRalized Autocalibrated Partially Parallel Acquisitions (GRAPPA) [33], but there are many more, and lots of new emerging ones every day [30, 34].

In this paper we will review the main statistical models commonly used in MRI, under the assumption of a *direct* acquisition, i.e., we will assume that: (1) data are acquired in the **k**-space using a regular Cartesian sampling; (2) the different contributions of noise are all independent, so that the total noise in the system is the sum of the noise from each individual source; and (3) post processing schemes such as EPI correction are not applied. Though these assumptions may seem unrealistic for certain applications, they are common in the literature, and otherwise necessary to achieve a reasonable trade-off between the accuracy of the model and its generalization capabilities. Note the corrections mentioned in the previous paragraph are usually manufacturer-dependent or they even depend on the particular magnet devised, hence they will need a more in-depth study which is far from the scope of this paper. In many cases, however, such study can be derived from the general models here described.

With the aforementioned limitations, we present a comprehensive study of the noise models arising in the most common MRI protocols currently used, mainly single- and multiple-coil acquisitions without **k**-space subsampling, SENSE, and GRAPPA. In first place we review the popular Rician model as a first order statistical model for the voxel-wise distribution of noise, and introduce the nc-χ model as a necessary generalization in many multiple-coil systems. pMRI methods require a more careful study, presented in the next section, with the identical conclusion that the nc-χ generalization is mandatory for GRAPPA. The other widely accepted assumption in noise modeling, i.e., the noise power being constant for the whole image domain, is reviewed next. We reason that a non-stationary behavior must be assumed in many cases due to the inhomogeneous sensitivity of the receiving coils,

the existing correlations between the thermal noise produced by each of them, and the reconstruction process itself with GRAPPA. With SENSE, the reconstruction process introduces a different side effect, the spatial correlation of the noise pattern, which is studied in the final part of the manuscript. To conclude, we show the actual relevance of the features studied through representative examples, and discuss their impact over a number of MRI processing pipelines, explaining how to adapt the existing methods based on stationary Rician noise to more complex models.

2 First-order statistical models for fully sampled signals

2.1 Complex single– and Multiple–Coil MR signals

The $\mathbf{k}$–space data are acquired in multiple-shot acquisitions through the repeated application of excitation pulses with a different phase encoding gradient for each readout gradient. Each sampled line of the $\mathbf{k}$−space is frequency encoded, and the measured signal is uniformly sampled at the desired rate. The points in the $\mathbf{k}$-space measured from the MRI scanner are thus independent samples of the RF signal received by each coil. The primary origin of random fluctuation is the so-called thermal noise [35], whose variance depends on the following parameters:

$$\sigma^2_{\text{thermal}} \propto 4k_B T \, R_{\text{eff}} \, B_W, \tag{1}$$

where k_B is Boltzmann's constant, T is the absolute temperature of the resistor, R_{eff} is the effective resistance of the coil loaded by the object to scan, and B_W is the bandwidth of the noise-voltage detecting system.

Under the assumption that the noise affects equally to all the frequencies, it is independent for each source, and independent on the signal, it can be modeled as a complex Additive White Gaussian Noise (AWGN) process, with zero mean and variance $\sigma^2_{K_l}$ [36, 37]:

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; 0, \sigma^2_{K_l}(\mathbf{k})), \quad l = 1, \cdots, L, \tag{2}$$

with $a_l(\mathbf{k})$ the noise-free signal at the l-th coil (of a total of L coils) and $s_l(\mathbf{k})$ the received (noisy) signal. If the noise in the RF signal is considered stationary, it makes sense to consider n_l itself stationary (which implies that $\sigma^2_{K_l}(\mathbf{k}) = \sigma^2_{K_l}$ is a constant), so that we may write:

$$n_l(\mathbf{k}; 0, \sigma^2_{K_l}(\mathbf{k})) \equiv n_l(\mathbf{k}; 0, \sigma^2_{K_l}) = n_{l_r}(\mathbf{k}; 0, \sigma^2_{K_l}) + j \cdot n_{l_i}(\mathbf{k}; 0, \sigma^2_{K_l}).$$

The complex image domain is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice and at each coil. Under the assumption that the data is sampled on a Cartesian lattice, and the iDFT is applied without any kind of interpolation/filtering/apodization/zero-padding, the iDFT will be an orthogonal

linear transformation, and the noise in the complex image domain will still be Gaussian for each receiving coil:

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; 0, \sigma_l^2), \quad l = 1, \ldots, L, \tag{3}$$

where $N_l(\mathbf{x}; 0, \sigma_l^2) = N_{l_r}(\mathbf{x}; 0, \sigma_l^2) + j \cdot N_{l_i}(\mathbf{x}; 0, \sigma_l^2)$. Since the iDFT is applied to each coil, it will not introduce any correlation *per sé*. However, there may be an initial noise correlation between the receiver coils due to electromagnetic coupling [38, 39, 40]. As a consequence, the noise pattern in the complex image domain may be seen as a complex multivariate (one variable per coil) AWGN process, with zero mean and covariance matrix $\mathbf{\Sigma}$ [41]:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1L}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \cdots & \sigma_{2L}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L1}^2 & \sigma_{L2}^2 & \cdots & \sigma_L^2 \end{pmatrix}, \tag{4}$$

with $\sigma_{ij}^2 = \rho_{ij}^2 \sigma_i \sigma_j$ and ρ_{ij}^2 the coefficient of correlation between coils i-th and j-th. While this coefficient of correlation depends only on the electromagnetic coupling between coils i and j, the variance of noise for each coil may be easily predicted from that in the $\mathbf{k}-$space [36, 42, 43]:

$$\sigma_l^2 = \frac{1}{|\Omega|} \sigma_{K_l}^2, \tag{5}$$

where $|\Omega|$ is the size of the Field of View (FOV), i.e. the number of points used in the 2D iDFT.

2.2 Fully sampled single coil: the Rician distribution

For a single–coil acquisition the complex model in eq. (3) simplifies to:

$$S(\mathbf{x}) = A(\mathbf{x}) + N(\mathbf{x}; 0, \sigma^2),$$

with $N(\mathbf{x}; 0, \sigma^2) = N_r(\mathbf{x}; 0, \sigma^2) + j \cdot N_i(\mathbf{x}; 0, \sigma^2)$ a complex AWGN with zero mean and variance σ^2. The magnitude signal $M(\mathbf{x})$ is the Rician distributed envelope of the complex signal [5]:

$$M(\mathbf{x}) = |S(\mathbf{x})|. \tag{6}$$

The probability density function (PDF) of the Rician distribution is defined as [44][2]:

---

[2] To simplify the notation, the dependency with $\mathbf{x}$ has been removed from the PDFs.
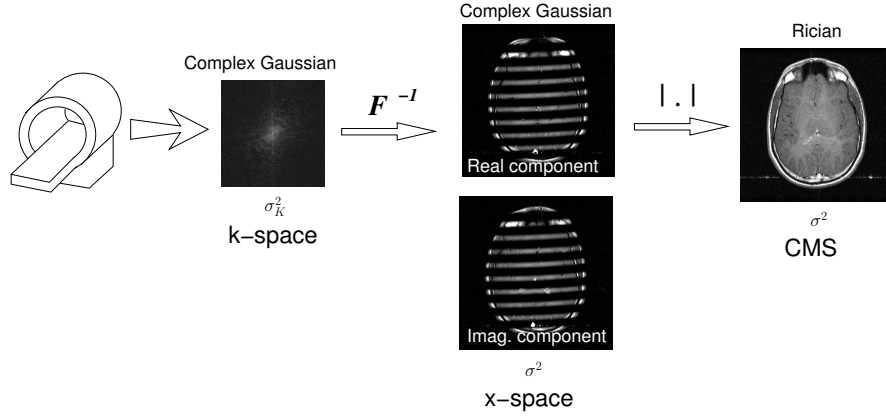
Fig. 1 Single-coil acquisition process. The data in both the **k**-space and the image domain follow a Gaussian distribution. The final signal after the magnitude is taken will follow a Rician distribution.

$$p_M(M|A_T, \sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{M^2 + A_T^2}{2\sigma^2}\right) I_0\left(\frac{A_T M}{\sigma^2}\right) u(M), \qquad (7)$$

where we call $I_n(.)$ the n-th order modified Bessel function of the first kind, $u(.)$ the Heaviside step function, and $A_T(\mathbf{x}) = |A(\mathbf{x})|$. In the image background, where the SNR is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution with PDF:

$$p_M(M|\sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{M^2}{2\sigma^2}\right) u(M). \qquad (8)$$

For the sake of illustration, a pipeline with the distributions involved in single coil acquisitions is depicted in Fig. 1.

2.3 Fully sampled, uncorrelated multiple coils: the noncentral-χ distribution

In a multiple coil system, if the **k**$-$space is fully sampled, the Composite Magnitude Signal (CMS) must be reconstructed from the L complex signals from every coil, $S_l(\mathbf{x})$, with $l = 1 \cdots, L$. One of the most used methods is the so-called Sum of Squares (SoS), which can be directly applied over eq. (3) [27, 38]:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l(\mathbf{x})|^2}. \qquad (9)$$

In an ideal scenario the variance of noise is the same for all the coils, which are assumed to produce uncorrelated samples. The covariance matrix $\mathbf{\Sigma}$ in eq. (4) is hence diagonal with identical eigenvalues:

$$\mathbf{\Sigma} = \sigma^2 \cdot \mathbf{I},$$

where $\mathbf{I}$ is the $L \times L$ identity and $\sigma^2 = \frac{1}{|\Omega|}\sigma_K^2$. Under these assumptions, $M_L(\mathbf{x})$ follows a noncentral χ (nc-χ) distribution [27, 45] with PDF:

$$p_{M_L}(M_L|A_T,\sigma,L) = \frac{A_T^{1-L}}{\sigma^2} M_L^L \exp\left(-\frac{M_L^2 + A_T^2}{2\sigma^2}\right) I_{L-1}\left(\frac{A_T M_L}{\sigma^2}\right) u(M_L), \tag{10}$$

with $A_T^2(\mathbf{x}) = \sum_{l=1}^{L} |A_l(\mathbf{x})|^2$. Obviously, for $L = 1$, the nc-χ reduces to the Rician distribution. In the background, this PDF simplifies to a central χ (c-χ) distribution with PDF:

$$p_{M_L}(M_L|\sigma,L) = \frac{2^{1-L}}{\Gamma(L)} \frac{M_L^{2L-1}}{\sigma^{2L}} \exp\left(-\frac{M_L^2}{2\sigma^2}\right) u(M_L), \tag{11}$$

which reduces to Rayleigh for $L = 1$.

2.4 Fully sampled, correlated multiple coil: the noncentral-χ approximation

The nc-χ distribution proposed in the previous section has been used to model the noise in MRI when the signals at different receiving coils are combined with SoS [27, 46, 47, 48]. However, this CMS will only show nc-χ statistics if the variance of noise is the same for all coils, and no correlation exists between them. Although it is well known that in phased array systems noise correlations do exist [41, 38, 39, 40], this effect is usually left aside due to their minimal effect and practical considerations, as stated in [27].

However, for modern acquisition systems comprising up to 32 or 64 coils, the receivers will show in general a certain coupling. This means that the noisy samples at each $\mathbf{k}-$space location will be correlated from coil to coil. Assuming such correlation is frequency-independent (i.e. the same for all $\mathbf{k}-$space samples), the linear iDFT operator will clone this exact same correlation between coils in the complex image domain, so that $\mathbf{\Sigma}$ becomes a non-diagonal, symmetric, positive definite matrix. The off-diagonal elements stand for the correlations between each pair of coils. In this case, the actual PDF is not strictly a nc-χ, though for small correlations it is expected that such model remains approximately valid [49]. Even when the nc-χ assumption is feasible, correlations will affect the number of Degrees of Freedom (DoF) of the distribution. If SoS is used, the PDF of the CMS can indeed be accurately approximated with the traditional nc-χ model in eq. (10) whenever effective
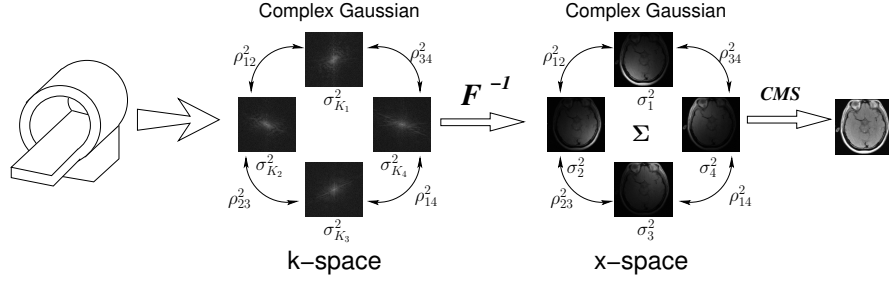
Fig. 2 Multiple-coil acquisition process. The data in both the **k**-space and the image domain follow a Gaussian distribution in each coil. The final composite magnitude signal will follow different distributions depending on the method employed to aggregate coils and on the possible correlations.

parameters (reduced L and increased σ) are used [49]:

$$L_{\text{eff}}(\mathbf{x}) = \frac{A_T^2(\mathbf{x})\,\text{tr}\,(\mathbf{\Sigma}) + (\text{tr}\,(\mathbf{\Sigma}))^2}{\mathbf{A}^*(\mathbf{x})\,\mathbf{\Sigma}\,\mathbf{A}(\mathbf{x}) + ||\mathbf{\Sigma}||_F^2}; \tag{12}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\text{tr}\,(\mathbf{\Sigma})}{L_{\text{eff}}(\mathbf{x})}, \tag{13}$$

with $||.||_F$ the Frobenius norm and $\mathbf{A}(\mathbf{x}) = [A_1(\mathbf{x}), A_2(\mathbf{x}), \cdots, A_L(\mathbf{x})]^T$. Note that, with this approximation, the $L_{\text{eff}}(\mathbf{x})$ becomes a non integer number.

Thus, for multiple correlated coils the nc-χ model is just an approximation of the real distribution, and effective parameters must be considered. The parameters of the final distribution are signal-dependent, and therefore they become harder to estimate than simpler models [50].

For the sake of illustration, a pipeline with the distributions involved in multiple-coil acquisitions is depicted on Fig. 2. Note the correlations between coils are the same in both the **k**-space and the image domain. These correlations are hardware-dependent and thus inevitable.

3 First-order statistical models for pMRI acquistions

In the previous section we have reviewed the noise model for multiple–coil systems when the $\mathbf{k}-$space is fully sampled. However, pMRI protocols usually increase the acquisition rate by subsampling the $\mathbf{k}-$space data [30, 31], while reducing phase distortions when strong magnetic field gradients are present. The immediate effect of the $\mathbf{k}-$space subsampling is the appearance of aliased replicas in the image domain retrieved at each coil. In order to suppress or correct this aliasing, pMRI combines the redundant information from several coils to reconstruct a single non-aliased image domain.

The previously presented Rician and nc-χ models do not necessarily hold in this case. Depending on the way the information from each coil is combined, the statistics of the image will follow different distributions. It is therefore necessary to study the behavior of the data for a particular reconstruction method. In this paper we focus on two of the most popular methods, SENSE [32] and GRAPPA [33], in their most basic formulation.

In the following sections we will assume that $s_l^{\mathcal{S}}(\mathbf{k})$ is the subsampled signal at the l-th coil of the $\mathbf{k}-$space, $S_l^{\mathcal{S}}(\mathbf{x})$ is the subsampled signal in the image domain, and r is the subsampling rate. Since $s_l^{\mathcal{S}}(\mathbf{k})$ is just a subsampled version of the $\mathbf{k}-$space signal, it is still corrupted with AWGN. If the iDFT is directly applied to the subsampled signal, we will have an AWGN process [43] with variance (compare with eq. (5)):

$$\sigma_l^2 = \frac{r}{|\Omega|}\sigma_{K_l}^2.$$

with $|\Omega|$ the final number of pixels in the FOV. Note the final noise power is greater than in the fully sampled case due to the reduced $\mathbf{k}-$space averaging, as it will be the case with SENSE (see below). On the contrary, the iDFT may be computed after zero-padding the missing (not sampled) $\mathbf{k}-$space lines, and then we have [42]:
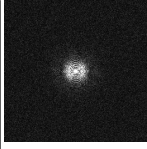
$$\sigma_l^2 = \frac{1}{|\Omega| \cdot r}\sigma_{K_l}^2.$$
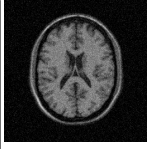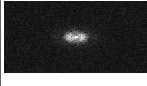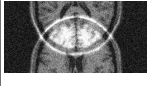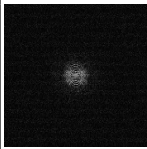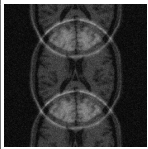
In the latter case the noise power is reduced with respect to the fully sampled case, since we average exactly the same number of samples but only 1 of each r of them contributes a noise sample (this will also be the case with GRAPPA), see Table 1 for a more thorough description. Finally, note that although the level of noise is smaller in GRAPPA due to the zero padding, the SNR does not increase, due to a reduction of the level of the signal.

3.1 Statistical model in GRAPPA reconstructed images

GRAPPA estimates the missing lines in a subsampled $\mathbf{k}-$space by linear interpolation of the complex data [30, 33, 34]. While the sampled data $s_l^{\mathcal{S}}(\mathbf{k})$ remain the same, the reconstructed lines $s_l^{\mathcal{R}}(\mathbf{k})$ are estimated through a linear combination of the existing nearby samples from all the available coils. Given a neighborhood $\eta(\mathbf{k})$ of $\mathbf{k}$, the interpolation reads:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} \sum_{\mathbf{c} \in \eta(\mathbf{k})} s_m^{\mathcal{S}}(\mathbf{k} - \mathbf{c})\omega_{ml}(\mathbf{c}), \tag{14}$$

where $\omega_{ml}(\mathbf{c})$ is a complex coefficient weighting the contribution of the measured signal at the m-th coil in the interpolation of a missing line at the l-th coil, given an offset $\mathbf{c}$ between the measured and the missing samples. These coefficients are

Noise relations				
k-space	Parameters		**x**-space	Relation
	Fully sampled, $\sigma^2_{K_l}$ **k**-size: $|\Omega|$			$\sigma^2_l = \frac{1}{|\Omega|}\sigma^2_{K_l}$, **x**-size: $|\Omega|$
	Subsampled r, $\sigma^2_{K_l}$ **k**-size: $|\Omega|/r$			$\sigma^2_l = \frac{r}{|\Omega|}\sigma^2_{K_l}$, **x**-size: $|\Omega|/r$
	Subsampled r, $\sigma^2_{K_l}$ **k**-size: $|\Omega|$ (zero padded)			$\sigma^2_l = \frac{1}{|\Omega|\cdot r}\sigma^2_{K_l}$, **x**-size: $|\Omega|$

Table 1 Relations between the variance of noise in complex MR data for each coil in the **k**-space and the image domain.

learnt from the low-frequency spectrum (as it will show the highest SNR), the so-called Auto Calibration Signal (ACS) lines, which are sampled at the Nyquist rate. Breuer *et al.* in [51] pointed out that eq. (14) can be rewritten using the *convolution* operator:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} s_m^{\mathcal{S}}(\mathbf{k}) \circledast \mathsf{w}_{ml}(\mathbf{k}), \tag{15}$$

where $\mathsf{w}_{ml}(\mathbf{k})$ is a convolution kernel that can be easily built from the GRAPPA weights $\omega_{ml}(\mathbf{k})$. For the sake of simplicity in the analysis, these weights are usually considered as constant (non-stochastic). Since a (circular) convolution in the **k**-space is equivalent to a product in the image domain, we can write:

$$S_l^{\mathcal{R}}(\mathbf{x}) = |\Omega| \sum_{m=1}^{L} S_m^{\mathcal{S}}(\mathbf{x}) \times W_{ml}(\mathbf{x}) \tag{16}$$

$$= \underbrace{|\Omega| \sum_{m=1}^{L} A_m^{\mathcal{S}}(\mathbf{x}) \times W_{ml}(\mathbf{x})}_{\text{Reconstructed Signal}} + \underbrace{|\Omega| \sum_{m=1}^{L} N_m(\sigma_n^2) \times W_{ml}(\mathbf{x})}_{\text{Gaussian Noise}} \tag{17}$$

$$= A_l^{\mathcal{R}}(\mathbf{x}) + N_l^{\mathcal{R}}(\mathbf{x}), \tag{18}$$

with $W_{ml}(\mathbf{x})$ the 2D iDFT of $\mathsf{w}_{ml}(\mathbf{k})$. This result assumes the signal plus noise model in each coil used in the previous section, a Cartesian sampling of the $\mathbf{k}$-space, and the GRAPPA weights are considered as constant (non-stochastic). The first important conclusion is that the noise power at each image location $\mathbf{x}$ will be different, since $W_{ml}(\mathbf{x})$ is spatially variant. Since the convolution in eq. (15) is again a linear operator, the noise in the image domain is still an AWGN process. In addition, even assuming that the coils are initially independent, the signals $S_l^{\mathcal{R}}(\mathbf{x})$ will become correlated when the signals from each coil are mixed through $\omega_{ml}(\mathbf{c})$.

The composite magnitude image $M_L(\mathbf{x})$ can be obtained using SoS as shown in eq. (9). Following a similar reasoning to the one for fully sampled correlated coils, we can conclude that the resultant distribution is not strictly a nc-χ. Again, it can be modeled as such with a small error if effective values are taken into account [42]:

$$L_{\text{eff}}(\mathbf{x}) = \frac{A_T^2 \operatorname{tr}\left(\mathbf{C}_X^2\right) + \left(\operatorname{tr}\left(\mathbf{C}_X^2\right)\right)^2}{\mathbf{A}^* \, \mathbf{C}_X^2 \, \mathbf{A} + ||\mathbf{C}_X^2||_F^2}; \tag{19}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\operatorname{tr}\left(\mathbf{C}_X^2\right)}{L_{\text{eff}}}, \tag{20}$$

where $\mathbf{C}_X^2(\mathbf{x}) = \mathbf{W}(\mathbf{x}) \, \boldsymbol{\Sigma} \, \mathbf{W}^*(\mathbf{x})$ is the covariance matrix of the interpolated data at each spatial location; $\mathbf{A}(\mathbf{x}) = [A_1^{\mathcal{R}}(\mathbf{x}), A_2^{\mathcal{R}}(\mathbf{x}), \cdots, A_L^{\mathcal{R}}(\mathbf{x})]^T$ is the noise-free reconstructed signal; $A_T^2(\mathbf{x}) = \sum_{i=1}^{L} |A_i^{\mathcal{R}}(\mathbf{x})|^2$, and $\mathbf{W}(\mathbf{x})$ is a matrix arranged by the set of complex weights $W_{ml}(\mathbf{x})$:

$$\mathbf{W}(\mathbf{x}) = \begin{pmatrix} W_{11}(\mathbf{x}) & W_{12}(\mathbf{x}) & \cdots & W_{1L}(\mathbf{x}) \\ W_{21}(\mathbf{x}) & W_{22}(\mathbf{x}) & \cdots & W_{2L}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ W_{L1}(\mathbf{x}) & W_{L2}(\mathbf{x}) & \cdots & W_{LL}(\mathbf{x}) \end{pmatrix}. \tag{21}$$

The reduced number of DoF in the nc-χ model is originated by the correlation and inhomogeneous variance of the complex Gaussians, i.e. by $\mathbf{C}_X^2$. In GRAPPA, this artifact mainly comes from the interpolation matrix $\mathbf{W}$ and not from the covariance matrix $\boldsymbol{\Sigma}$.

Summarizing, the nc-χ model does not hold for GRAPPA reconstructed data. However, this distribution can be used as a good approximation of the actual one. For this approximation to hold, effective parameters have to be considered which represent an equivalent, non-subsampled configuration with a smaller number of coils (DoF) and, consequently, a greater level of noise. Note the effective parameters L_{eff} and σ_{eff}^2 are signal dependent (hence spatial-dependent), which is further discussed in the next section.
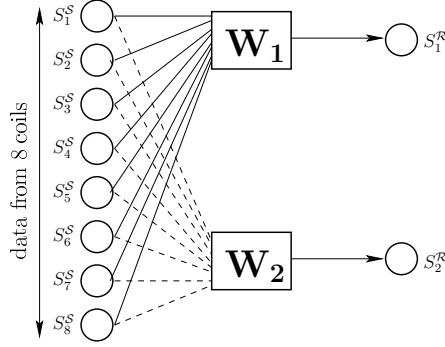
Fig. 3 Example of the SENSE interpolation for 8 coils and an acceleration factor $r = 2$.

3.2 Statistical model in SENSE reconstructed images

Most of the noise related studies in SENSE are usually done from a SNR or a g-factor (noise amplification) point of view [32, 34]. In what follows we present an equivalent reformulation, implicit in previous studies [43, 52, 46] but more coherent with the signal and noise analysis done for the other modalities reviewed in this paper.

For the sake of simplicy, let us assume that SENSE [32] is only be applied to MRI data regularly subsampled by a factor r. The reconstruction takes place in the image domain. Assuming an original size $|\Omega| = M_x \times M_y$, the subsampled signal $S_l^{\mathcal{S}}(\mathbf{x}) = S_l^{\mathcal{S}}(x, y)$ are the (complex) Fourier inverse transform of $s_l^{\mathcal{S}}(\mathbf{k})$, of size $M_x \times (M_y/r)$.

In multiple coil scanners, the image received in coil l-th, $S_l(x, y)$, can be seen as an *original image* $S_0(x, y)$ weighted by the sensitivity of that specific coil:

$$S_l(x, y) = C_l(x, y)S_0(x, y), \quad l = 1, \cdots, L \tag{22}$$

An accelerated pMRI acquisition with a factor r will reduce the matrix size of the image at every coil. The signal in one pixel at location (x, y) of l-th coil can be now written as [34]:

$$S_l(x, y) = C_l(x, y_1)S_0(x, y_1) + \cdots + C_l(x, y_r)S_0(x, y_r) \tag{23}$$

In SENSE, the reconstructed image $S^{\mathcal{R}}(x, y)$ can be seen as an estimator of the original image $S^{\mathcal{R}}(x, y) = \widehat{S_0}(x, y)$ that can be obtained from eq. (23). For instance, for $r = 2$ for pixel (x, y), $S^{\mathcal{R}}(x, y)$ is obtained as

$$\begin{bmatrix} S_1^{\mathcal{R}} \\ S_2^{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 \ \mathbf{W}_2 \end{bmatrix} \times \begin{bmatrix} S_1^{\mathcal{S}} \ \cdots \ S_L^{\mathcal{S}} \end{bmatrix}. \tag{24}$$

An example for 8 coils and $r = 2$ is depicted in Fig. 3. In matrix form for each

output pixels an arbitrary r

$$S_i^{\mathcal{R}} = \mathbf{W}_i \times \mathbf{S}^{\mathcal{S}} \quad i = 1, \cdots, r. \tag{25}$$

$$\mathbf{S}^{\mathcal{R}} = \mathbf{W} \times \mathbf{S}^{\mathcal{S}}, \tag{26}$$

with $\mathbf{W}(x, y) = [\mathbf{W}_1, \cdots \mathbf{W}_r]$ a reconstruction matrix created from the sensitivity maps at each coil. These maps, $\mathbf{C}(x, y) = [\mathbf{C}_1, \cdots, \mathbf{C}_L]$ are estimated through calibration right before each acquisition session. Once they are known, the matrix $\mathbf{W}(x, y)$ reduces to a least-squares solver for the overdetermined problem $\mathbf{C}(x, y) \times \mathbf{S}^{\mathcal{R}}(x, y) \simeq \mathbf{S}^{\mathcal{S}}(x, y)$ [32, 34]:

$$\mathbf{W}(x, y) = (\mathbf{C}^*(x, y)\mathbf{C}(x, y))^{-1}\mathbf{C}^*(x, y). \tag{27}$$

The correlation between coils may be incorporated in the reconstruction as a pre-whitening matrix for the measurements, and $\mathbf{W}(x, y)$ becomes then a weighted least squares solver with correlation matrix $\mathbf{\Sigma}$:

$$\mathbf{W}(x, y) = (\mathbf{C}^*(x, y)\mathbf{\Sigma}^{-1}\mathbf{C}(x, y))^{-1}\mathbf{C}^*(x, y)\mathbf{\Sigma}^{-1}.$$

The SNR of the fully sampled image and the image reconstructed with SENSE are related by the so-called g-factor [52, 34]:

$$\text{SNR}_{\text{SENSE}} = \frac{\text{SNR}_{\text{full}}}{\sqrt{r \cdot g}} \tag{28}$$

However, we will focus on the actual noise model underlying the SENSE reconstruction and on the final variance of noise. The final signal for each of the r reconstructed pixles, $S_i^{\mathcal{R}}$, is obtained as a linear combination of the samples in each coil, $S_l^{\mathcal{S}}$, where the noise is Gaussian distributed. Thus, the resulting signal is also Gaussian, with variance:

$$\sigma_i^2 = \mathbf{W}_i^*\mathbf{\Sigma}\mathbf{W}_i. \tag{29}$$

Since $\mathbf{W}_i$ is position dependent, i.e. $\mathbf{W}_i = \mathbf{W}_i(x, y)$, so will be the variance of noise, $\sigma_i^2(x, y)$. For further reference, when the whole image is taken into account, let us denote the variance of noise for each pixel in the reconstructed data by $\sigma_{\mathcal{R}}^2(\mathbf{x})$.

Note now that all the lines $S_i^{\mathcal{R}}$ reconstructed from the same data $S_l^{\mathcal{S}}$ will be strongly correlated, since they are basically different linear combinations of the same Gaussian variables. In that case, the covariance between $S_i^{\mathcal{R}}$ and $S_j^{\mathcal{R}}$, $i \neq j$ can be calculated as

$$\sigma_{i,j}^2 = \mathbf{W}_i^*\mathbf{\Sigma}\mathbf{W}_j, \tag{30}$$

and the correlation coefficient is derived straight forward:

$$\rho_{i,j}^2 = \frac{\sigma_{i,j}^2}{\sigma_i \sigma_j} = \frac{\mathbf{W}_i^*\mathbf{\Sigma}\mathbf{W}_j}{\sqrt{(\mathbf{W}_i^*\mathbf{\Sigma}\mathbf{W}_i)(\mathbf{W}_j^*\mathbf{\Sigma}\mathbf{W}_j)}}, \tag{31}$$

However, each r correlated lines are far enough within the final image, so we can neglect this correlation effect for processing purposes. In other pMRI modalities that carry out the interpolation in the **k**-space, such as GRAPPA, there will also be spatial correlations. However, the iDFT will sparsely distribute them among the entire image domain, considerably reducing their impact when compared to those in **x**-space reconstruction methods. Accordingly, they are usually left aside.

All in all, noise in the final reconstructed signal $S^{\mathcal{R}}(\mathbf{x})$ will follow a complex Gaussian distribution. If the magnitude is considered, i.e. $M(\mathbf{x}) = |S^{\mathcal{R}}(\mathbf{x})|$, the final CMS will follow a Rician distribution, just like single-coil systems.

We can summarize our developments as follows:

1. Subsampled multi coil MR data reconstructed with Cartesian SENSE follow a Rician distribution at each point of the image.
2. The resulting distribution is non-stationary. This means that the variance of noise will vary from point to point across the image.
3. The final value of the variance of noise at each point will only depend on the covariance matrix of the original data and on the sensitivity map.
4. Each pixel in the final image will be strongly correlated with all those pixels reconstructed from the same original data. Each pixel is correlated with $r - 1$ other pixels. Due to the distance between correlated pixels, this correlation may be left aside.

For the particular case in which there is no initial correlation between coils and all the coils have the same noise variance σ^2, we can write eq. (29) as:

$$\sigma_i^2 = \sigma^2 \times |\mathbf{W}_i|^2. \tag{32}$$

Since σ^2 is the noise variance of the subsampled data in the image domain, according to eq. (5), it is related to the original noise level without subsampling, say σ_0^2, by the subsampling rate:

$$\sigma^2 = r \cdot \sigma_0^2,$$

and therefore

$$\sigma_i = \sqrt{r} \cdot \sigma_0 \times |\mathbf{W}_i|, \tag{33}$$

which is totally equivalent to the formulations for SNR reduction in literature [32, 43].

4 Spatial variation of the noise distribution parameters: Noise maps

We have concluded in the previous sections that the noise pattern in certain multiple-coil systems may show spatially-variant, or even signal-dependent distributions. The origin of this artifact may be the initial correlation (coupling) between the receiver coils in the MRI scanner, or the reconstruction process in pMRI protocols. In all

these cases we are interested in characterizing the variability of the parameters of the noise distribution. From now on, the values of these parameters for each pixel at each coil or in the final CMS will be referred to as *noise maps*.

4.1 Noise maps for fully sampled k−spaces from multiple correlated coils

The existence of correlations between the noise in different coils compels using effective parameters for the nc-χ distribution, whose value is given in eq. (13):

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\mathbf{A}^*(\mathbf{x})\mathbf{\Sigma}\mathbf{A}(\mathbf{x}) + ||\mathbf{\Sigma}||_F^2}{A_T^2(\mathbf{x}) + \text{tr}(\mathbf{\Sigma})}.$$

This magnitude depends on the signal itself, hence it is spatially variant, as it is also the case for $L_{\text{eff}}(\mathbf{x})$. This fact makes it very difficult developing statistical models for signal processing. On the other hand, the value of L_{eff} or σ_{eff}^2 is usually not as relevant as the value of their product, $L_{\text{eff}}\,\sigma_{\text{eff}}^2$ which can be found in some of the moments of the nc-χ.

From eqs. 12 and 13, it is easy to check:

$$L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x}) = \text{tr}(\mathbf{\Sigma}), \tag{34}$$

which depends now only on $\mathbf{\Sigma}$. In the particular case in which the variance of noise is equal in each coil :

$$L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x}) = L \cdot \sigma^2$$

which is similar to the uncorrelated case.

Hence, the parameter of interest is both signal-independent and spatially-invariant, and building a noise map is worthless. On the contrary, suppose we are interested in the individual value of σ_l^2. Assuming $\sigma_l^2 = \sigma^2$ is the same for every coil, the effective variance of noise becomes largely different in the background (SNR=0) and signal (SNR$\to \infty$) areas [49]:

- Background: $\sigma_{\text{eff},B}^2 = \sigma^2\,(1 + \langle\rho^4\rangle(L-1))$.
- Signal: $\sigma_{\text{eff},S}^2 = \sigma^2\,(1 + \langle\rho^2\rangle(L-1))$.

with $\langle.\rangle$ the sample mean operator, so that:

$$\langle\rho^n\rangle = \frac{1}{L(L-1)} \sum_{i\neq j} \rho_{ij}^n.$$

Since $\rho_{ij}^2 \in [0, 1]$, the effective variance of noise in the signal areas will be greater or equal than in the background, i.e. $\sigma_{\text{eff},S}^2 \geq \sigma_{\text{eff},B}^2$. This reasoning suggests the definition of the noise ratio as:
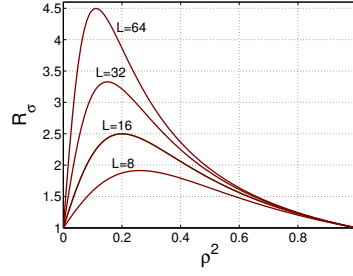
Fig. 4 Noise ratio analytically computed for different numbers of coils.

$$\mathcal{R}_\sigma = \frac{\sigma_{\text{eff},S}^2}{\sigma_{\text{eff},B}^2} = \frac{1 + \langle \rho^2 \rangle (L-1)}{1 + \langle \rho^4 \rangle (L-1)} \tag{35}$$

For the sake of illustration, this ratio is depicted in Fig. 4 for different values of the true L and the particular case $\langle \rho^n \rangle = \rho^n$. Even for small values of the correlation coefficient ρ, there is a large variability of σ_{eff}^2 between signal and background areas, which disappears when we consider $L_{\text{eff}} \cdot \sigma_{\text{eff}}^2$.

4.2 Noise maps for GRAPPA

Unlike the previous case, the GRAPPA interpolation makes the signal at each coil non-stationary even before SoS is done. According to eq. (17), the noise at the l-th coil in the image domain will follow a complex non-stationary Gaussian distribution, whose noise map for either the real or the imaginary component is defined by:

$$\sigma_l^2(\mathbf{x}) = \mathbf{W}_l^*(\mathbf{x}) \, \mathbf{\Sigma} \, \mathbf{W}_l(\mathbf{x}), \tag{36}$$

with $\mathbf{W}_l(\mathbf{x}) = [W_{1l}(\mathbf{x}), \cdots, W_{Ll}(\mathbf{x})]^T$. If there are no initial correlations between coils, this equation simplifies to:

$$\sigma_l^2(\mathbf{x}) = \sum_{m=1}^{L} \sigma_m^2 |W_{ml}(\mathbf{x})|^2. \tag{37}$$

This noise map does not depend on the signal, but on the original covariance matrix and on the GRAPPA reconstruction coefficients. Once the SoS is taken, the effective map of noise becomes signal dependent; from eq. (20):

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\mathbf{A}^* \mathbf{C}_X^2 \mathbf{A} + ||\mathbf{C}_X^2||_F^2}{A_T^2 + \text{tr}\left(\mathbf{C}_X^2\right)}.$$

If we consider now the product $L_{\text{eff}}\,\sigma_{\text{eff}}^2$ in the nc-χ distribution, eqs. 19 and 20 show that this parameter is at least signal-independent. As opposed to the fully sampled case, this product is indeed spatially variant, hence it makes sense defining a noise map:

$$L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x}) = \text{tr}\left(\mathbf{C}_X^2(\mathbf{x})\right).$$

4.3 Noise maps for SENSE

After SENSE reconstruction, the final signal follows a non-stationary complex Gaussian, which becomes a Rician when its magnitude is taken. Thus, the same parameters and noise maps can be considered for both. The final noise map will follow eq. (29):

$$\sigma_i^2(\mathbf{x}) = \mathbf{W}_i^*(\mathbf{x})\,\mathbf{\Sigma}\,\mathbf{W}_i(\mathbf{x}) \quad i = 1, \cdots, r.$$

The noise map in SENSE does not depend on the signal, but only on the original covariance matrix and on the sensitivity maps of the coils. This means that, knowing the sensitivity maps and the acceleration rate beforehand, it is possible to predict the output noise map for SENSE.

In addition, note that in the case of SENSE, the noise map is not the only interesting noise-related map to take into account. Due to the reconstruction process, there will be a high correlation between adjacent lines. A correlation map can also be defined following eq. (31).

4.4 Some practical examples
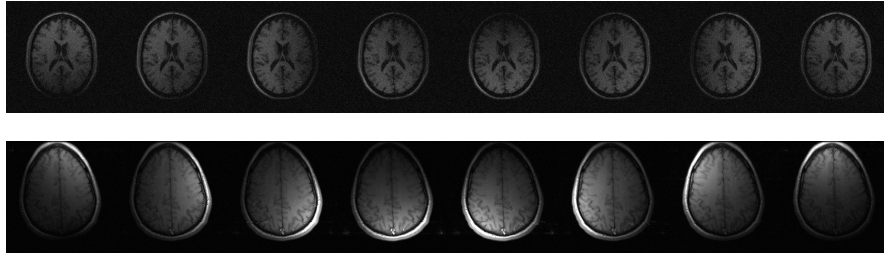
Fig. 5 MR data used for illustration: Top: Synthetic 8-coil acquisition with correlated noise. Bottom: Actual brain imaging acquisition from a GE Signa 1.5T scanner with 8 coils.

In this section we will show some of the noise maps previously defined. For the sake of illustration, two different data sets will be considered:
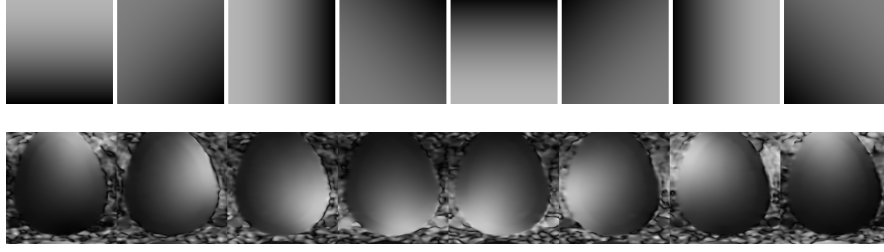
Fig. 6 Sensitivity Maps used for the experiments. Top: Synthetic sensitivity map created so that the SoS of the maps gives a constant image. Bottom: Sensitivity map estimated from an actual brain imaging acquisition in a GE Signa 1.5T scanner with 8 coils.

1. In order to know before hand all possible parameters such as $\mathbf{\Sigma}$, number of coils, A_i and so on, a synthetic phantom mimicking a parallel acquisition will be used, as shown in Fig. 5. The starting point is a 2D synthetic slice $A_0(\mathbf{x})$ in the image domain from the BrainWeb MR volume [53], with intensity values in $[0, 255]$. The average intensity value for the White Matter is 158, for the Gray Matter is 105, for the cerebrospinal fluid 36 and 0 for the background. An 8-coil system is simulated using an artificial sensitivity map coded for each coil so that $A_T^2(\mathbf{x}) = \sum_{l=1}^{8} |A_l(\mathbf{x})|^2 = A_0(\mathbf{x})$, as shown in Fig. 6 (top). The image domain for each coil is corrupted with complex Gaussian noise with $\sigma_l^2 = 100$ and a correlation coefficient of $\rho^2 = 0.05$ between all coils, so that:

$$\mathbf{\Sigma} = 100 \times \begin{pmatrix} 1 & 0.05 & \cdots & 0.05 \\ 0.05 & 1 & \cdots & 0.05 \\ \vdots & \vdots & \ddots & \vdots \\ 0.05 & 0.05 & \cdots & 1 \end{pmatrix}.$$

2. A real T1 acquisition of a brain, as shown in Fig. 5, done in a GE Signa 1.5T EXCITE, FSE pulse sequence, 8 coils, TR=500 msec, TE=13.8 msec, image size 256×256 and FOV: 20cm×20cm. We will consider the sensitivity maps estimated by the scanning software, see Fig. 6 (bottom).

First we will consider the fully sampled case (FS) for the synthetic data, using SoS to calculate the CMS. Using eq. (12) and eq. (13) we can calculate the maps for $\sigma_{\text{eff}}(\mathbf{x})$ and $L_{\text{eff}}(\mathbf{x})$, we show them in Fig. 7-(a) and 7-(b). There are basically two different areas to take into account: the background area and the signal area. The background has a smaller level of noise than the signal. The product between the two maps is a constant value for the whole image:

$$\sigma_{\text{eff}}^2(\mathbf{x}) \cdot L_{\text{eff}}(\mathbf{x}) = \text{tr}(\mathbf{\Sigma}) = \sigma_l^2 \cdot L = 100 \times 8,$$

as shown in Fig. 7-(c). From the same data in each coil, the $\mathbf{k}$-space is subsampled using an acceleration rate $r = 2$, keeping 32 ACS lines for calibration, and GRAPPA

(a) FS: $\sigma_{\mathrm{eff}}(\mathbf{x})$

(b) FS: $L_{\mathrm{eff}}(\mathbf{x})$

(c) FS: $\sqrt{L_{\mathrm{eff}}(\mathbf{x})\sigma^2_{\mathrm{eff}}(\mathbf{x})}$

(d) GRAPPA: $\sigma_{\mathrm{eff}}(\mathbf{x})$

(e) GRAPPA: $L_{\mathrm{eff}}(\mathbf{x})$
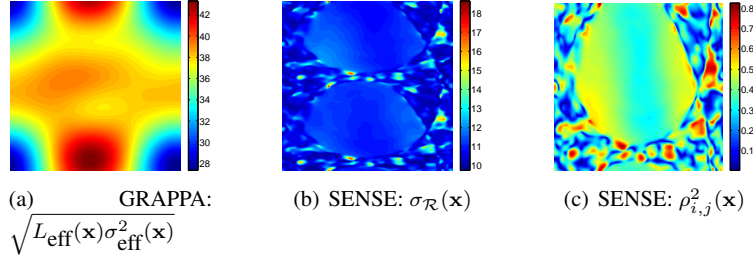

(f) GRAPPA: $\sqrt{L_{\mathrm{eff}}(\mathbf{x})\sigma^2_{\mathrm{eff}}(\mathbf{x})}$

(g) SENSE: $\sigma_{\mathcal{R}}(\mathbf{x})$

(h) SENSE: $\rho^2_{i,j}(\mathbf{x})$
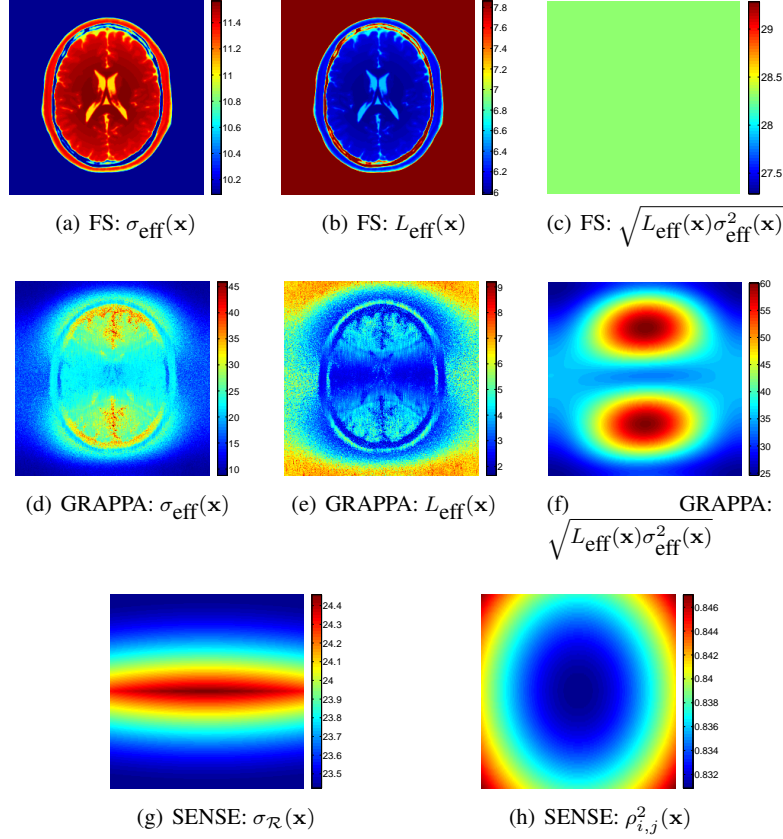
Fig. 7 Maps of noise for the synthetic example. (a) Effective noise map for the fully sampled case. (b) Effective number of coils for the fully sampled case. (c) The map of $\sqrt{L_{\mathrm{eff}}(\mathbf{x})\sigma^2_{\mathrm{eff}}(\mathbf{x})}$ for the fully sampled case. (d) Effective noise map for the GRAPPA reconstruction. (e) Effective number of coils for the GRAPPA reconstruction. (f) The map of $\sqrt{L_{\mathrm{eff}}(\mathbf{x})\sigma^2_{\mathrm{eff}}(\mathbf{x})}$ for GRAPPA reconstruction. (g) Noise map for SENSE reconstruction. (h) Coefficient of correlation for the SENSE reconstruction.

is used for reconstruction. The final CMS is obtained using SoS. Again, we can calculate the maps for noise and number of coils using eq. (20) and eq. (19). Results for GRAPPA are shown in Fig. 7-(c) and 7-(d). Note that there is a great range of values for $\sigma_{\mathrm{eff}}(\mathbf{x})$, mainly due to the GRAPPA coefficients. Most of the background lies in values around 10, while there are signal areas with levels of noise around 45. In addition, the background and the signal do not show a pattern as regular as the FS case. The product $\sigma^2_{\mathrm{eff}}(\mathbf{x}) \cdot L_{\mathrm{eff}}(\mathbf{x})$ for GRAPPA will not be a constant either, but a map with different values for different $\mathbf{x}$. The advantage of this product is that it does not depend on the signal, but only on the reconstruction coefficients and on

(a)   GRAPPA: (b) SENSE: $\sigma_{\mathcal{R}}(\mathbf{x})$ (c) SENSE: $\rho_{i,j}^2(\mathbf{x})$
$\sqrt{L_{\text{eff}}(\mathbf{x})\sigma_{\text{eff}}^2(\mathbf{x})}$

Fig. 8 Maps of noise for the real acquisition. (a) The map of $\sqrt{L_{\text{eff}}(\mathbf{x})\sigma_{\text{eff}}^2(\mathbf{x})}$ for GRAPPA reconstruction. (b) Noise map for SENSE reconstruction. (c) Coefficient of correlation for the SENSE reconstruction.

the original covariance matrix. In Fig. 7-(e) we show the map for the synthetic data, and in Fig. 8-(a) the same map is calculated for the real acquisition. For the real acquisition we have just used the GRAPPA coefficients and assumed a Σ matrix like the one in the synthetic experiment. (There are methods allowing the estimation of the actual Σ, but for the sake of illustration and comparison we will consider the same for the two experiments).
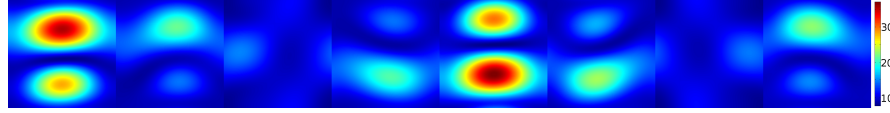
Fig. 9 Synthetic example: $\sigma_l(\mathbf{x})$, map of noise in each coil for GRAPPA before SoS.
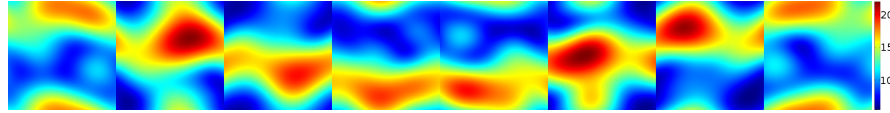
Fig. 10 Real acquisition: $\sigma_l(\mathbf{x})$, map of noise in each coil for GRAPPA before SoS.

Even assuming the same Σ matrix, note that the resultant map differs from the two cases, which gives an idea of the great influence of the GRAPPA coefficients over the final distribution of noise. Using eq. (36) we can also calculate the variance of noise of the complex Gaussian distributions in each coil before the SoS. They are depicted in Fig. 9 (synthetic data) and in Fig. 10 (real acquisition).

Finally, the **k**-space is again subsampled using an acceleration rate $r = 2$, and SENSE is used for reconstruction using the sensitivity maps of each of the data sets. Results for $\sigma_{\mathcal{R}}(\mathbf{x})$ and $\rho_{i,j}^2(\mathbf{x})$ are shown in Fig. 7-(f) and 7-(g)(synthetic data) and in Fig. 8-(b) and 8-(c) (real acquisition). Note that the uniformity of the noise map

depends highly on the sensitivity maps. For the synthetic case, the levels of noise take values in a small range. Thus, with a small error, we can assume the noise to be stationary, and work as if it were a single coil acquisition. However, for the real acquisition, the range of values is broader.

5 Discussion

While many of the features previously discussed for the noise pattern in MRI are well known, many other are not equally familiar to the image processing community. In this section we briefly discuss some important implications of the properties explained before, and provide useful guidelines for the design of image processing pipelines involving the statistical characterization of noise in MRI. The discussion is divided in three blocks corresponding to the voxel-wise (first order) statistics, the spatially variant noise patterns, and the spatial correlations (second order statistics), each of them arising challenging problems in image processing tasks such as image denoising, quantitative diffusion MRI, functional MRI, perfusion MRI, and others.

5.1 The Rician and nc-χ models in voxel-wise statistics

The Rician distribution has been widely accepted in the MRI literature as a suitable model for the noise in magnitude images [5]. In spite of this, the complexity of Rician statistics has favored the approximation of this distribution with Gaussian functions. This simplification is justified as long as for large SNRs both distributions are virtually identical. The main advantage of the Gaussian approximation relies on the huge background on image processing techniques based on this model. In a statistical sense, Gaussian corrupted signals fulfill two useful properties:

$$E\{M(\mathbf{x})\} = A_T(\mathbf{x}); \tag{38}$$
$$E\{(M(\mathbf{x}) - A_T(\mathbf{x}))^2\} = \sigma^2, \tag{39}$$

i.e. the expected value of the observation equals the magnitude of interest itself, and the expected value of the estimation error does not depend on the actual signal $A_T(\mathbf{x})$. This behavior does not hold for Rician-distributed signals, for which we have:

$$E\{M(\mathbf{x})\} = \sqrt{\frac{\pi}{2}} L_{1/2}\left(-\frac{A_T(\mathbf{x})^2}{2\sigma^2}\right)\sigma, \tag{40}$$

with $L_n(x) = {}_1F_1(-n; 1; x)$, and the estimation error in this case does depend on the actual value of $A_T(\mathbf{x})$. The use of the Gaussian approach to estimate the value $A_T(\mathbf{x})$ will introduce a systematic error (bias) in the estimation, which will decrease

when $A_T(\mathbf{x}) \gg \sigma$. Though this bias does not have an impact on the visual assessment of MRI volumes for diagnostic purposes, it will negatively affect quantitative modalities such as diffusion, perfusion, or functional MRI, especially when the SNR is decreased. One very illustrative example is the case of diffusion MRI. In the DWIs image intensities are lower for those gradient directions aligned with the major directions of diffusion. This means that the relevant information is coded in low SNR areas, for which $M(\mathbf{x})$ is systematically greater than $A_T(\mathbf{x})$ and the diffusivity will be underestimated. This effect is even more noticeable in High Angular Resolution DWI (HARDI), where the sensitizing gradients are typically stronger and the SNR is dramatically worsened due to the diffusion-induced attenuation [54].

Many authors have pointed out before this consideration, introducing the Rician statistics in the estimation of diffusion models [17, 55], curve fitting for quantitative perfusion measurements [56, 57, 58], hypothesis tests to assess the activation level in functional MRI [24, 25], or in a preprocessing (denoising) step to get rid of the aforementioned bias for the subsequent processing stages [48, 59, 60, 61]. These methods are usually based on non-linear, iterative procedures to recover $A_T(\mathbf{x})$ from $M(\mathbf{x})$ in eq. (40), or they alternatively use the Rician distribution in eq. (7) to estimate $A_T(\mathbf{x})$ from noisy measurements. This complexity can be reduced by alternatively considering the squared magnitude signal, whose relation with the squared original signal is linear [4]:

$$E\{M^2(\mathbf{x})\} = A_T^2(\mathbf{x}) + 2\sigma^2.$$

The so-called CA implies the actual signal can be directly computed from the expected value of the squared magnitude signal by simply subtracting the bias term:

$$A_T(\mathbf{x}) = \sqrt{E\{M^2(\mathbf{x})\} - 2\sigma^2}, \tag{41}$$

which has been intensively used in the MRI denoising literature [9, 11, 12, 13, 62].

On the other hand, the nc-χ model is not equally popular in the literature (see Table 2), despite of the fact that it has been proved a suitable model for many multiple coils systems, for example those using GRAPPA reconstruction. In these cases, using Rician statistics may lead to great errors [10, 50], and yet a Gaussian distribution might be a preferable model. Nevertheless, for multiple coil systems, both Rician and Gaussian models will cause a certain bias to appear in the estimation of $A_T(\mathbf{x})$, since the expected value of the magnitude signal is not equal to the magnitude to estimate:

$$E\{M_L(\mathbf{x})\} = \sqrt{\frac{\pi}{2}} L_{1/2}^{L-1} \left(-\frac{A_T(\mathbf{x})^2}{2\sigma^2} \right). \tag{42}$$

The biases induced by nc-χ signals are typically larger due to the combined effect of all the Gaussian sources (effective coils), and they are also dramatically increased when the data sets exhibit a poor SNR [15]. In this case the CA can be easily extrapolated [50]:

$$E\{M_L^2(\mathbf{x})\} = A_T^2(\mathbf{x}) + 2L\sigma^2 \Rightarrow A_T(\mathbf{x}) = \sqrt{E\{M_L^2(\mathbf{x})\} - 2L\sigma^2}, \tag{43}$$
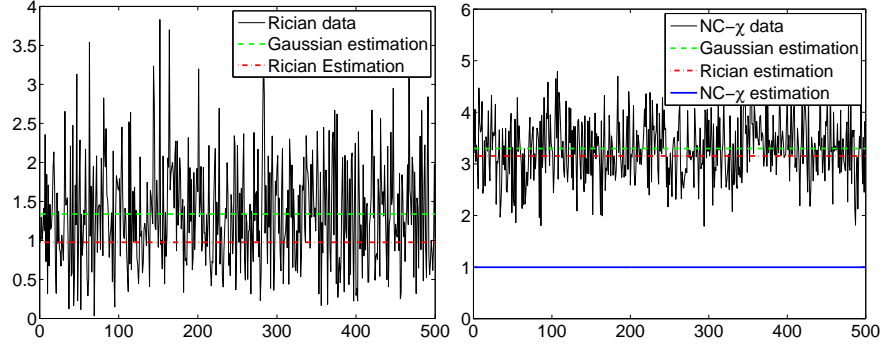
Fig. 11 1D example of estimation of A_T using different models.

while quantitative methods based on the non-squared magnitude signal for diffusion MRI [17, 55], perfusion MRI [56, 57, 58], functional MRI [24, 25], and others [48, 59, 60, 61] need a deeper reformulation. As a matter of fact, the popular Weighted Least Squares method for diffusion tensor estimation is still usable with nc-χ signals [15]. As described in Section 2.4, L and σ have to be replaced by their corresponding effective values in case that correlations between coils and/or GRAPPA interpolation are considered (and the nc-χ model becomes then an approximation). This consideration adds an extra layer of complexity to the problem, since neither L_{eff} nor σ_{eff} are known beforehand in general, and they must be estimated. Fortunately, the product $L_{\text{eff}}\sigma_{\text{eff}}^2$ is typically the relevant magnitude, since it accounts for the total amount of noise in the magnitude image instead of the independent contribution of each coil separately. Consider for example eq. (43), which will translate into:

$$A_T(\mathbf{x}) = \sqrt{E\{M_L^2(\mathbf{x})\} - 2L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x})} = \sqrt{E\{M_L^2(\mathbf{x})\} - 2\text{tr}(\boldsymbol{\Sigma})}, \quad (44)$$

Indeed, it is usually simpler estimating $L_{\text{eff}}\sigma_{\text{eff}}^2$ from a noisy data set than it is assessing each of them isolated, in a way that the estimation problem has a similar complexity to the one in the Rician case (although the equations involved are completely different, see the review in [47]).

As an illustration of how a mismatch in the choice of the noise model can affect the estimation, see Fig. 11. 500 samples of two 1D random signals were generated, following a Rician and a (non correlated) nc-χ distribution, with $A_T = 1$ and $\sigma_i = 0.8$. The signal value A_T is estimated using eq. (38), eq. (41) and eq. (44), respectively assuming a Gaussian, Rician and nc-χ model. Note that since we are working in a low SNR scenario, the mismatch between the chosen model and the real data yields to a bias in the estimation of the actual signal value.

5.2 Noise non-stationarity and noise maps

The main assumption for single coil Rician acquisitions is that the noise is stationary, and therefore a single value of σ characterizes the whole data set. This premise is extensible to multiple coils systems if the correlations between the receiving coils are neglected. In this case, two unique values of L and σ equally describe the noise properties at all imaged voxels. In case appreciable correlations do exist, however, the model has to be redefined in terms of effective nc-χ parameters that are spatially variant. The noise map for the synthetic example in Fig 7-(a) illustrates the great mismatch of noise values between the signal and background areas. Additionally, L_{eff} and σ_{eff} are not only spatially dependent but also signal-dependent. Fortunately, the product $L_{\text{eff}}\sigma_{\text{eff}}^2$ (the overall noise contribution for all coils) is constant, thus signal-independent, for the whole image domain. As previously stated, this means that most of the quantitative MRI techniques based on local moments computations will be equally valid for single and multiple coils systems whenever fully sampled $\mathbf{k}-$spaces are acquired. For example, in eq. (44) $\boldsymbol{\Sigma}$ is a constant.

Parallel reconstructions from partially sampled $\mathbf{k}-$spaces raise more serious issues. The non-stationarity is even more severe, and signal-dependent artifacts are also present. In this paper we have analyzed the effect of GRAPPA and SENSE over the noise in the final CMS, but a similar study can be carried out for other techniques, such as SMASH, PILS or non-Cartesian SENSE [30, 34]. With SENSE, the noise map mainly depends on the original level of noise and the sensitivity map of each coil. The corresponding map $\sigma(\mathbf{x})$ provides an estimation of how far from the Rician stationary case we are. In the synthetic example previously presented in Fig. 7-(f), for instance, the assumption that σ is constant for the whole image will not produce a great error, while it will considerably simplify the computations.

With GRAPPA, both the initial correlations between coils and the reconstruction interpolation favor the appearance of signal-dependent noise. Contrary to the fully sampled case, considering $L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x}) = \text{tr}\left(\mathbf{C}_X^2(\mathbf{x})\right)$ will not remove the spatial variability. Even so, $\text{tr}\left(\mathbf{C}_X^2(\mathbf{x})\right)$ depends only on $\boldsymbol{\Sigma}$ and the GRAPPA reconstruction coefficients, so that $L_{\text{eff}}\sigma_{\text{eff}}^2$ is at least signal independent and hence easier to work with [63].

The actual impact of the non-stationarity of noise in quantitative MRI will depend on whether the processing needs a prior estimation of the distribution parameters. For example, ML estimators for Rician PDFs can embed the assessment of σ at each voxel themselves [17]. The most of the algorithms for the estimation of diffusion MRI models are run voxel by voxel, and they do not require the actual knowledge of the noise distribution, mainly because they assume an underlying Gaussian distribution of noise [14, 15, 55, 64]. In all these cases, the non-stationarity is not a serious issue as long as the computations involve a unique voxel or a neighborhood small enough to consider an approximately homogeneous noise distribution whose parameters are not relevant. On the contrary, other diffusion MRI methods include some sort of spatial regularization, and in this case the spatial variability of noise urges a reformulation [22]. In other quantitative methods based on Rician statistics, though they treat each voxel independently, the computations require a

prior knowledge of the noise power; hence they will suffer from noise inhomogeneities [24, 25, 56, 57, 58]. This is the case also for denoising and other local moments-based techniques [9, 10, 11, 12, 13, 48, 60, 61, 62]: local moments are computed as sample estimates in a small neighborhood, mixing up noise samples with different statistical properties. Luckily, noise maps are smooth enough, so that the variability of the distributions parameters can be considered negligible inside those neighborhoods. In any case, the statistical model used for denoising must vary across the image voxels, as it may be checked from one of the simplest cases in eq. (44).

With these considerations, we may conclude that most of the quantitative techniques in MRI may be generalized to the non-stationary case once a proper estimate of the spatial map of distribution parameters is known. Accordingly, the problem of properly estimating $\sigma(\mathbf{x})$ (or $L_{\mathrm{eff}}(\mathbf{x})$ and $\sigma_{\mathrm{eff}}(\mathbf{x})$) earns a capital importance over the remaining applications. Many methods have been proposed for noise estimation, most of them showing a very good performance in terms of accuracy and low variance of the estimates [47]. Unfortunately, they usually assume a stationary pattern, hence they do not directly apply to those noise models accounting for non-stationarity.

Certain works have appeared addressing the different properties of noise in signal and background areas (especially when the latter are artificially removed), but they still estimate individual noise parameters for the entire MRI volume [8, 65, 66]. There have been proposals to carry out a rough estimation of non-stationary noise maps assuming a Rician models. However, most of these approaches require extra information beyond the simple magnitude signal: multiple acquisitions or different signals are required ([67, 68, 69]), a biophysical model must be defined ([69]), or even acquisition information such the estimated sensitivity of the coils is needed ([70]). This need of extra information has supposed a drawback in the usage of more complex noise models. Recently, there have been some novel proposals to estimate these non-homogeneous maps of noise out of a single image, not only in the MRI context. The most usual approaches are based on wavelet decomposition or local moments[71, 72, 73, 68, 74], on a homomorphic decomposition of the nosie components [75]. Similar efforts can be found for GRAPPA [70, 76, 77].

As a summary of this section, the traditional noise estimation methods will work "as they are" for single- and multiple-coils systems without subsampling or inter-coil noise correlations. For the other scenarios, some reformulations will be required. We have presented here how to predict the spatial pattern of noise parameters, so that the complexity of noise estimation may be dramatically reduced from the model-free approaches in [78, 73] to a problem where only one degree of freedom has to be estimated. In order to predict the spatial patterns, we need to *a priori* model, for SENSE, the sensitivity maps of each coil (which is inherent to this method), meanwhile with GRAPPA we need to characterize the interpolation weights. Once the maps are derived, local moments-based estimators of noise may be generalized by embedding a regularization step based on the prior knowledge of the spatial patterns.

Model	Application	References
Rician	Signal and noise models	[5, 44, 79]
	Noise Estimation	[7, 8, 37, 47, 80, 81, 82, 83, 84, 66, 65, 78, 85]
	Noise Filtering	[4, 7, 9, 11, 12, 13, 80, 82, 86, 87, 88, 89, 6, 59, 62, 73, 90, 91, 92]
	Segmentation	[20, 21, 37, 93]
	Gaussian Correction	[48, 60, 61, 55]
	Diffusion Tensor Estimation	[14, 16, 18, 19, 23]
	fMRI	[24, 25]
	Perfussion Analysis	[56, 57, 58]
nc-χ	Signal and noise models	[27, 49, 46]
	Noise Estimation	[27, 46, 47, 48, 94, 50, 95, 77]
	Noise Filtering	[10, 48, 96, 97, 50]
	Gaussian Correction	[48, 60]
	Diffusion Tensor Estimation	[15]
GRAPPA	Signal and noise models	[33, 42, 43, 51]
	Noise Estimation	[63, 70, 76, 77]
	Noise Filtering	[50]
SENSE	Signal and noise models	[32, 43, 35]
	Noise Estimation	[69, 70, 72, 98, 99, 100, 101, 67, 68, 75]
	Noise Filtering	[99, 100, 102]

Table 2 Some examples of usage of the models proposed in this paper. Most of them have been focus of the stationary Rician model.

To sum up, in Table 2 we summarize some of the different applications in literature that make use of the statistical models reviewed in this paper. Note that, from a practical point of view, most of the authors have assumed a simple Rician model for most applications, while procedures that take into account pMRI statistics are scarce. This gives a global idea of the potential of future research in MRI modeling, since the amount of unfinished problems is still large.

6 Conclusion

The proper modeling of the statistics of thermal noise in MRI is crucial for many image processing and computer aided diagnosis tasks. While the stationary Rician model has been the keystone of statistical signal processing in MR for years, neither the stationarity assumption nor the Rician distribution are always valid. The deviations of the actual statistics of noise with respect to the traditional model will imply very different outcomes depending on the final application, as we describe in the previous section. As a consequence, each application will need a careful choice for a proper and realistic model for the data sets at hand, attending to the need for

modeling the actual voxel-wise PDF and/or the spatial distribution of its parameters and possible second-order artifacts (noise correlations).

The existing models for noise statistics surveyed in this paper are suitable for many MRI acquisition sequences, as long as the $\mathbf{k}-$space is formed by Cartesian sampling, the phase-encoded lines independently acquired, and the image domain retrieved by means of a linear operator. On the contrary, single-shot acquisition sequences like EPI require post-processing schemes, ghost-correction, and others [28, 29]. Further post-processing stages may include customary noise reduction systems or corrections to the sensitivity inhomogeneities. In all these cases, large deviations of the noise statistics and the correlation between samples with respect to the AWGN model will appear. This might be the case also for the emerging techniques beyond pMRI, like non-linear GRAPPA [103] or compressed sensing-based reconstructions [104, 105, 106]. Even when the $\mathbf{k}-$space was contaminated with AWGN, the retrieval of the image domain is in these cases heavily non-linear, so that the Rician or nc-χ models are not necessarily appropriate.

Finally, in Table 3 a survey of the different models reviewed in this paper is carried out.

Composite Magnitude Signal				
Number of coils	**Acquisition**	**Statistical Model**	**Stationarity**	**Parameters**
1 coil	Single coil	Rician	Stationary	σ^2
Multiple coils (Uncorrelated)	No subsampling + SoS	nc-χ	Stationary	σ^2 L (Number of coils)
Multiple coils (correlated)	No subsampling + SoS	nc-χ (approx.)	Non-stationary	$\sigma^2_{\text{eff}}(\mathbf{x})$ $L_{\text{eff}}(\mathbf{x})$
Multiple coils	pMRI + SENSE	Correlated Rician	Non-stationary	$\sigma^2_{\mathcal{R}}(\mathbf{x})$ $\rho^2_{i,j}(\mathbf{x})$
Multiple coils	pMRI + GRAPPA+ SoS	nc-χ (approx.)	Non-stationary	$\sigma^2_{\text{eff}}(\mathbf{x})$ $L_{\text{eff}}(\mathbf{x})$

Table 3 Survey of noise models in the final composite magnitude signal for different acquisition schemes.

Influence of Noise Correlation in Multiple-Coil Statistical Models with Sum of Squares Reconstruction

Santiago Aja-Fernández, Antonio Tristán-Vega*

Abstract Noise in the composite magnitude signal (CMS) from multiple-coil systems is usually assumed to follow a noncentral χ distribution when Sum of Squares (SoS) is used to combine images sensed at different coils. However, this is true only if the variance of noise is the same for all coils, and no correlation exists between them. We show how correlations may be obviated from this model if effective values are considered. This implies a reduced effective number of coils and an increased effective variance of noise. In addition, the effective variance of noise becomes signal-dependent.

1 Introduction

The noncentral χ (nc-χ) distribution has been extensively used to model the noise in MRI when the signals at different receiving coils are combined as the Sum of their Squares (SoS) [27, 46, 47, 107, 48]. However, this Composite Magnitude Signal (CMS) will only show nc-χ statistics if the variance of noise is the same for all coils, and no correlation exists between them. Although it is well known that in phased array systems noise correlations do exist [38, 39, 40, 41], the effect of noise correlations has been usually left aside due to their minimal effect and practical considerations, as stated in [27]. With the break-up of modern acquisition systems comprising up to 32 or 64 coils this simplification is at stake.

We study here how correlations effect the nc-chi model, and under what assumptions this approximation is still valid. The deviations from ideal statistics with GRAPPA reconstructions due to k-space interpolation artifacts have been studied in a recent paper [42]. We [the authors] concluded that the divergence with respect to

---

the ideal demeanor can be obviated with the introduction of certain effective statistics. In a similar fashion, we show in the current paper how inter-coil correlations can be modeled by means of an effective (increased) number of coils and an effective (decreased) noise variance. The resulting model accurately fits the actual statistics found in real data sets [46].

2 Theory

2.1 Statistical Model in Multiple–Coil MR signals

The **k**-space data at each coil can be accurately described by an Additive White Gaussian Noise (AWGN) process, with zero mean and variance σ_K^2:

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; \sigma_K^2), \quad l = 1, \cdots, L \tag{1}$$

with $a_l(\mathbf{k})$ the noise-free signal and $n_l(\mathbf{k}; \sigma_K^2) = n_{l_r}(\mathbf{k}; \sigma_K^2) + j n_{l_i}(\mathbf{k}; \sigma_K^2)$ the AWGN process, which is initially assumed stationary so that σ_K^2 does not depend on **k**. The complex **x**-space is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice or volume, so the noise in the complex **x**-space is still Gaussian:

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; \sigma_n^2), \quad l = 1, \cdots, L$$

where $N_l(\mathbf{x}; \sigma_n^2) = N_{l_r}(\mathbf{x}; \sigma_n^2) + j N_{l_i}(\mathbf{x}; \sigma_n^2)$ is also a complex AWGN process (note we are assuming that there are not any spatial correlations) with zero mean and covariance matrix $\boldsymbol{\Sigma}$. If the **k**-space is fully sampled, SoS can be directly used to recover the MRI [27, 38]:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l(\mathbf{x})|^2}. \tag{2}$$

In the same fashion as [42], we consider $M_L^2(\mathbf{x})$ instead, and distinguish two different scenarios regarding the covariance matrix $\boldsymbol{\Sigma}$:

1. In the ideal scenario, the variance of noise at each coil is the same, and no correlations exist: $\boldsymbol{\Sigma}$ is therefore diagonal with identical eigenvalues:

$$\boldsymbol{\Sigma} = \sigma_n^2 \cdot \mathbf{I}$$

where $\mathbf{I}$ is the $L \times L$ identity, $\sigma_n^2 = \frac{1}{|\Omega|} \sigma_K^2$ is the original variance of noise and $|\Omega|$ is the number of points used to compute the iDFT. Since the variance of noise is the same for all coils and they are uncorrelated, $M_L^2(\mathbf{x})$ follows a noncentral χ^2 (nc-χ^2) distribution [45] with Probability Density Function (PDF):

$$p_X(x|A_T, \sigma_n, L) = \frac{A_T^{1-L}}{2\sigma_n^2} x^{\frac{1}{2}(L-1)} e^{-\frac{x+A_T^2}{2\sigma_n^2}} I_{L-1}\left(\frac{A_T\sqrt{x}}{\sigma_n^2}\right) u(x), \quad (3)$$

with $A_T^2(\mathbf{x}) = \sum_{l=1}^{L} |A_l(\mathbf{x})|^2$, $I_L(.)$ the Lth order modified Bessel function of the first kind, and $u(.)$ the Heaviside step function. In the background, this PDF simplifies to a central χ^2 (c-χ^2). The mean and the variance of the nc-χ^2 distribution are:

$$E\{M_L^2\} = A_T^2 + 2\sigma_n^2 L \qquad (4)$$
$$\text{Var}\{M_L^2\} = 4A_T^2\sigma_n^2 + 4L\sigma_n^4. \qquad (5)$$

2. In the general case, $\boldsymbol{\Sigma}$ is an arbitrary, symmetric, positive definite matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1L} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L1} & \sigma_{L2} & \cdots & \sigma_L^2 \end{pmatrix},$$

the off-diagonal elements standing for the correlations between each pair of coils. Though the corresponding PDF cannot be theoretically derived, the mean and variance are easily computed as [42]:

$$E\{M_L^2(\mathbf{x})\} = A_T^2 + 2\,\text{tr}\,(\boldsymbol{\Sigma})\,; \qquad (6)$$
$$\text{Var}\{M_L^2(\mathbf{x})\} = 4\,\mathbf{A}^*\boldsymbol{\Sigma}\mathbf{A} + 4\,||\boldsymbol{\Sigma}||_F^2. \qquad (7)$$

with $||.||_F$ the Frobenius norm and $\mathbf{A} = [A_1, A_2, \cdots, A_L]^T$.

This case will be found in coils systems where correlations do exist. As a consequence, the actual PDF is not strictly a nc-χ^2, though for small correlations it is expected that such model remains approximately valid. Correlations will affect the number of Degrees of Freedom (DoF) of the distribution. To deduce an *effective nc-χ^2* PDF, we use the method of moments: eqs. 4 and 5 are respectively equated to 6 and 7 to solve for the following effective parameters:

$$L_{\text{eff}} = \frac{A_T^2\,\text{tr}\,(\boldsymbol{\Sigma}) + (\text{tr}\,(\boldsymbol{\Sigma}))^2}{\mathbf{A}^*\boldsymbol{\Sigma}\mathbf{A} + ||\boldsymbol{\Sigma}||_F^2}\,; \qquad (8)$$
$$\sigma_{\text{eff}}^2 = \frac{\text{tr}\,(\boldsymbol{\Sigma})}{L_{\text{eff}}}. \qquad (9)$$

2.2 Simplified scenarios

Consider the following covariance matrix:

$$\Sigma = \sigma_n^2 \cdot \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1L} \\ \rho_{21} & 1 & \cdots & \rho_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{L1} & \rho_{L2} & \cdots & 1 \end{pmatrix} \tag{10}$$

where $\rho_{ij} \in [0, 1]$ is the correlation coefficient between coils i and j. We assume that the variance of noise $\sigma_i^2 = \sigma_n^2$ is the same for every coil. Although this premise might be not completely realistic, its study provides a simple, yet powerful insight on the behavior of the true statistics with regard to noise correlations. Under this assumption, eq. 8 and eq. 9 read:

$$L_{\text{eff}} = L \frac{A_T^2 + L\,\sigma_n^2}{A_T^2 + L\,\sigma_n^2 + \sum_{i \neq j} A_i^* A_j \rho_{ij} + \sigma_n^2 \sum_{i \neq j} \rho_{ij}^2} = L \frac{1}{C_r(L)} \tag{11}$$

$$\sigma_{\text{eff}}^2 = \sigma_n^2 \left(1 + \frac{\sum\limits_{i \neq j} A_i^* A_j \rho_{ij} + \sigma_n^2 \sum\limits_{i \neq j} \rho_{ij}^2}{A_T^2 + L\,\sigma_n^2} \right) = \sigma_n^2\, C_r(L) \tag{12}$$

where $C_r(L)$ can be seen as a *correction factor*. Since $C_r(L) \geq 1$, $L_{\text{eff}} \leq L$ and $\sigma_{\text{eff}}^2 \geq \sigma_n^2$. As expected, noise correlations induce effective numbers of coils smaller than the actual one. The effective variance of noise, on the contrary, is increased due to the reduced DoF to be averaged. A further simplification is to assume $A_i = A_j$ for all i, j:

$$L_{\text{eff}} = L \left(1 + (L - 1) \frac{A_T^2 \langle \rho \rangle + L\,\sigma_n^2 \langle \rho^2 \rangle}{A_T^2 + L\,\sigma_n^2} \right)^{-1} \tag{13}$$

$$\sigma_{\text{eff}}^2 = \sigma_n^2 \left(1 + (L - 1) \frac{A_T^2 \langle \rho \rangle + L\,\sigma_n^2 \langle \rho^2 \rangle}{A_T^2 + L\,\sigma_n^2} \right), \tag{14}$$

with $\langle . \rangle$ the sample mean operator, so that:

$$\langle \rho^n \rangle = \frac{1}{L(L-1)} \sum_{i \neq j} \rho_{ij}^n$$

Results in eqs. 13 and 14 make clear that the effective values of L and σ_n^2 do depend on the signal value $A_T(\mathbf{x})$. If $A_T^2(\mathbf{x})\langle \rho \rangle$ is comparable to $L\,\sigma_n^2\,\langle \rho^2 \rangle$, i.e. the SNR is low, the effective parameters will depend on the position $\mathbf{x}$. Therefore the data is no longer stationary, the noise power varying along with $M_L^2(\mathbf{x})$. Consider these extreme cases:

1. In the background, where no signal is present and hence SNR=0, the effective values are:
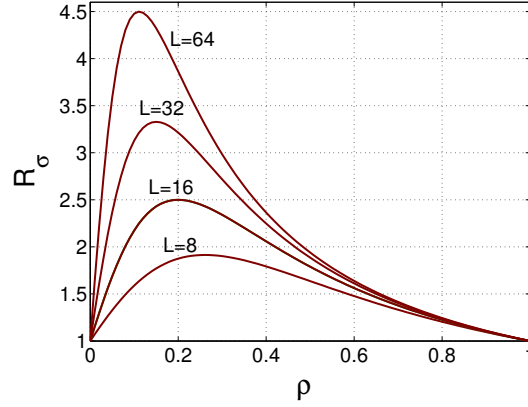
Fig. 1 Noise ratio analytically computed for different number of coils.

$$L_{\text{eff},B} = \frac{L}{1 + \langle \rho^2 \rangle (L-1)} \tag{15}$$

$$\sigma^2_{\text{eff},B} = \sigma^2_n \left(1 + \langle \rho^2 \rangle (L-1)\right). \tag{16}$$

2. For high SNR areas, say $\frac{A_T^2}{\sigma^2_n} \to \infty$, we have:

$$L_{\text{eff},S} = \frac{L}{1 + \langle \rho \rangle (L-1)} \tag{17}$$

$$\sigma^2_{\text{eff},S} = \sigma^2_n \left(1 + \langle \rho \rangle (L-1)\right). \tag{18}$$

Since $\rho_{ij} \in [0,1]$, the effective variance of noise in the signal areas will be greater or equal than in the background, i.e. $\sigma^2_{\text{eff},S} \geq \sigma^2_{\text{eff},B}$. This reasoning suggests the definition of the noise ratio as:

$$\mathcal{R}_\sigma = \frac{\sigma^2_{\text{eff},S}}{\sigma^2_{\text{eff},B}} = \frac{1 + \langle \rho \rangle (L-1)}{1 + \langle \rho^2 \rangle (L-1)} \tag{19}$$

For illustration this ratio is depicted in Fig. 1 for different values of the true L and the particular case $\langle \rho^n \rangle = \rho^n$. The maximum is found in low values of the correlation coefficient, meaning that the non-stationarity is more noticeable for moderately low (but not null) values of ρ.

We can summarize our main points up to now as follows:

1. When correlation between coils is considered, the nc-χ model does not hold.
2. The nc-χ can be used as a good approximation as long as effective parameters are considered.
3. The effective number of coils is decreased and the effective variance of noise is increased.

4. The final distribution is signal-dependent. Thus, the resulting distribution becomes non-stationary, and its parameters are spatially dependent for low SNR and/or large numbers of receiving coils.

3 Material and Methods

3.1 Synthetic experiments

We present some experiments supporting the initial hypotheses. **First**, we numerically validate the nc-χ approximation proposed and measure the error introduced with this approach. The SoS of 128×128 Gaussian correlated synthetic images are considered. We assume 0 mean and the following covariance matrix:

$$\mathbf{\Sigma} = \sigma_n^2 \cdot \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

with $\sigma_n^2 = 1$, $\rho \in [0-1]$, and $L \in [4, 32]$. The analysis in terms of ρ is used to characterize whether or not a nc-χ model can be accurately fitted in each particular situation (since we are considering 0 mean, it will actually be a central-χ). In addition, note that, in this case, the assumption of $\rho_{ij} = \rho$ is irrelevant for the experiment. Our target measure is the relative Mean Squared Error (rMSE):

$$\text{rMSE} = \frac{\int |g_\chi(x) - \widetilde{g_\chi}(x)|^2 dx}{\int |\widetilde{g_\chi}(x)|^2 dx}$$

where $g_\chi(x)$ is the real distribution of the CMS obtained by SoS and $\widetilde{g_\chi}(x)$ is the distribution of an equivalent nc-χ. 100 different realizations are considered and the average of the error is calculated.

In the **second** experiment we test areas with different SNR. 10000 samples of Gaussian correlated synthetic images are considered. We assume two areas, the former with 0 mean (background) and the latter with $A_i = 1$ (signal), the same covariance matrix as before, $\rho \in [0-1]$ and $L \in [8, 32]$. The CMS is obtained using SoS. A nc-χ is fitted to the signal area and a c-χ to the background. We measure rMSE, L_{eff}, and σ_{eff}^2 as the average among 100 independent trials.
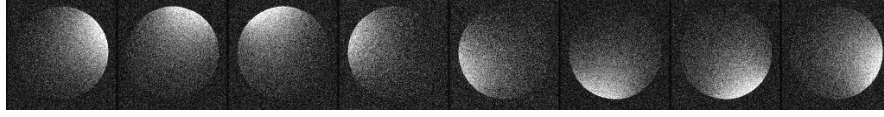
Fig. 2 Slice of an 8-coil 2D acquisition of the phantom used for the experiments.

3.2 Experiments with MR Data-Sets

A final set of experiments is carried out using real data: 100 repetitions of the same slice of a phantom, see Fig. 2, scanned in an 8-channel head coil on a GE Signa 1.5T EXCITE 12m4 scanner with FGRE Pulse Sequence to generate low SNR, Matrix size= 128x128, TR/TE=8.6/3.38 ms, FOV 21x21cm, slice thickness = 1mm.

First we test the existence of noise correlation between coils. Since, in real cases, ρ_{ij} will be different for each pair of coils, we will no longer assume the Σ matrix of the synthetic experiments, but the following one in eq. 10. The sample correlation coefficient is calculated for each pixel using the 100 samples:

$$\widehat{\rho}_{x,y}(\mathbf{x}) = \frac{\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)\sigma_x\sigma_y}$$

The global $\widehat{\rho}_{x,y}$ for each pair of coils is estimated as the average for all pixels.

Once we assure that correlations exist, we expect that the CMS will show a higher variance of noise and a value of the effective number of coils smaller than 8. For the sake of simplicity we will focus on the background, though results may easily be extrapolated to the signal area. The original noise variance σ_n^2 can be initially estimated using the variance of the real part of every coil of every sample, where the noise is known to be AWGN [108, 47]. The effective values are calculated fitting a c-χ to the histogram of the background data of the CMS, which is obtained by SoS. For statistical purposes, the data of the 100 images are used for estimation.

4 Results

The average rMSE for the first experiment is plotted in Fig. 3. Fig. 3-A shows the rMSE for effective parameters, meanwhile Fig. 3-B corresponds to the original L and σ_n^2. According to the plots, it seems clear that the original parameters cannot be used to model the actual distribution, since the error committed will be huge. The use of effective parameters significantly reduces the error of the nc-χ model, or almost cancels it for low correlations. The error grows with ρ, and it is more significant for larger coil numbers. Finally, for high correlation coefficients, the situation is similar to have an identical image in each coil or, alternatively, to have a single coil system. In this case the noise is Rician distributed, and thus the error decreases.
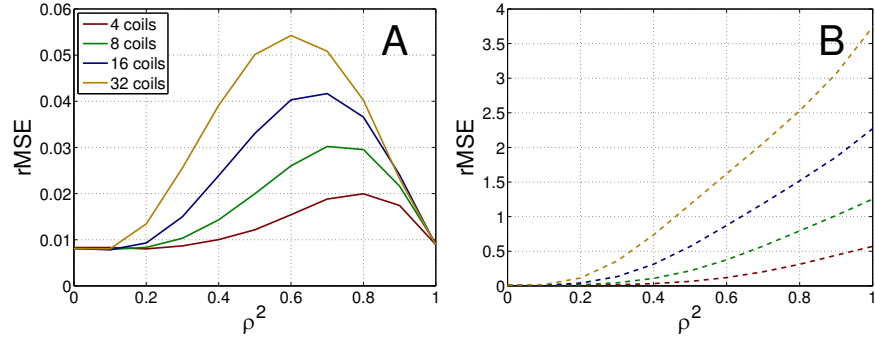
Fig. 3 Relative errors in the PDF for the central χ approximation (Montecarlo simulations from synthetic data), as a function of the correlation coefficient. A) Using effective parameters. B) Using the original parameters.
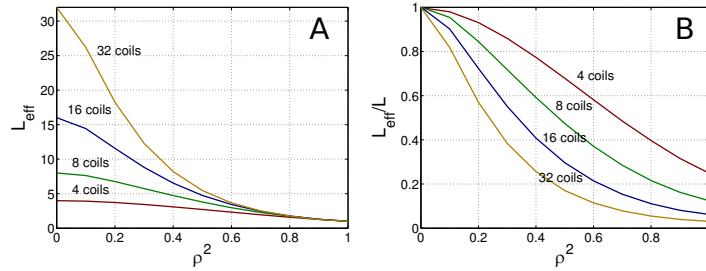
Fig. 4 Effective number of coils as a function of the coefficient of correlation, analytically computed for the setting-up of the first experiment. A) Absolute value. B) Relative Value.

Complementary, on Fig. 4 the effective number of coils for the experiment is depicted. As the correlation grows, all the values tend to one, i.e., for very high correlation the system is equivalent to a single coil. Those configurations with a greater number of coils are the ones with the greater slope (see the relative values in Fig. 4-B): this means they are much more affected by correlations. For instance, a 32-coil configuration with $\rho = 0.2$ will be equivalent to a 18-coil configuration, which gives an idea of the importance of the correlations in the performance of the scanner. Little correlations will provide higher effective number of coils, and therefore smaller effective noise powers.

The results for the second synthetic experiment are shown in Fig. 5. The background (solid line) and signal area (dashed line) have been treated separately. The experiment shows that there is a mismatch between what occurs in the background and what occurs in the signal areas. The effective number of coils, for instance, decays faster in the signal areas, even when a moderate SNR (SNR=1) is considered. With a larger numbers of coils (see $L = 32$) the difference between the signal areas and the background is very noticeable for all ρ. The results are similar for the variance of noise. The different behavior of the effective parameters between the
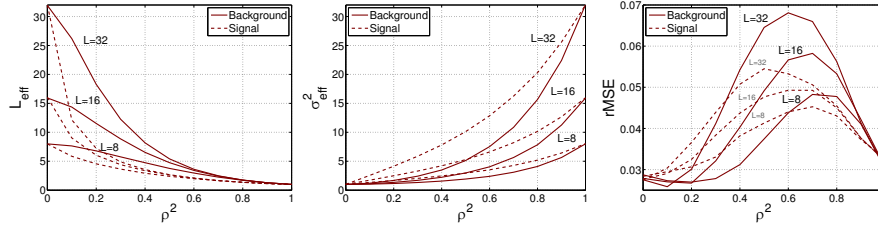
Fig. 5 Effective number of coils, effective variance of noise (analytically computed for the setting-up of the second experiment), and rMSE (numerically computed) for the second experiment.

background and the signal areas will potentially affect noise estimation algorithms based on the background, see [47].

For the final experiment with the MRI phantom, the estimated covariance matrix is:

$$\mathbf{\Sigma} = \sigma_n^2 \cdot \begin{pmatrix} 1 & 0.0060 & 0.0025 & 0.0043 & 0.0070 & 0.0107 & 0.0068 & 0.0132 \\ 0.0060 & 1 & 0.0270 & 0.0017 & 0.0029 & 0.0144 & 0.0220 & 0.0025 \\ 0.0025 & 0.0270 & 1 & 0.0277 & 0.0058 & 0.0126 & 0.0113 & 0.0056 \\ 0.0043 & 0.0017 & 0.0277 & 1 & 0.0056 & 0.0035 & 0.0112 & 0.0003 \\ 0.0070 & 0.0029 & 0.0058 & 0.0056 & 1 & 0.0359 & 0.0006 & 0.0107 \\ 0.0107 & 0.0144 & 0.0126 & 0.0035 & 0.0359 & 1 & 0.0192 & 0.0000 \\ 0.0068 & 0.0220 & 0.0113 & 0.0112 & 0.0006 & 0.0192 & 1 & 0.0234 \\ 0.0132 & 0.0025 & 0.0056 & 0.0003 & 0.0107 & 0.0000 & 0.0234 & 1 \end{pmatrix}$$

As expected, there are some correlations between coils, although their values are bounded by $\rho \leq \rho_{5,6} = 0.0359$. Thus, the effective values are expected to be close to the actual ones. The actual variance of noise σ_n^2 is estimated using the variance of the real part of every coil and location [47]. The effective values estimated over the background are as follows:

$$L = 8; \; L_{\mathrm{eff}} = 7.3073; \; \sigma_n = 0.0430; \; \sigma_{\mathrm{eff}} = 0.0447.$$

Even for relatively low correlations, the reduction in the effective number of coils (respectively, the increase of the effective variance of noise) is clearly noticeable. We would like to stress that this effect has been previously reported in [46] in a merely empirical basis.

In order to study the non-stationarity of the noise we can also estimate the noise ratio between the signal areas and the background

$$\mathcal{R}_\sigma = \frac{1 + \langle \rho \rangle (L - 1)}{1 + \langle \rho \rangle (L - 1)} = 1.0721.$$

The small value means that the non-stationarity is barely noticeable.

5 Discussion and conclussions

Although correlations between coils are usually left aside in multiple coil systems, we have shown that they can seriously affect the statistical distribution of data, especially for modern machinery with a large number of receiving antennae. The effect has been quantified for SoS, although it seems likely to persist with other algorithms. We have evidenced that the CMS can indeed be accurately approximated with the traditional nc-χ model whenever effective parameters (reduced L and increased σ_n) are used. This variation of the effective parameters can have a deep impact on the effectiveness of certain coil architectures depending on the resulting correlations.

Since the effective parameters depend on the value of the signal at each point, the final variance of noise will also depend on the signal, and hence on the position within the image. Although this effect has been previously described for parallel imaging techniques such as SENSE and GRAPPA, we may now conclude it is not a particularity of these protocols, being present even without **k**-space subsampling. Actually, the spatial variability in GRAPPA described in [42] responds to the same correlation phenomenon reported here.

The clearest artifact related to this non-stationarity is the different variance of noise in the background and the signal areas. One could argue that what happens in the background is not important, since the information is in the signal areas. Nonetheless, the background has been traditionally used as a source for parameter estimation and calibration. Many noise estimation techniques are based on the assumption that the variance of noise in the background and the signal areas are the same, and hence the estimation is done using the background pixels. If this underestimated value is used over the signal areas, proper corrections must be made.

The study carried out is also relevant in image denoising. Many of the most popular filtering algorithms are based on the statistical characterization of noise, underlying the assumption that it can be considered as a stationary process. We have shown this is not a realistic assumption in multiple coil systems even without subsampling, and there exists a certain asymmetry between areas with different SNR. These differences are more important as the number of coils grows, the reason being two-folded: first, the change in the effective values directly depends on L; second, as the number of coils in the system increases, it is known that stronger correlations between them will appear, which is the other factor influencing the non-stationarity of the signal.

To sum up, the main consequence of the correlation between coils is a reduction in the effective number of coils in the system. For devices with a large number of coils this reduction could be quite significant, and it is followed by the increase of the effective variance of noise. Hence, to improve the effectiveness of the scanner, the increase of the number of coils should go together with a proper architecture design to reduce the correlation between them.

Acknowledgments

Statistical Noise Analysis in GRAPPA using a parametrized non-central chi approximation model

Santiago Aja-Fernández, Antonio Tristán-Vega and W. Scott Hoge*

Abstract The characterization of the distribution of noise in the magnitude MR image is a very important problem within image processing algorithms. The Rician noise assumed in single-coil acquisitions has been the keystone for SNR estimation, image filtering, or diffusion tensor estimation for years. With the advent of parallel protocols such as SENSE or GRAPPA that allow accelerated acquisitions, this noise model no longer holds. Since GRAPPA reconstructions yield the combination of the squared signals recovered at each receiving coil, noncentral Chi statistics have been previously proposed to model the distribution of noise. However, we prove in this paper that this is a weak model due to several artifacts in the acquisition scheme, mainly the correlation existing between the signals obtained at each coil. Alternatively, we propose to model such correlations with a reduction in the number of degrees of freedom of the signal, which translates in an equivalent non-accelerated system with a minor number of independent receiving coils and, consequently, a lower SNR. With this model, a non-central Chi distribution can be assumed for all pixels in the image, whose effective number of coils and effective variance of noise can be explicitly computed in a closed form from the GRAPPA interpolation coefficients. Extensive experiments over both synthetic and *in vivo* data sets have been performed to show the goodness of fit of out model.

1 Introduction

Statistical models of signal and noise play a fundamental role in medical image processing. In particular, many different applications in the Magnetic Resonance (MR) processing field rely on a well defined prior statistical model of the data. Many ex-

---

43

amples of these model–based methods may be found in literature: noise removal and signal estimation methods as the Conventional Approach [4], signal and noise maximum likelihood estimation [7], Linear Minimum Mean Square Error (LMMSE) filtering based schemes [8] or unbiased non-local mean filters [11]; noise estimation techniques that assume an homogeneous noisy background which follows a Rayleigh or central Chi distribution [27, 47, 81, 37, 8, 84]; Weighted Least Squares methods to estimate the Diffusion Tensor (DT) in DTI, which has proved to be optimal when the data follows a Rician [14] or a non-central Chi distribution [109]; and new procedures for DTI tensor estimation based on an underlying statistical model [110].

For practical purposes, this modeling is usually done assuming noise in MR data is a zero-mean uncorrelated Gaussian process with equal variance in both the real and imaginary parts in each acquisition coil. As a result, in single coil systems magnitude data in the spatial domain are modeled using a stationary Rician distribution [5]. A natural extension can be made to stationary non–central Chi models when more that one coil is present and the **k**–space is fully sampled [27]. Multiple coil systems were developed to enhance the Signal to Noise Ratio (SNR) of the magnitude image while maintaining a large Field of View (FOV). Parallel MRI (pMRI) techniques [30, 31] extended the applicability of these systems by increasing the image acquisition rate via subsampling of the **k**–space data. The **k**–space subsampling creates aliasing and underlying artifacts in the spatial domain image. In order to suppress or correct these artifacts, reconstruction methods such as Sensitivity Encoding (SENSE) [32] and GeneRalized Autocalibrated Partially Parallel Acquisitions (GRAPPA) [33] are employed.

If **k**–space data is subsampled and reconstructed with some pMRI method, the noise power in the final image will become non-stationary: it will vary across the image and also be different in each receiving coil image. Depending on the way the information from each coil is combined (i.e., depending on the reconstruction method), the statistics of the image no longer follow the non–central Chi distribution. It is therefore necessary to study the behavior of the data for a particular reconstruction method. For instance, the reconstruction in **SENSE** takes place into the spatial domain [32] and it can be seen as a combination of the subsampled coils weighted by some factors dependent on the sensitivity map. As a result, the reconstructed image will follow a complex Gaussian distribution [46], and as a consequence of different weighting for each pixel, this distribution will be non-stationary. When taking the magnitude value, this Gaussian becomes a non-stationary Rician distribution. For GRAPPA, some simulation based [43] and preliminary studies [46] have been made, but these methods still lack a theoretical statistical model.

In this paper, we will focus on GRAPPA. This method performs the reconstruction in the **k**-space domain, it does not need an explicit coil sensitivity field map estimation to perform the reconstruction, and supports reconstruction of variable density sub–sampling. We will study how the reconstruction method affects the assumed prior noise models. As a result, we will propose a parametrized non-stationary non-central chi distribution to model the composite magnitude image with GRAPPA reconstruction and sum-of-squares (SoS) coil combination. This will nat-

urally lead to the definition of an *effective number of coils* and an *effective variance of noise* which depend on both the original chi distribution parameters of the measured data and on the GRAPPA reconstruction weights.

2 Theory

2.1 Statistical model in single– and multiple–coil MR signals

k–space data is typically acquired through repeated application of excitation pulses with a different phase encoding gradient for each readout gradient. Each sampled line of **k**–space is "frequency encoded", and the measured signal is uniformly sampled at the desired rate. As a consequence, points in the **k**–space measured from the MRI scanner are independent samples of an RF signal measured with a receiver coil, for which the noise can be modeled as a complex stationary Additive White Gaussian Noise (AWGN) process, with zero mean and variance σ_K^2.

For a single–coil acquisition the complex spatial MR data is therefore also modeled as a complex Gaussian process and the magnitude signal $M(\mathbf{x})$ is the Rician distributed envelope of the complex signal [5]. In the image background, where the signal-to-noise ratio is zero due to the lack of water-proton density in the air, the Rician simplifies to a Rayleigh distribution. In a multiple–coil MR acquisition system, from a merely statistical point of view, the acquired signal in the **k**–space data for each coil can be modeled as:

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; \sigma_K^2), \quad l = 1, \cdots, L \tag{1}$$

with $a_l(\mathbf{k})$ the original signal if no noise is present and $n_l(\mathbf{k}; \sigma_K^2) = n_{l_r}(\mathbf{k}; \sigma_K^2) + jn_{l_i}(\mathbf{k}; \sigma_K^2)$ complex uncorrelated Gaussian processes with zero mean and variance σ_K^2. The noise is initially assumed to be stationary, and therefore parameter σ_K^2 does not depend on **k**.

The complex **x**–space image is obtained via the 2D inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice. Since the iDFT is a linear operator, the noise in the complex signal in the **x**–space for each coil will also be Gaussian, and it can be expressed as:

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; \sigma_n^2),$$

where $N_l(\mathbf{x}; \sigma_n^2) = N_{l_r}(\mathbf{x}; \sigma_n^2) + jN_{l_i}(\mathbf{x}; \sigma_n^2)$ is a complex Gaussian process with zero mean and variance σ_n^2. It is easy to prove the relation [111]:

$$\sigma_n^2 = \frac{1}{|\Omega|}\sigma_K^2, \tag{2}$$

where $|\Omega|$ is the size of the image in each coil, i.e. the number of points used in the 2D iDFT.

If the **k**–space is fully sampled, the composite magnitude image may be obtained using methods such as SoS [27, 38]:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l(\mathbf{x})|^2}. \tag{3}$$

Defining $A_T(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |A_l(\mathbf{x})|^2}$, and assuming the noise components to be identically and independently distributed, the envelope of the magnitude signal M_L will follow a non–central Chi (nc-χ) distribution [27] with probability density function (PDF):

$$p_{M_L}(M_L|A_T, \sigma_n, L) = \frac{A_T^{1-L}}{\sigma_n^2} M_L^L e^{-\frac{M_L^2 + A_T^2}{2\sigma_n^2}} I_{L-1}\left(\frac{A_T M_L}{\sigma_n^2}\right) u(M_L), \tag{4}$$

with $I_L(.)$ the Lth order modified Bessel function of the first kind and $u(.)$ the Heaviside step function. Eq. 4 reduces to the Rician distribution [5] for $L = 1$. In the background, this PDF simplifies to a central Chi distribution with PDF:

$$p_{M_L}(M_L|\sigma_n, L) = \frac{2^{1-L}}{\Gamma(L)} \frac{M_L^{2L-1}}{\sigma_n^{2L}} e^{-\frac{M_L^2}{2\sigma_n^2}} u(M_L). \tag{5}$$

As a final remark, note that three requirements are needed for the SoS of Gaussian random variables (RV) to be modeled as a stationary nc-χ:

1. The noise is stationary in each of the Complex Gaussian **x**–space images. If the **k**–space data is fully sampled, the noise variance will be the same for all the points in the image in both the **k**–space and **x**–space domains, and the noise can be assumed to be stationary.
2. The variance of noise σ_n^2 is the same for each of the coils.
3. No correlation is assumed between the Gaussian RVs.

We will see that these requirements are not fulfilled when GRAPPA is employed. One of the aims of parallel imaging is precisely to accelerate the image acquisition process by sub–sampling data in each coil. From a statistical point of view, such a reconstruction will affect the underlying model and the stationarity of the noise in the reconstructed data.

2.2 Statistical model in GRAPPA reconstructed images

The GRAPPA reconstruction strategy estimates the missing lines in a sub–sampled **k**–space data [33, 30, 34] acquisition for each coil. While the sampled data $s_l^S(\mathbf{k})$ remain the same, the reconstructed lines $s_l^R(\mathbf{k})$ are estimated through a linear combination of the existing samples. Weighted data in a neighborhood $\eta(\mathbf{k})$ around the

estimated pixel from several coils is used for such an estimation:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} \sum_{\mathbf{c} \in \eta(\mathbf{k})} s_m^{\mathcal{S}}(\mathbf{k} - \mathbf{c})\omega_m(l, \mathbf{c}), \tag{6}$$

with $s_l(\mathbf{k})$ the complex signal from coil l at point $\mathbf{k}$ and $\omega_m(l, \mathbf{k})$ the complex reconstruction coefficients for coil l. In self-referenced reconstructions, these coefficients are determined from the low-frequency coordinates of $\mathbf{k}$–space, termed the Auto Calibration Signal (ACS) lines, which are sampled at the Nyquist rate (i.e. unaccelerated). Breuer *et al.* in [51] pointed out that eq. 10 can be rewritten using the *convolution* operator:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} s_m^{\mathcal{S}}(\mathbf{k}) \circledast \mathsf{w}_m(l, \mathbf{k}), \tag{7}$$

where $\mathsf{w}_m(l, \mathbf{k})$ is a convolution kernel that can be easily built from the GRAPPA weight set $\omega_m(l, \mathbf{k})$. Since a (circular) convolution in the $\mathbf{k}$–space is equivalent to a product into the $\mathbf{x}$-space, we can write [51]:

$$S_l^{\mathcal{R}}(\mathbf{x}) = |\Omega| \sum_{m=1}^{L} S_m^{\mathcal{S}}(\mathbf{x}) \times W_m(l, \mathbf{x}) \tag{8}$$

$$= |\Omega| \sum_{m=1}^{L} \left[A_m^{\mathcal{S}}(\mathbf{x}) + N_m(\sigma_n^2) \right] \times W_m(l, \mathbf{x})$$

$$= \underbrace{|\Omega| \sum_{m=1}^{L} A_m^{\mathcal{S}}(\mathbf{x}) \times W_m(l, \mathbf{x})}_{\text{Reconstructed Signal}} + \underbrace{|\Omega| \sum_{m=1}^{L} N_m(\sigma_n^2) \times W_m(l, \mathbf{x})}_{\text{Gaussian Noise}} \tag{9}$$

$$= A_l^{\mathcal{R}}(\mathbf{x}) + N_l^{\mathcal{R}}(\mathbf{x}) \tag{10}$$

with $W_m(l, \mathbf{x})$ the GRAPPA reconstruction coefficients in the $\mathbf{x}$-space. The variance of noise in the $\mathbf{x}$-space of the sampled data is $\sigma_n^2 = \frac{\sigma_K^2}{|\Omega| \cdot R}$, with R the reduction rate. Note that:

1. As a consequence of the coefficients $W_m(l, \mathbf{x})$ (which have different values for each pixel) the reconstructed signal in each coil $S_l^{\mathcal{R}}(\mathbf{x})$ becomes a non–stationary Gaussian process, i.e. a complex distributed Gaussian image with different parameters (mean and variance) in each point of the image:

$$E\{S_l^{\mathcal{R}}(\mathbf{x})\} = A_l^{\mathcal{R}}(\mathbf{x}) = |\Omega| \sum_{m=1}^{L} A_m^{\mathcal{S}}(\mathbf{x}) \times W_m(l, \mathbf{x}) \tag{11}$$

$$\mathrm{Var}\{S_l^{\mathcal{R}}(\mathbf{x})\} = |\Omega|^2 \left[\sum_{m=1}^{L} \sigma_m^2 \times |W_m(l, \mathbf{x})|^2 + 2 \sum_{p \neq q} \sigma_{pq}^2 \mathcal{R}e\{W_p(l, \mathbf{x}) W_q^*(l, \mathbf{x})\} \right]$$

$$= \sigma_l^{2\mathcal{R}}(\mathbf{x}). \tag{12}$$

Therefore, the stationarity requirement in described in section 2.2 fails.

2. For the same pixel $\mathbf{x}$, the mean and the variance will also vary along the coil dimension, and therefore the second requirement for the nc-χ model fails.
3. Even assuming that the coils are initially uncorrelated, signals $S_l^{\mathcal{R}}(\mathbf{x})$ will be correlated due to the GRAPPA reconstruction. Note that the signal in each coil is a linear combination of the signals at all remaining coils, hence the third requirement also fails.

The composite magnitude image $M_L(\mathbf{x})$ can be obtained using SoS. For convenience, in what follows we will equivalently consider $M_L^2(\mathbf{x})$ instead of $M_L(\mathbf{x})$ (if M_L follows a nc-χ distribution, M_L^2 will follow a nc-χ squared).

After GRAPPA reconstruction, the interpolated lines at each coil $S_l^{\mathcal{R}}(\mathbf{x})$ can be seen as strongly correlated non-stationary complex Gaussian RVs. Therefore, from a theoretical point of view the magnitude image after SoS, $M_L(\mathbf{x})$, will not strictly be a nc-χ, although it is usually modeled that way in the literature [46].

The SoS image can be expressed as

$$M_L^2(\mathbf{x}) = \sum_{l=1}^{L} |S_l^{\mathcal{R}}(\mathbf{x})|^2 \tag{13}$$

$$= \sum_{l=1}^{L} |S_l^{\mathcal{S}}(\mathbf{x})|^2 \cdot C_l^2(\mathbf{x}) + 2 \sum_{p>q} \mathcal{R}e\{S_p^{\mathcal{S}}(\mathbf{x}) \left(S_q^{\mathcal{S}}(\mathbf{x})\right)^* C_{pq}^2(\mathbf{x})\} \tag{14}$$

with

$$C_l^2(\mathbf{x}) = |\Omega|^2 \sum_{i=1}^{L} W_i(l, \mathbf{x}) \cdot W_i^*(l, \mathbf{x}) = |\Omega|^2 \sum_{i=1}^{L} |W_i(l, \mathbf{x})|^2$$

$$C_{pq}^2(\mathbf{x}) = |\Omega|^2 \sum_{i=1}^{L} W_i(p, \mathbf{x}) \cdot W_i^*(q, \mathbf{x}) = C_{qp}^2(\mathbf{x})^*.$$

Signal $M_L^2(\mathbf{x})$ can be seen as the SoS of the sum of weighted Normal variables. A deep study of the resultant distribution is done in Appendix 6.1. The mean and the variance (for each $\mathbf{x}$) of the resulting distribution are:

$$E\{M_L^2(\mathbf{x})\} = A_T^2 + 2 \, \mathrm{tr}\left(\mathbf{C}_X^2\right); \tag{15}$$

$$\mathrm{Var}\{M_L^2(\mathbf{x})\} = 4 \, \mathbf{A}^* \mathbf{C}_X^2 \mathbf{A} + 4 \, ||\mathbf{C}_X^2||_F^2. \tag{16}$$

where $\mathbf{C}_X^2 = \mathbf{W}\Sigma\mathbf{W}^*$ is the covariance matrix of the **interpolated** data at each spatial location, $\mathbf{M}(\mathbf{x}) = [M_1^2, M_2^2, \cdots, M_L^2]^T$. Accordingly, Σ^2 is the covariance matrix of the **original** data at each coil. $||.||_F$ is the Frobenious norm, and $\mathbf{A}(\mathbf{x}) = [A_1, A_2, \cdots, A_L]^T$ is the noise-free reconstructed signal, from which we define:

$$A_T^2(\mathbf{x}) = \sum_{i=1}^{L} |A_i^{\mathcal{R}}(\mathbf{x})|^2 = \sum_{l=1}^{L} |A_l^{\mathcal{S}}(\mathbf{x})|^2 \cdot C_l^2(\mathbf{x}).$$

Although the resultant distribution is not strictly a nc-χ^2, our hypothesis is that its behavior will be very similar and could be modeled as such (the error in this approximation will be quantified in the results section). However, the high correlations between the reconstructed signals in each coil should translate in a decrease of the number of Degrees of Freedom (DoF) of the distribution. In Appendix 6.1, the method of the moments is used to show that the reduction in the number of DoF translates in an *effective number of coils* (obviously reduced, since the number of coils L is clearly related to the DoF) and an *effective variance of noise*:

$$L_{\text{eff}}(\mathbf{x}) = \frac{A_T^2 \operatorname{tr}\left(\mathbf{C}_X^2\right) + \left(\operatorname{tr}\left(\mathbf{C}_X^2\right)\right)^2}{\mathbf{A}^* \mathbf{C}_X^2 \mathbf{A} + ||\mathbf{C}_X^2||_F^2}; \tag{17}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\operatorname{tr}\left(\mathbf{C}_X^2\right)}{L_{\text{eff}}}. \tag{18}$$

Finally, note that the origin of the reduced DoF of the nc-χ^2 model is in the correlation and inhomogeneous variance of the complex Gaussians, i.e. in $\mathbf{C}_X^2$ not being of the form $\sigma^2\mathbf{I}$ with $\mathbf{I}$ the identity matrix. With GRAPPA the distortion comes mainly from the interpolation matrix $\mathbf{W}$ which is not diagonal. For unaccelerated acquisitions we may write $\mathbf{W} = \mathbf{I}$, so that $\mathbf{C}_X^2 = \Sigma^2$. Even in this case, Σ^2 is not necessarily diagonal, since a certain correlation does exist between the signals at each channel especially for those systems with a large number of receiving coils L. This means that effective parameters should be considered even if no subsampling is performed, as has been empirically shown by [46]. We can summarize our developments as follows:

1. The nc-χ model does not hold for GRAPPA reconstructed data. However, this distribution can be used as a good approximation of the actual one.
2. For this approximation to hold, effective parameters have to be considered which represent an equivalent, non-subsampled configuration with a smaller number of coils (DoF) and, consequently, a greater level of noise.
3. Even when the nc-χ model is feasible, the resulting distribution is non-stationary (the effective parameters are spatially dependent).
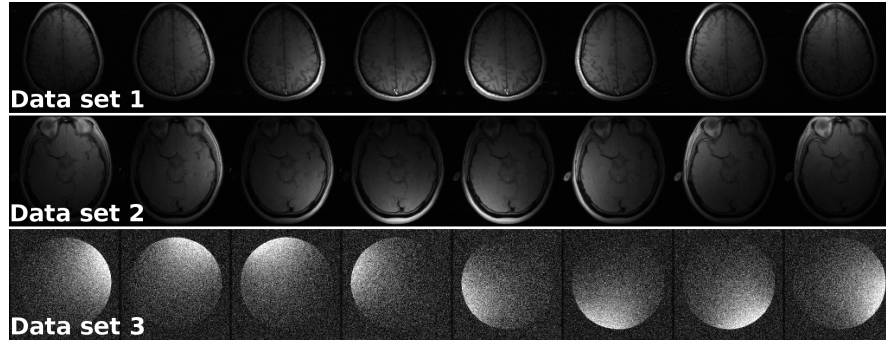
Fig. 1 Data sets for the experiments.

3 Materials and Methods

3.1 Data sets

In order to validate the hypothesis posed in the previous section, the following data sets are considered for the experiments, see Fig. 1:

- Data set 1: 1 repetition of 8-channel head coil data acquired on a GE Signa 1.5T EXCITE 11m4 scanner, FGRE Pulse Sequence, TR=500 ms, TE= 13.8 ms, matrix size= 256×256, FOV=20×20 cm, slice thickness= 5 mm.
- Data set 2: the MR brain data-set from [112] (the authors provide their data set on-line[2]); eight-channel head array from 3 Tesla GE scanner, fast spoiled gradient-echo sequence, TR/TE=300/10 ms,RBW= 16 kHz, matrix size = 256×256, FOV 22×22 cm.
- Data set 3: 60 repetitions of the same slice of a phantom, scanned in an 8-channel head coil on a GE Signa 1.5T EXCITE 12m4 scanner with FGRE Pulse Sequence to generate low SNR. matrix size= 128×128, TR/TE=8.6/3.38 ms, FOV 21x21cm, slice thickness = 1mm.

  In addition, synthetically generated random variables will also be considered.

3.2 Synthetic Experiments

First, we numerically validate the nc-χ approximation proposed and measure the error introduced with this approach. To that end synthetically generated RVs are considered in the following experiments and the following distributions are compared:

---

[2] `http://www.ece.tamu.edu/~jimji/pulsarweb/index.htm`

- A nc-χ^2 distribution with the *effective* parameters L_{eff} and σ^2_{eff}.
- A nc-χ^2 distribution with the original parameters ($L = 8$).
- A Normal distribution fitted using the mean and the variance in eqs. 15 and 16.

The validation experiments are as follows:

1. For the sake of illustration, we fit the histogram of the real distribution to the proposed three statistical models. We consider $5 \cdot 10^4$ samples of 8 complex Gaussian RVs with zero mean and unitary variance, $N(0, 1)$. The RVs are combined using 8×8 complex random weights ($\mathbf{W} \sim N(1, 4)$) and then combined using the (square) SoS. The experiment is repeated with the same parameters, but considering a signal value of $A_i = 3$, $i = 1, \cdots 8$.

2. We will analyze the error committed with the nc-χ^2 approximation for different configurations of matrix $\mathbf{C}^2$. To avoid any effect of the signal over the noise, we first assume that $\mu_i = 0$. We define 10^4 samples of 8 complex Gaussian RVs and a 8×8 coefficient matrix $\mathbf{W}$, so that three different scenarios are considered (see Appendix for more details about the relation between matrix $\mathbf{C}^2$ configuration and the final distribution):

   a. Matrix $\mathbf{C}^2$ is diagonal (different elements in the diagonal).
   b. Matrix $\mathbf{C}^2$ is not diagonal and the diagonal elements are greater than the rest of the elements (diagonal dominant).
   c. Matrix $\mathbf{C}^2$ is totally random.

   Since we are assuming unitary noise variance and no correlation, note that $\mathbf{C}^2 = \mathbf{C}^2_X$. These configurations will be measured by two parameters:

   a. The coefficient of variation of the elements of the diagonal: $\text{Cv}^2\{C_l^2\}$.
   b. The ratio between the determinant of $\mathbf{C}^2$ and the trace of the matrix: $||\mathbf{C}^2||/\text{Tr}(\mathbf{C}^2)$. This is a measure of the weight of the non-diagonal elements over the diagonal. Note that for the first case, this parameter is a constant.

   The RVs are combined following eq. 14. The analysis in terms of the elements of matrix $\mathbf{C}^2$ is used to characterize whether a nc-χ^2 can be accurately fitted in each particular situation[3]. The following (relative) Mean Squared Error (rMSE) is measured:

   $$\text{rMSE} = \frac{\int_\infty^\infty |p_O(x) - p_z(x)|^2 dx}{\int_\infty^\infty |p_z(x)|^2 dx} \tag{19}$$

   where $p_O(x)$ is the PDF of the original data (estimated from the histogram of the 10^4 samples for each pixel) and $p_z(x)$ is the distribution to fit.

3. We will test the model fit for different SNR values with a given set of weights. We consider 10^4 samples of 8 complex Gaussian RVs with unitary variance and mean in the range $\mu \in (0, 3)$, i.e $\sim N(\mu, 1)$. The RVs are combined using 8×8 complex random weights (selected in a range similar to an actual GRAPPA

---

[3] Note that since $A_i = 0$, we are really fitting a central χ^2 rather than a nc-χ^2, but the reasoning and the conclusions are similar.

reconstruction) and the SoS is considered. We calculate the error of fitting a nc-χ^2 and a Gaussian for different SNR values, defining the original SNR in each coil as SNR= $\frac{\mu}{\sigma} = \mu$.

3.3 Experiment with the MR data-sets

In a second stage, GRAPPA coefficients from actual MR acquisitions will be used. In each case, a rectangular 2 k_y -by- 5 k_x GRAPPA convolution kernel $\eta(k)$ is employed. The following experiments are considered:

1. Data sets 1 and 2: The **k**–space data are $2\times$ subsampled with 32 ACS lines–skipping every other line except in the central region. From the ACS lines the GRAPPA coefficients $\omega_m(l, \mathbf{k})$ are estimated. The **k**–space domain coefficients are then *transformed* to **x**-space, $W_m(l, \mathbf{x})$. The experiments proposed for synthetic data are repeated using the calculated GRAPPA coefficients. For each of the pixels in the image we consider 10^4 samples of 8 complex Gaussian RVs with unitary variance and mean in the range $\mu \in (0, 4)$, so that original SNR $\in (0, 4)$. To simulate the reconstruction process, for each of the values of **x** the RVS are combined following eq. 9, and the composite magnitude image is obtained by SoS. For the sake of simplicity, $M_l^2(\mathbf{x})$ will be used. After reconstruction, 10^4 samples of the 256×256 image are available.The distribution of the data is tested by fitting a theoretical probability distribution to the real data. To that end, the rMSE is measured. The same three distributions used in the synthetic experiments (nc-χ^2 with effective parameters, nc-χ^2 with the original parameters and Gaussian) are considered.
2. Data set 1: To further test it, the experiment for Data set 1 is repeated with SNR= 1. A variable number of samples is considered, and a Kolmogorov–Smirnov test at a significance level of $\alpha = 0.05$ is used to decide whether the sample data follow any of the proposed distributions.
3. Data set 3: The GRAPPA reconstruction coefficients are derived from one of the 60 samples, using $2\times$ subsampling and 32 ACS lines. All the 60 samples are $2\times$ subsampled and reconstructed with the same GRAPPA coefficients and the composite magnitude image is obtained by SoS. A Golden Standard (to obtain the A_i values) is created by averaging the complex data in each coil. The original σ_n and σ_K values for each coil are estimated from the background of the fully sampled images [47]. Theoretical values of $L_{\text{eff}}(\mathbf{x})$ and $\sigma^2_{\text{eff}}(\mathbf{x})$ are calculated for each point **x**. A Kolmogorov–Smirnov test at a significance level of $\alpha = 0.05$ is used to decide whether each **x** point of the 60 samples follow a nc-χ distributions with parameters $L_{\text{eff}}(\mathbf{x})$ and $\sigma^2_{\text{eff}}(\mathbf{x})$.
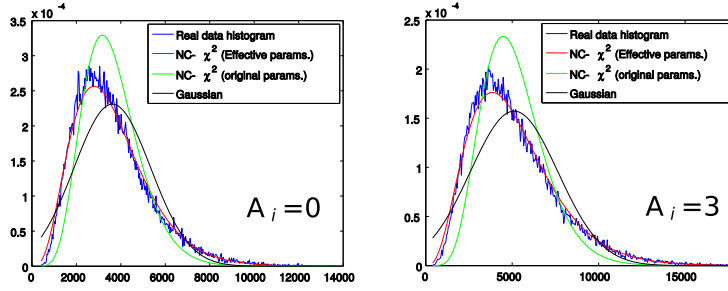
Fig. 2 Actual distribution of the SoS of the sum of weighted Normal RVs compared to nc-χ^2 distribution approximation and Gaussian distribution.

4 Results

4.1 Synthetic Experiments

First, we will consider the experiments with synthetically generated weights. Results for the first experiment are shown in Fig. 2. Note that the nc-χ^2 distribution with effective parameters is the distribution that better follows the actual variations of the data for both experiments. The more *asymmetric* the distribution, the better is the behavior of the nc-χ^2 when compared to the Gaussian.

The second experiment measures the error of the different approximations as a function of the configuration of matrix $\mathbf{C}^2$, i.e. the $C^2_{pq}(\mathbf{x})$ coefficients. Results are depicted in Fig. 3. In all the cases, the nc-χ^2 distribution with effective values is the one that shows a lower rMSE for a wider range of variation of parameters. In all the cases, the error of approximating the actual distribution by a nc-χ^2 is always smaller than 0.1, usually around 0.05. In Fig. 3-(a) (diagonal case) we can see that as the values in the diagonal are more different, the Gaussian and the nc-χ^2 with original parameters produce a greater error. However, the approximation with effective parameters remains constant. Fig. 3-(b)-(c) shows the results for the dominant diagonal, and Fig. 3-(d)-(e) the results for the random $\mathbf{C}^2$. An interesting effect in these two experiments is that for the original nc-χ^2 and the Gaussian approaches the error increases when the ratio $||\mathbf{C}^2||/\mathrm{Tr}(\mathbf{C}^2)$ grows. It means that, as the diagonal becomes less significant, these two approaches are not a good representation of the data. The nc-χ^2 approximation with effective parameters due precisely to the fine tune of these effective parameters, is able to properly fit the actual distribution.

The evolution of the effective number of coils for the experiment in Fig 3-(d)-(e) is reproduced in Fig. 4. As the ratio $||\mathbf{C}^2||/\mathrm{Tr}(\mathbf{C}^2)$ grows, the correlation between variables also grows. As a consequence, the number of DoF of the system will decrease together with the effective number of coils, L_{eff}. This reduction in L_{eff} implies an increase in σ^2_{eff} and consequently a decrease in the SNR.
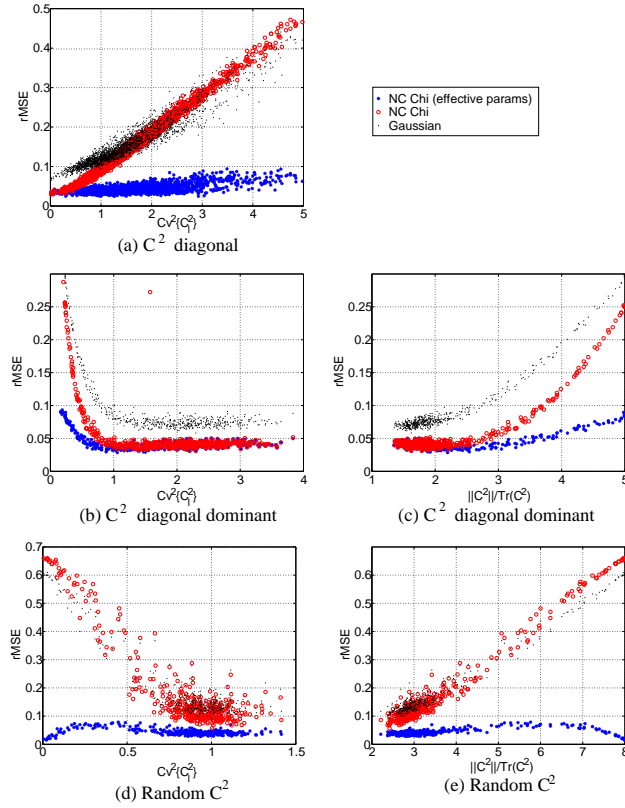
Fig. 3 Relative errors in the PDF for the non-central Chi approximation with effective parameters. For the sake of comparison, the non-central Chi with fix parameters and the Gaussian distribution are also considered.
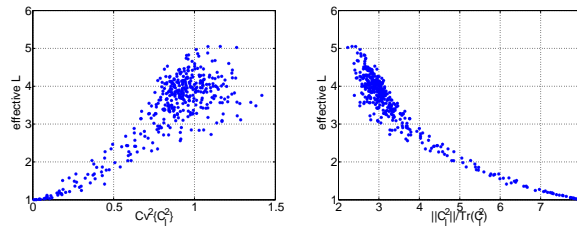

Fig. 4 Evolution of the effective value of parameter L for the experiments in Fig 3-(d)-(e).

The third synthetic experiment deals with the error of the approximation for different SNR configurations. Results are depicted in Fig. 5. For large SNR values (SNR> 2), the nc-χ^2 and the Gaussian approximation converges, as expected.
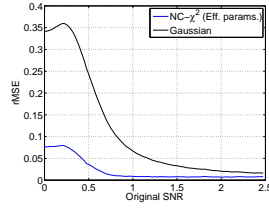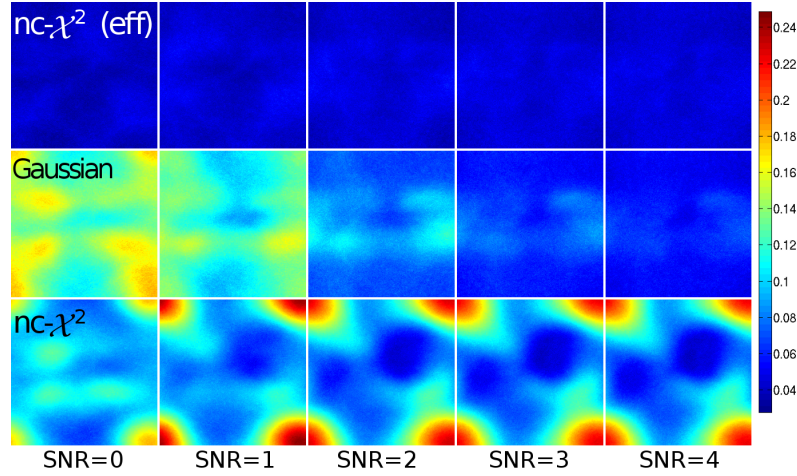
Fig. 5 Relative errors for the nc-χ^2 approximation with effective parameters and Gaussian for different SNR values.
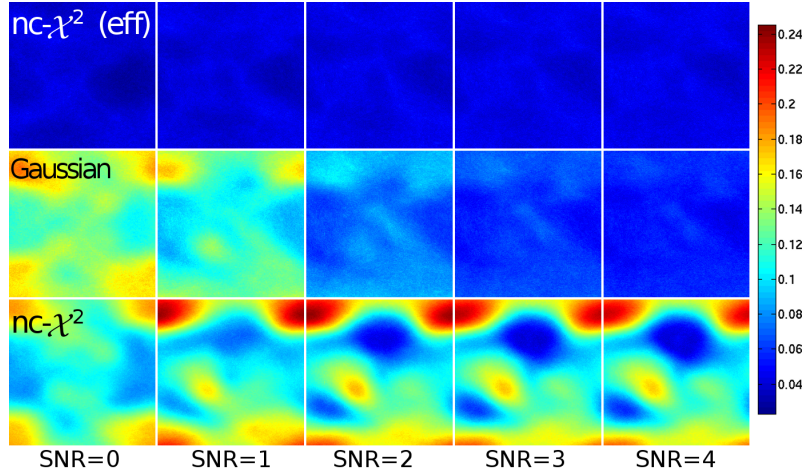
Fig. 6 rMSE of different statistical approximations using the GRAPPA coefficients of dataset 1. Different original SNR are considered. Top row: nc-χ^2 with effective parameters. Middle row: Gaussian. Bottom row: nc-χ^2 with original parameters.

4.2 Experiment with the MR data-sets

The next step is to analyze the results generated by estimated GRAPPA coefficients. Results for the data fit for different distributions are shown in Fig. 6 for the data set 1 and in Fig. 7 for the data set 2. In both cases, the rMSE for each of the points of the final image is shown. For better illustration, the average of the rMSE in the whole image is depicted in Fig. 8. Regardless of the SNR level, the error committed when approximating the distribution by a nc-χ^2 with effective parameters is always lower than 6%, while the other two models (Gaussian and nc-χ^2) show greater values. What is more, the rMSE for nc-χ^2 is quite independent of the SNR value. As expected, the Gaussian model is a good approximation when SNR increases. On the other hand, if nc-χ^2 is considered but the effective parameters are not used, the error even increases together with the SNR in certain areas of the image.
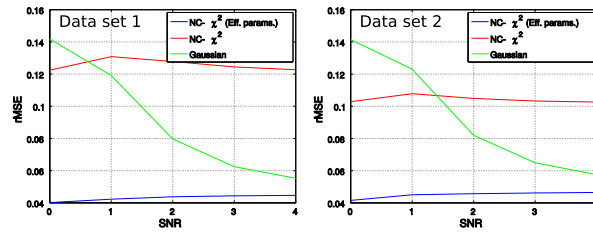
Fig. 7 rMSE of different statistical approximations using the GRAPPA coefficients of dataset 2. Different original SNR are considered. Top row: nc-χ^2 with effective parameters. Middle row: Gaussian. Bottom row: nc-χ^2 with original parameters.


**Fig. 8** Average of the rMSE in the whole image for the different SNR values.

From these results, we can conclude that the nc-χ model (with effective parameters) proposed gives a small error when considering the GRAPPA reconstruction coefficients from actual parallel acquisitions.

It is also interesting analyzing the distribution of the effective parameters along the image. Note that these parameters only depend on the GRAPPA coefficients. The effective number of coils (L_{eff}) is shown in Fig 9 and the effective variance of noise (σ^2_{eff}) is shown in Fig. 10. Note that those cases with low SNR also show low levels of the effective number of coils. Note also that the distribution of L_{eff} correlates with the error committed when considering nc-χ^2 without effective parameters. The points with lower rMSE are logically those with greatest values of L_{eff}.

The variance of noise, on the other hand, shows a great variation along the image, which confirms the non-stationarity assumption previously done. Values of σ_n range between 2 and 6. However, note that the variation of this parameter across the image is *soft*. So, in some applications (like filtering methods) local stationarity may be assumed if needed.
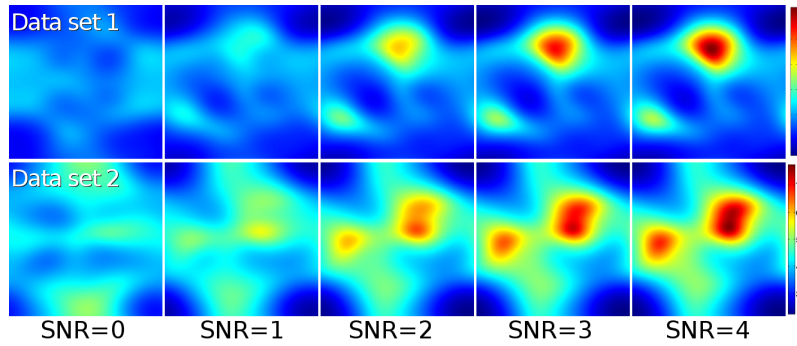
Fig. 9 L effective for different SNR values. Top: Data set 1. Bottom: Data set 2.
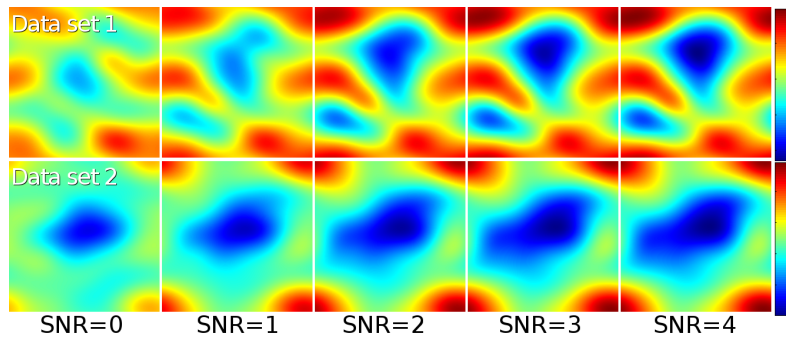
Fig. 10 σ_n effective for different SNR values. Top: Data set 1. Bottom: Data set 2.
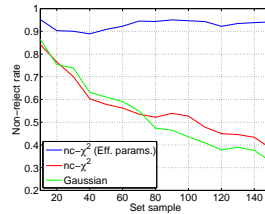
Fig. 11 Rate of sets passing the Kolmogorov–Smirnov test for the distributions under study as a function of the size of the population. Significance level is $\alpha = 0.05$, and a SNR=1. The data are synthetically generated from the GRAPPA coefficients of data set 1.

In Fig. 11 the acceptance rate by the Kolmogorov-Smirnov test for different set sizes is depicted. The acceptance rate for the nc-χ^2 model (with effective parameters) is over the 90% even for large population sizes. The larger the size of the sample, the easier for the test to reject the null hypothesis. For the Gaussian and nc-χ^2 models the rate goes down very fast, so that they can only be accepted for very small populations. It may be concluded that the effective nc-χ^2 model accurately fits data generated following the GRAPPA reconstruction algorithm.

Finally, the Kolmogorov-Smirnov test is run over the 60 GRAPPA reconstructed samples of the MR phantom (data set 3). The 90.13% of the points in the image passed the test, which correlates well with the results of the previous experiment, and confirms the nc-χ (with effective parameters) as a suitable distribution to accurately model data after GRAPPA reconstruction and SoS.

5 Discussion

The statistical characterization of the composite magnitude signal after GRAPPA reconstruction as a parameterized non–stationary nc-χ distribution has been derived. In fully sampled images, the derivation of the magnitude signal distribution is a straight transformation from the prior Gaussian model for each coil. With pMRI techniques, the final model strongly depends on the particular algorithm used. In general, those reconstructing one single image domain drive to a Rician distribution, since the magnitude is computed as the modulus of a Gaussian variable. Contrary to conventional MRI, the noise in the image cannot be considered stationary strictly speaking, since its variance shows fluctuations across the image. Besides, noisy samples cannot be considered independent either, due to interpolation artifacts.

The analysis of GRAPPA reconstructed signals is more difficult. Although its feature of providing one complete image domain for each receiving coil assimilates it to conventional multiple coils systems, the non–central Chi distribution does not directly apply to this case. Apart from the aforementioned interpolation artifacts related to non–stationary noisy patterns and correlations, the noise variance for each receiving coil is different with this kind of reconstruction. Hence, a non–central Chi model cannot be assumed in general. For brain images, which are the data sets of interest in this dissertation, it has been shown that in fact this distribution is a very accurate approximation to the actual statistics of noise. However, as a consequence of the correlation between the reconstructed data in each coil, reduced effective parameters, namely the *effective number of coils* and the *effective variance of noise*, must be taken into account. Thus, the final distribution for each point is equivalent to an acquisition with a fewer number of coils and a greater variance of noise.

We want to remark the importance of the suitable statistical model for GRAPPA proposed here. As an example. we will propose some future applications. (1) In the case of tensor estimation in DTI, in [109] authors show that Weighted Least Squares (WLS) is theoretically a good method to estimate the DT if the data is nc-χ distributed. GRAPPA reconstructed data is being widely used for these purposes, but knowing the models we assure a good estimation, and we can measure the bias and variance of the estimation error. (2) Although the signal is non-stationary, we can assume a small spatial variation of the parameters. This way, we can calculate local statistics and reformulate Rician noise filtering methods, like the conventional approach [4], ML [7], or LMMSE [8]. These algorithms will need as an input the number of coils L and the variance of noise σ_n^2. When applied over GRAPPA im-

ages, the new derived *effective* parameters must be used instead.(3) Noise estimation is usually done assuming that the noise is stationary in the background [113], i.e. the value of the variance of noise is homogenous for all the pixels in the background. As a consequence, those methods cannot be directly applied with GRAPPA. However, with the *effective* σ_n value derived for each **x** value in the image, the inhomogeneous background in GRAPPA can be easily corrected and used for estimation, straightly applying some of the methods in [47].

It is also convenient to stress the limitations of the study performed in this paper. First, the analysis carried out in this paper is only valid for the SoS to combine the images from each coil. Other methods used by newer systems, such as linear channel combination or B1 maps have not been considered. With these methods eq. 14 will differ, and a new study of parameters will be needed. Finally, certain MRI techniques, such as EPI, which can show their own peculiarities with regard to the statistical characterization of noise, has been removed from the study.

6 Conclusions

The main contribution of this paper is the statistical characterization of the composite magnitude signal after GRAPPA reconstruction as a parameterized non–stationary nc-χ distribution. We have shown that although data does not strictly follows a non–central Chi model, in practical cases this distribution is a very accurate approximation to the actual statistics of noise in the magnitude composite signal when SoS is used. To really fit the distribution to the data, reduced effective parameters are considered: the *effective number of coils* and the *effective variance of noise*. As a final remark we want to point out that the studied carried out is totally valid if initial correlations are assumed between coils.

6.1 About the SoS of the sum of weighted Normal variables

Let $N_l(\mu_l, \sigma_l^2)$, $l = \{1, \cdots, L\}$, be a set of independent Gaussian complex random variables (RVs) with mean μ_l and variance σ_l^2. For simplicity, let us initially assume a diagonal covariance matrix, i.e. there is not any initial correlation between variables:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_L^2 \end{pmatrix}$$

We define a linear combination of these RVs (as the one done for each pixel in GRAPPA in the **x**-space, see eq. 1) using complex weights as

$$M_i = \sum_{l=1}^{L} N_l \cdot W_{il}$$

with W_{il} a set of known (generic) complex weights which are arranged into the following matrix:

$$\mathbf{W} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1L} \\ W_{21} & W_2 & \cdots & W_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ W_{L1} & W_{L2} & \cdots & W_{LL} \end{pmatrix} \tag{20}$$

The (square) sum of squares (SoS) is defined as:

$$X_2 = \sum_{i=1}^{L} |M_i|^2 \tag{21}$$

$$= \sum_{l=1}^{L} |N_l|^2 \cdot C_l^2 + 2 \sum_{p>q} \mathcal{R}e\left\{ N_p N_q^* \cdot C_{pq}^2 \right\} \tag{22}$$

with

$$C_l^2 = \sum_{i=1}^{L} W_{il} \cdot W_{il}^* = \sum_{i=1}^{L} |W_{il}|^2$$

$$C_{pq}^2 = \sum_{i=1}^{L} W_{ip} \cdot W_{iq}^* = (C_{qp}^2)^*$$

Note that $C_l^2 = C_{ll}^2$. Let us define matrix $\mathbf{C}^2$ as

$$\mathbf{C}^2 = \begin{pmatrix} C_1^2 & C_{12}^2 & \cdots & C_{1L}^2 \\ C_{21}^2 & C_2^2 & \cdots & C_{2L}^2 \\ \vdots & \vdots & \ddots & \vdots \\ C_{L1}^2 & C_{L2}^2 & \cdots & C_L^2 \end{pmatrix} = \mathbf{W}\mathbf{W}^*$$

There are three possible scenarios directly related to the three requirements for the nc-χ described in section 2.2:

1. Matrix $\mathbf{C}^2$ is diagonal and with all the elements in the diagonal equal and the variance of all the Gaussian variables is the same, i.e. $\sigma_l^2 = \sigma^2$. All the requirements are fulfilled and X_2 follows a non-central Chi square (nc-χ^2) distribution [45] with PDF:

$$p_X(x|A_T, \sigma, L) = \frac{A_T^{1-L}}{2\sigma^2} x^{\frac{1}{2}(L-1)} e^{-\frac{x+A_T^2}{2\sigma^2}} I_{L-1}\left(\frac{A_T\sqrt{x}}{\sigma^2}\right) u(x), \tag{23}$$

with $A_i = \sum_{l=1}^{L} \mu_l \cdot W_{il}$ and $A_T^2 = \sum_{i=1}^{L} |A_i|^2 = \sum_{l=1}^{L} |\mu_l|^2 \cdot C_l^2$. The mean and the variance of the distribution will be

$$E\{X_2\} = A_T^2 + 2\sigma^2 L \tag{24}$$

$$\text{Var}\{X_2\} = 4A_T^2\sigma^2 + 4L\sigma^4. \tag{25}$$

2. Matrix $\mathbf{C}^2$ is diagonal, but the variance of each variable σ_i^2 is different, and therefore the elements of the diagonal are different. In this case, the second requirement is not fulfilled and the distribution of X_2 is no longer nc-χ^2. For each variable, $|M_i|^2$ will follow a nc-χ^2 distribution as the one in eq.23 with $L = 1$ and its characteristic function will be

$$G_{M_i^2}(\omega) = \frac{1}{1 - 2j\omega\sigma_i^2 C_i^2} \exp\left(\frac{|A_i|^2 j\omega}{1 - 2j\omega\sigma_i^2 C_i^2}\right).$$

Therefore, the characteristic function for X_2 will become:

$$G_{X_2}(\omega) = \prod_{i=1}^{L} G_{M_i^2}(\omega) = \prod_{i=1}^{L} \frac{1}{1 - 2j\omega\sigma_i^2 C_i^2} \exp\left(\frac{|A_i|^2 j\omega}{1 - 2j\omega\sigma_i^2 C_i^2}\right)$$

The mean and the variance of this distribution are:

$$E\{X_2\} = A_T^2 + 2\sum_{i=1}^{L} \sigma_i^2 C_i^2 \tag{26}$$

$$\text{Var}\{X_2\} = 4A_T^2 \sum_{i=1}^{L} \sigma_i^2 C_i^2 + 4\left(\sum_{i=1}^{L} \sigma_i^2 C_i^2\right)^2 \tag{27}$$

The greater the fluctuation of the noise power σ_l^2 along the coils, the greater the error committed when approximated by a nc-χ^2. The error in the approximation can be measured by the variability of σ_l^2.

3. Matrix $\mathbf{C}^2$ is not diagonal. Neither the second nor the third requirements are fulfilled. In this case, the final PDF or the characteristic function become hard to derive theoretically. The mean and the variance can be calculated:

$$E\{X_2\} = A_T^2 + 2\sum_{l=1}^{L} \sigma_l^2 \cdot C_l^2 \tag{28}$$

$$\text{Var}\{X_2\} = 4\sum_{l=1}^{L} \sigma_l^2 |\alpha_l|^2 + 4\sum_{p=1}^{L}\sum_{q=1}^{L} |C_{pq}^2|^2 \sigma_p^2 \sigma_q^2. \tag{29}$$

with

$$\alpha_l = \sum_{i=1}^{L} A_i W_{il}^* = \sum_{i=1}^{L} \mu_i \cdot C_{il}^2$$

If we define the covariance matrix of X_2 as [51]

$$\mathbf{C}_X^2 = \mathbf{W} \cdot \mathbf{\Sigma} \cdot \mathbf{W}^* \tag{30}$$

the mean and the variance can be rewritten as

$$E\{X_2\} = A_T^2 + 2\,\mathrm{tr}\left(\mathbf{C}_X^2\right) \tag{31}$$

$$\mathrm{Var}\{X_2\} = 4\,\mathbf{A}^*\mathbf{C}_X^2\mathbf{A} + 4\,||\mathbf{C}_X^2||_F^2. \tag{32}$$

with $||.||_F$ the Frobenious norm and $\mathbf{A} = [A_1, A_2, \cdots, A_L]^T$.

The third case is the one usually found when dealing with GRAPPA reconstruction. As an effect of the correlation between the RVS M_i, the number of Degrees of Freedom (DoF) of the final distribution would decrease. In order to fit the new distribution to a nc-χ^2, the method of the moments is used over the theoretical mean and variance in eqs. 24, 25, 28 and 29. As a result, the final nc-χ^2 shows new *effective* values for the parameters L and σ:

$$L_{\mathrm{eff}} = \frac{A_T^2\,\mathrm{tr}\left(\mathbf{C}_X^2\right) + \left(\mathrm{tr}\left(\mathbf{C}_X^2\right)\right)^2}{\mathbf{A}^*\mathbf{C}_X^2\mathbf{A} + ||\mathbf{C}_X^2||_F^2} \tag{33}$$

$$\sigma_{\mathrm{eff}}^2 = \frac{\mathrm{tr}\left(\mathbf{C}_X^2\right)}{L_{\mathrm{eff}}} \tag{34}$$

The *effective* L parameter is directly related to the DoF of the distribution, and it will decrease. As a consequence, parameter σ will increase.

This solution has been obtained assuming that there is no initial correlation between coils. However, it can be easily extrapolated to the correlated case. Note that even if $\mathbf{\Sigma}$ is not diagonal, it is an Hermitian, positive-definite matrix, and therefore it can be decomposed as:

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^*$$

with D the real, diagonal matrix with positive eigenvalues. Thus, we can write:

$$\mathbf{C}_X = (\mathbf{W}\mathbf{U}) \cdot \mathbf{D} \cdot (\mathbf{W}\mathbf{U})^*.$$

So, although $\mathbf{\Sigma}$ is not diagonal, we can *decorrelate* it by assuming an *effective* weight matrix $\mathbf{W}_e = \mathbf{W}\mathbf{U}$ and an *effective* decorrelated initial covariance matrix $\mathbf{\Sigma}_e = \mathbf{D}$. The final covariance matrix will be

$$\mathbf{C}_X = \mathbf{W}_e \cdot \mathbf{\Sigma}_e \cdot \mathbf{W}_e^*.$$

With this formulation, results in eqs. 33-34 are totally valid for the correlated case.

Acknowledgments

Statistical Noise Analysis in SENSE Parallel MRI

Santiago Aja-Fernández, Gonzalo Vegas-Sánchez-Ferrero, Antonio Tristán-Vega*

Abstract A complete first and second order statistical characterization of noise in SENSE reconstructed data is proposed. SENSE acquisitions have usually been modeled as Rician distributed, since the data reconstruction takes place into the spatial domain, where Gaussian noise is assumed. However, this model just holds for the first order statistics and obviates other effects induced by coils correlations and the reconstruction interpolation. Those effects are properly taken into account in this study, in order to fully justify a final SENSE noise model. As a result, some interesting features of the reconstructed image arise: (1) There is a strong correlation between adjacent lines. (2) The resulting distribution is non-stationary and therefore the variance of noise will vary from point to point across the image. Closed equations for the calculation of the variance of noise and the correlation coefficient between lines are proposed. The proposed model is totally compatible with g-factor formulations.

1 Introduction

An accurate statistical model of signal and noise is the keystone for many different applications in medical image processing and in the Magnetic Resonance (MR) field in particular. Traditionally, noise filtering techniques are based on a well-defined prior data statistical model. Many examples can be found in literature, such as the Conventional Approach [4], ML [7] and LMMSE [8, 9, 10] estimators or unbiased non-local mean filters [11, 12, 13]. A proper noise modeling may be useful not only for filtering purposes, but for many other processing techniques. Lately, for instance, Weighted Least Squares methods to estimate the Diffusion Tensor have proved to

---

be nearly optimal when the data follows a Rician [14] or a non-central Chi (nc-χ) distribution [109, 15].

For practical purposes, the modeling is usually done assuming noise in MR data is a zero-mean spatially uncorrelated Gaussian process with equal variance in both the real and imaginary parts in each acquisition coil. As a result, in single coil systems magnitude data in the spatial domain are modeled using a stationary Rician distribution [5]. When multiple coils are considered and the **k**-space is fully sampled, the natural extension of the Rician model yields to a stationary nc-χ distribution, whenever the different images are combined using sum of squares, the variance of noise is the same for all coils, and no correlations exist between them. However, multiple coils systems are preferably set up to work with subsampled **k**-space data, which have to be combined by some means to avoid the inherent aliasing artifact introduced by the under-completeness of the Fourier domain. The most popular algorithms to accomplish this task are GRAPPA [33] and SENSE [32], the latter focusing the discussion in the present paper.

While the nc-χ has been used to describe noise in GRAPPA [51, 42], SENSE acquisitions have been usually modeled as Rician distributed, owing to the computation of the modulus of the complex signal linearly obtained from the array of coils [43]. However, this model stands exclusively for the first order statistics of noise, and obviates some other side effects induced by coils correlations and **k**-space interpolation. For example, in GRAPPA reconstructions, both the initial inter-coil codependence and the **k**-space interpolation introduce a strong correlation between the signals to be combined, and the nc-χ model is not strictly fulfilled. This inaccuracy can be worked around introducing an effective value of the power of noise and an effective number of receivers in the nc-χ distribution [42, 49], both of them being spatially dependent.

For SENSE, the subsampling/interpolation effects have been previously described through the so-called g-factor, a global parameter that explicitly measures the SNR degradation in the acquisition process [52]. In this work we aim at fully characterizing the first and second order statistics of noise in SENSE reconstructed images, including the effects of **k**-space subsampling and inter-coil noise correlations. Since the reconstruction will take place into the spatial domain, and it can be seen as a weighted combination of the subsampled coils, the reconstructed image will be modeled as a complex Gaussian distribution. Its magnitude will be a nonstationary Rician distribution, with a spatial pattern that we can predict from certain imaging parameters such as the coils sensitivities and the speed-up factor (first order characterization). The study will show, in addition, another interesting feature of the reconstructed image: there exists a strong correlation between adjacent lines in the reconstructed volume, with an extent directly dependent on the acceleration factor (second order characterization).

2 Theory

2.1 Statistical Model of MR signals

The **k**-space data at each coil of the MR scanner can be accurately described by a noise-free signal plus an Additive White Gaussian Noise (AWGN) process, with zero mean and variance $\sigma_{K_l}^2$:

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; 0, \sigma_{K_l}^2), \quad l = 1, \cdots, L \tag{1}$$

with $a_l(\mathbf{k})$ the noise-free signal and $n_l(\mathbf{k}; 0, \sigma_{K_l}^2) = n_{l_r}(\mathbf{k}; 0, \sigma_{K_l}^2) + j n_{l_i}(\mathbf{k}; 0, \sigma_{K_l}^2)$ the AWGN process, which is initially assumed stationary so that $\sigma_{K_l}^2$ does not depend on **k**. The complex **x**-space is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice or volume, so the noise in the complex **x**-space is still assumed to be Gaussian:

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; 0, \sigma_l^2), \quad l = 1, \cdots, L$$

where $N_l(\mathbf{x}; 0, \sigma_l^2) = N_{l_r}(\mathbf{x}; 0, \sigma_l^2) + j N_{l_i}(\mathbf{x}; 0, \sigma_l^2)$ is also a complex AWGN process (assuming that there are not any spatial correlations) with zero mean and covariance matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1L}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \cdots & \sigma_{2L}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L1}^2 & \sigma_{L2}^2 & \cdots & \sigma_L^2 \end{pmatrix},$$

The variance of noise for each coil in **k**- and **x**-spaces are related through the number of points in the image:

$$\sigma_l^2 = \frac{1}{|\Omega|} \sigma_{K_l}^2 \tag{2}$$

with $|\Omega|$ the size of the image in each coil, i.e. the number of points used in the 2D iDFT. If the **k**-space is fully sampled, the Composite Magnitude Signal (CMS) can be directly obtained using SoS [27, 38]:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l(\mathbf{x})|^2}. \tag{3}$$

For a single–coil acquisition, the CMS, $M(\mathbf{x})$, is the Rician distributed envelope of the complex signal [5]. In the image background, where the signal-to-noise ratio is zero due to the lack of water-proton density in the air, the Rician simplifies to a Rayleigh distribution. For multiple coils, if the variance of noise is the same for all coils, no correlation exists between them, and the signals are combined using SoS, the CMS may be modeled as a nc-χ distribution [27, 46, 47, 42, 48]. In a more general case where correlations are taken into account, the nc-χ is only an
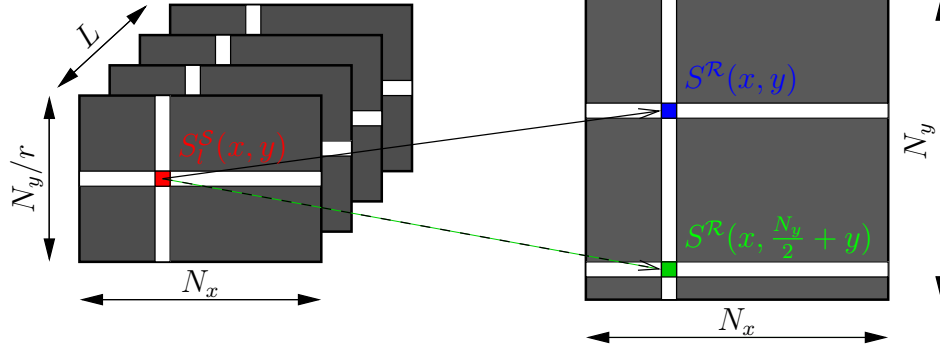
Fig. 1 Example of the SENSE interpolation for 4 coils and an acceleration factor $r = 2$.

approximation of the real distribution. It can be accurately approximated with this model if effective parameters (reduced number of coils and and increased variance of noise) are used [49].

2.2 Statistical model in SENSE reconstructed images

For the sake of simplicy, let us assume that SENSE [32] is only be applied to MRI data regularly subsampled by a factor r. The reconstruction takes place in the image domain. Assuming an original size $|\Omega| = M_x \times M_y$, the subsampled signal in the **x**-space $S_l^{\mathcal{S}}(\mathbf{x}) = S_l^{\mathcal{S}}(x, y)$ is the (complex) Fourier inverse transform of $s_l^{\mathcal{S}}(\mathbf{k})$, of size $M_x \times (M_y/r)$. Note than since the size of subsampled data in each coil in the **x**-space is reduced by a factor r, the variance of noise will be umplified by that same factor:

$$\sigma_l^2 = \frac{r}{|\Omega|}\sigma_{K_l}^2$$

In multiple coil scanners, the image received in coil l-th, $S_l(x, y)$, can be seen as an *original image* $S_0(x, y)$ weighted by the sensitivity of that specific coil:

$$S_l(x, y) = C_l(x, y)S_0(x, y), \quad l = 1, \cdots, L \tag{4}$$

An accelerated pMRI acquisition with a factor r will reduce the matrix size of the image at every coil. The signal in one pixel at location (x, y) of l-th coil can be now written as [34]:

$$S_l(x, y) = C_l(x, y_1)S_0(x, y_1) + \cdots + C_l(x, y_r)S_0(x, y_r) \tag{5}$$

In SENSE, the reconstructed image $S^{\mathcal{R}}(x, y)$ can be seen as an estimator of the original image $S^{\mathcal{R}}(x, y) = \widehat{S_0}(x, y)$ that can be obtained from eq. (4). For instance, for $r = 2$ for pixel (x, y), $S^{\mathcal{R}}(x, y)$ is obtained as
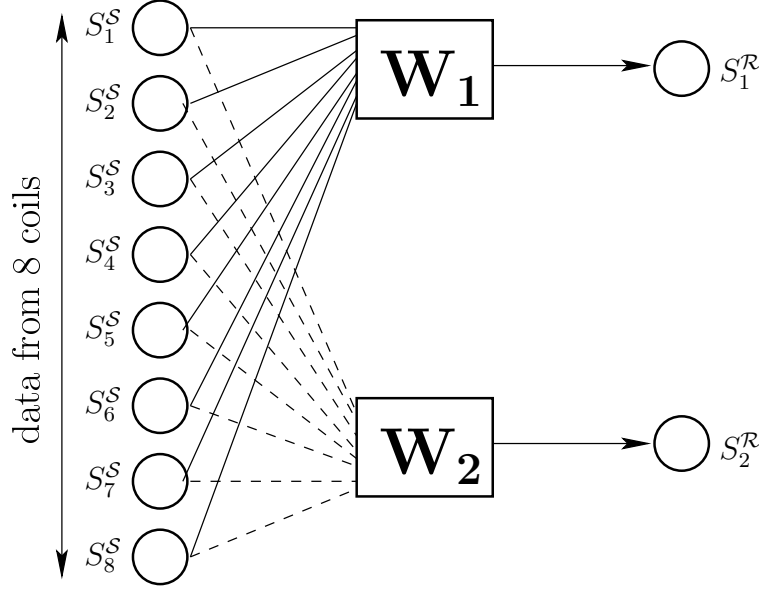
Fig. 2 Example of the SENSE interpolation for 8 coils and an acceleration factor $r = 2$.

$$\begin{bmatrix} S_1^{\mathcal{R}} \\ S_2^{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_2 \end{bmatrix} \times \begin{bmatrix} S_1^{\mathcal{S}} & \cdots & S_L^{\mathcal{S}} \end{bmatrix}. \tag{6}$$

In matrix form:

$$\mathbf{S}^{\mathcal{R}} = \mathbf{W} \times \mathbf{S}^{\mathcal{S}} \tag{7}$$

with $\mathbf{W}(x,y) = [\mathbf{W}_1, \cdots, \mathbf{W}_r]$ a reconstruction matrix created from the sensitivity map of each coil, $\mathbf{C}(x,y) = [\mathbf{C}_1, \cdots, \mathbf{C}_l]$:

$$\mathbf{W}(x,y) = (\mathbf{C}^*(x,y)\mathbf{C}(x,y))^{-1}\mathbf{C}^*(x,y)$$

If the correlation between coils is taken into account, the reconstruction matrix must incorporate the covariance matrix:

$$\mathbf{W}(x,y) = (\mathbf{C}^*(x,y)\mathbf{\Sigma}^{-1}\mathbf{C}(x,y))^{-1}\mathbf{C}^*(x,y)\mathbf{\Sigma}^{-1}$$

For the sake of simplicity, we will remove any pixel dependency, so that we can write for each output pixel:

$$S_i^{\mathcal{R}} = \mathbf{W}_i \times \mathbf{S}^{\mathcal{S}} \quad i = 1, \cdots, r \tag{8}$$

Two examples can be found on Fig. 1 and on Fig. 2.

The SNR of the fully sampled image and the image reconstructed with SENSE are related by the so-called g-factor, g [52, 34]:

$$\mathrm{SNR}_{\mathrm{SENSE}} = \frac{\mathrm{SNR}_{\mathrm{full}}}{\sqrt{r} \cdot g} \tag{9}$$

However, we will focus on the actual noise model underlying the SENSE reconstruction and on the final variance of noise. The final signal $S_i^{\mathcal{R}}$ is obtained as a linear combination of $S_l^{\mathcal{S}}$, where the noise is Gaussian distributed. Thus, the resulting signal is also Gaussian, with variance:

$$\sigma_i^2 = \mathbf{W}_i^* \boldsymbol{\Sigma} \mathbf{W}_i \tag{10}$$

Since $\mathbf{W}_i$ is position dependent, i.e. $\mathbf{W}_i = \mathbf{W}_i(x, y)$, so will be the variance of noise, $\sigma_i^2(x, y)$. For further reference, when the whole image is taken into account, let us denote the variance of noise for each pixel in the reconstructed data by $\sigma_{\mathcal{R}}^2(\mathbf{x})$.

Note now that all the lines $S_i^{\mathcal{R}}$ reconstructed from the same data $S_l^{\mathcal{S}}$ will be strongly correlated, since they are basically different linear combinations of the same Gaussian variables. In that case, the covariance between $S_i^{\mathcal{R}}$ and $S_j^{\mathcal{R}}$, $i \neq j$ can be calculated as

$$\sigma_{i,j}^2 = \mathbf{W}_i^* \boldsymbol{\Sigma} \mathbf{W}_j \tag{11}$$

and the correlation coefficient is derived straight forward:

$$\rho_{i,j}^2 = \frac{\sigma_{i,j}^2}{\sigma_i \sigma_j} = \frac{\mathbf{W}_i^* \boldsymbol{\Sigma} \mathbf{W}_j}{\sqrt{\left(\mathbf{W}_i^* \boldsymbol{\Sigma} \mathbf{W}_i\right)\left(\mathbf{W}_j^* \boldsymbol{\Sigma} \mathbf{W}_j\right)}}, \tag{12}$$

However, these correlations are not strongly affecting the data, since the correlated pixels are separated by N_y/r lines

All in all, noise in the final reconstructed signal $S^{\mathcal{R}}(x, y)$ will follow a complex Gaussian distribution. If the magnitude is considered, i.e. $M(x, y) = |S^{\mathcal{R}}(x, y)|$, the final CMS will follow a Rician distribution, just like single-coil systems.

We can summarize our developments as follows:

1. Subsampled multi coil MR data reconstruted with cartesian SENSE follows a Rician distribution in each point of the image.
2. The resulting distribution is non-stationary. This means that the variance of noise will vary from point to point across the image.
3. The variance of noise final value in each point will only depend on the covariance matrix of the original data and on the sensitivity map.
4. Each pixel in the final image will be strongly correlated with all those pixels reconstructed from the same original data. Each pixel is correlated with $r - 1$ other pixels. These correlated pixels are far enough and they can be neglected.

For the particular case in which there is no correlation between coils and all the coils has the same noise variance σ_n^2, we can write eq. (9) as:

$$\sigma_i^2 = \sigma_n^2 \times |\mathbf{W}_i|^2 \tag{13}$$

Since σ_n^2 is the noise variance for the subsampled data in the **-x**-space, according to eq. (2), it is related to the original noise level without subsampling, say σ_0^2, by the

subsampling rate:

$$\sigma_n^2 = r \cdot \sigma_0^2$$

and therefore

$$\sigma_i = \sqrt{r} \cdot \sigma_0 \times |\mathbf{W}_i| \tag{14}$$

which is totally equivalent to the g-factor formulations for SNR reduction in literature [32, 43].

3 Materials and Methods

For the sake of validation and illustration of the results in the previous section, the following experiments are considered:

First, we will study the statistical behavior of Gaussian data when a combination like the one in SENSE is done. To that end, we consider 10^5 samples of 8 correlated complex Gaussian RVs with zero mean and unitary variance, $N(0,1)$ and two different correlation coefficients, $\rho^2 = 0$ and $\rho^2 = 0.2$. (Note the correlation is between variables, not between samples of the same variable). The RVs are combined using real random weights, $\mathbf{W}_1$ and $\mathbf{W}_2$, both following a uniform distribution in $[0, 1]$ and normalized so that

$$|\mathbf{W}_i|^2 = 1, \quad i = 1, 2$$

Two new variables are created by using a combination like the one in eq. (6), obtaining two new RVs. The sample variance and correlation coefficient are estimated from the data and then compared to those calculated from eq. (9) and eq. (12).

Secondly, we will test how the values of $\sigma_{\mathcal{R}}^2(\mathbf{x})$ varies across the image. To that end, we will work with one sensitivity map synthetically generated, as shown in Fig. 1 (top). This map simulates an 8-coil system using an artificial sensitivity map coded for each coil so that $\sum_l |C_l(\mathbf{x})|^2 = 1$, with $l = 1 \cdots, 8$, and $C_l(\mathbf{x})$ the sensitivity map of coil l-th. For the experiment:

- We assume that each coil has an original variance of noise $\sigma_l^2 = 100$. We will simulate two different configurations, first, assuming that there is no initial correlation between coils, and second, assuming a correlation coefficient of $\rho^2 = 0.1$ between all coils, so that

$$\mathbf{\Sigma} = 100 \times \begin{pmatrix} 1 & 0.1 & \cdots & 0.1 \\ 0.1 & 1 & \cdots & 0.1 \\ \vdots & \vdots & \ddots & \vdots \\ 0.1 & 0.1 & \cdots & 1 \end{pmatrix}.$$

- From the data, and using the theoretical expressions in eq. (9) and eq. (12) we calculate the variance of noise for each pixel in the final image.
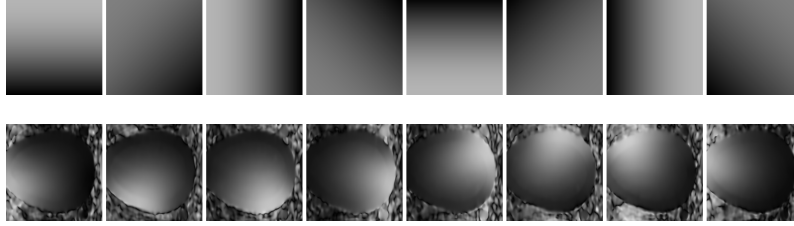
Fig. 3 Sensitivity Maps used for the experiment. Top: Synthetic sensitivity map created so that the SoS of the maps gives a constant image. Bottom: Sensitivity map estimated from an actual brain imaging acquisition in a SENSE Signa 1.5T scanner with 8 coils.

- In order to test the theoretical distributions, 5000 samples of 8 complex 256×256 Gaussian images with zero mean and covariance matrix $\boldsymbol{\Sigma}$ are generated. The **k**-space of the data is subsampled by a 2x factor and reconstructed using SENSE and the synthetic sensitivity field. We estimate the variance of noise in each point using the second order moment of the Rayleigh distribution [47]:

$$\sigma_{\mathcal{R}}^2(\mathbf{x}) = \frac{1}{2} E\{M^2(\mathbf{x})\}$$

We estimate the $E\{M^2(\mathbf{x})\}$ along the 5000 samples.

Last, the previous experiment is repeated for the correlated case, now using real sensitivity maps as shown in Fig. 1 (bottom). These maps are estimated from a real brain T1 acquisition done in a GE Signa 1.5T EXCITE, FSE pulse sequence, 8 coils, TR=500msec, TE=13.8msec, image size 256×256 and FOV: 20cm$\times$20cm.

4 Results

Experiment	Parameter	Sample Value	Theoretical
$\rho^2 = 0$	σ_1	1.0072	1.0000
	σ_2	1.0023	1.0000
	$\rho_{1,2}^2$	0.9262	0.9247
$\rho^2 = 0.2$	σ_1	1.4198	1.4121
	σ_2	1.4536	1.4412
	$\rho_{1,2}^2$	0.9706	0.9703

Table 1 Results from the first experiment. Standard deviation and correlation coefficient of the SENSE-like combination of synthetic Gaussian data. Theoretical and sample values.
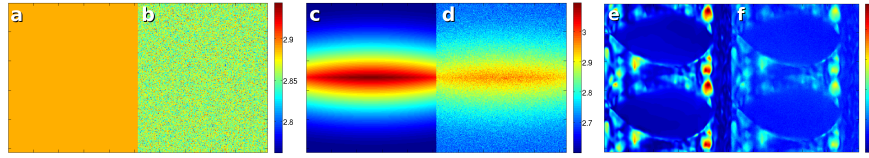
Fig. 4 Maps of standard deviation of noise $\sigma_{\mathcal{R}}(\mathbf{x})$ in the final image: (a-c-e): Theoretical values. (b-d-f): Estimated from samples. (a-b) Synthetic Sensitivity Map with no correlation. (c-d) Synthetic Sensitivity Map with correlation between coils. (e-f) Real sensitivity map with correlation between coils.

The **first** experiment studies the behavior of a SENSE-like combination of Gaussian data. Results of the standard deviation and correlation for the two resulting variables are collected in Table 1. As expected, the theoretical values match the estimation through samples. Note that the variances in the final RVs in the correlated case are higher than the ones in the case without correlation. This effect can be found in real data, where correlations exist and must be taken into account.

The **second** and **third** experiments deals with the non homogeneous spatial layout of the noise and the influence of the sensitivity map over the final noise distribution. Visual results are depicted in Fig 2. For the synthetic maps, when no correlations are considered, since $\sum_l |C_l(\mathbf{x})|^2 = 1$ for all pixels, the final variance of noise will not depend on the position $\mathbf{x}$. Therefore, in this particular case $\sigma_{\mathcal{R}}^2(\mathbf{x}) = \sigma_{\mathcal{R}}^2$. The estimated values in Fig 2-(b) show a noise pattern that slightly varies around the real value (note the small range of variation). In this very particular case, the noise can be considered to be spatially stationary, and the final image (leaving the correlation between pixels aside) is equivalent to one obtained from a single-coil scanner.

When correlations are taken into account, even using the same synthetic sensitivity map, results differ. In Fig. 2-(c), the theoretical value shows that the standard deviation of noise of the reconstructed data is not the same for every pixel, i.e., the noise is no longer spatial-stationary. The center of the image shows a larger value that decreases going north and south. So, in this more realistic case, the $\sigma_{\mathcal{R}}^2(\mathbf{x})$ will depend on $\mathbf{x}$, which can have serious implications for future processing, such as model based filtering techniques. The estimated value in Fig. 2-(d) shows exactly the same non-homogeneous pattern across the image.

In the last experiment, Fig. 2-(e) and Fig. 2-(f), a real sensitivity map is used, and correlation between coils is also assumed. Again, the noise is non-stationary. To increase the dynamic range of the images, the logarithm has been used to show the data.

5 Discussion and conclusions

Noise analysis in SENSE in literature is usually focused on the study of the SNR loss due to the acceleration process, being the so-called g-factor the most common measure. However, despite its proved utility, the g-factor is insufficient when trying to design certain MR applications. In those cases, the need of a complete statistical modeling arises naturally. In SENSE, the Rician model has been widely assumed by the MR community, as seen in recent literature [46, 43]. This assumption has traditionally been taken as a guarantee to use the methods designed for single coil MR data over multiple-coil SENSE reconstructed data.

The study of the noise distribution in SENSE carried out in this work has brought to light some serious implications that must have been taken into account when working with SENSE data. Even when the final distribution is always Rician, depending on the sensitivity maps and on the coils covariance matrix, this distribution is likely to be non-stationary. In the examples proposed, even in the optimal synthetic case, when correlations between coils were present, the non-stationarity arises. As a consequence, the variance of noise will differ from point to point across the image. The first implication of this feature is related with noise estimation. Since the variance of noise depends on the position, most noise estimation techniques, as the ones proposed in [47], have no longer sense. There is no longer a single value to estimate for the whole image, but one for each point. Thus, those algorithms that need an estimation of the level of noise, cannot be used either in its original shape. See, for instance, those filtering techniques reviewed in the Introduction.

As an example, let us consider a very simple noise reduction technique, the Conventional Approach (CA) for Rician data [4]:

$$\widehat{A}(\mathbf{x}) = \sqrt{E\{M_L^2(\mathbf{x})\} - 2\sigma_n^2},$$

whose same philosophy is shared by several other methods mainly based on NLM denoising [11, 12, 13]. When dealing with SENSE reconstructions, the variance of noise σ_n^2 will no longer be unique for the whole image, i.e. $\sigma_n^2(\mathbf{x})$. The noise estimation is no longer a simple task. Some prior data *regulation* like the one done in [63] could be necessary. Note that, even an approach as simple as the CA cannot be directly applied from the single coil Rician formulation over SENSE data.

To sum up, SENSE MR data is known to follow a Rician distribution, but non-homogeneity of the variance of noise arises due to the reconstruction process. A prior knowledge of the sensitivity maps and the coils covariance matrix will help to properly design applications to deal with this kind of data.

Acknowledgments

Part II
Noise estimators for non-accelerated acquisitions

Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images: A LMMSE approach

Santiago Aja-Fernández, Carlos Alberola-López and Carl-Fredrik Westin[*]

Abstract A new method for noise filtering in images that follow a Rician model —with particular attention to Magnetic Resonance Imaging (MRI)— is proposed. To that end, we have derived a (novel) closed-form solution of the Linear Minimum Mean Square Error (LMMSE) estimator for this distribution. Additionally, a set of methods that automatically estimate the noise power are developed. These methods use information of the sample distribution of local statistics of the image, such as the local variance, the local mean and the local mean square value. Accordingly, the dynamic estimation of noise leads to a recursive version of the LMMSE, which shows a good performance in both noise cleaning and feature preservation. The paper also includes the derivation of the probability density function of several local sample statistics for the Rayleigh and Rician model, upon which the estimators are built.

1 Introduction

Noise in magnitude Magnetic Resonance (MR) images is usually modeled by means of a Rician distribution, due to the existence of zero-mean uncorrelated Gaussian noise with equal variance in both the real and imaginary parts of the complex k-space data [44, 5]. This noise may affect the performance of different post-processing techniques applied to MR data, such as segmentation, registration or tensor estimation in Diffusion Tensor MRI (DT-MRI) [87]. Accordingly, a great amount of noise-removal methods has been reported in the literature.

One of the first attempts proposed to estimate the magnitude MR image from a noisy image is due to Henkelman [36] who investigated the effect of noise on MR

---

magnitude images. The author showed that the noise influence leads to an overestimation of the signal amplitude and proposed a correction scheme based on image intensities. The so-called *conventional approach* (CA) was proposed by McGibney *et al.* [4] utilizing the noise properties of the second-order moment. Sijbers *et al.* [80, 7, 82] estimate the Rician noise level and perform signal reconstruction using a maximum likelihood (ML) approach. A similar method is used by Jiang and Yang [114]. Expectation maximization (EM) formulations with Rician noise assumptions have been used in SAR imaging [98, 89]. Fillard *et al.* [110] recently proposed another ML-based scheme that uses an anisotropic regularization term to exploit the spatial correlation, and apply it to correct the bias in tensor estimation in DT-MRI.

Other approaches use wavelet-based methods for noise removal, as Nowak's [86] —in which the authors assume an underlying Rician model—- or the one due to Pižurica *et. al* [115]. Lysaker *et al.* [116] proposed a method for noise removal in MRI using fourth-order partial differential equations. McGraw *et al.* [117] use a weighted total-variation-norm denoising scheme and Ahn *et al.* [118] propose a template-based filtering procedure; none of this approaches use a Rician noise model. Basu *et al.* [87] use a Perona-Malik-like smoothing filter combined with a local Rician data attachment term (effectively trying to remove the intensity bias locally), assuming a known noise level for the Rician noise model. Using local statistics as priors, Awate and Whitaker [119] introduced a method to denoise a MR image using nonparametric neighborhood statistics. Recently, Koay and Basser in [48] developed a correction scheme that allows to analytically estimate the signal, also assuming the Rician model, and in [59] Martin-Fernandez *et al*. proposed a Wiener-filter approach with a correction for MRI data.

In this paper we propose a new method for noise filtering of MRI data, and by extension, of images that follows a Rician distribution, by using the Linear Minimum Mean Square Error (LMMSE) estimator for Rician noise.

As it is well-known, LMMSE estimators are a trade-off between optimality and simplicity; we obtain a closed-form analytical solution for our estimator which makes the filtering process far simpler than other estimation techniques that find the solution via an iterative optimization scheme. Results from our method are satisfactory as the experiments in Section 5 show. Additionally, the goodness of the filter will be intrinsically related to the accuracy of the estimation of the noise variance in the images. In the paper we will present four novel techniques to estimate this parameter; although three of them rely on the assumption of a uniform background, none of the cases require any sort of background segmentation.

The paper is organized as follows: Section 2 is a background section on Rician and Rayleigh distributions and their estimators. In Section 3 the LMMSE estimator for Rician noise is presented. Then, Section 4 describes the different procedures to estimate the noise. Some experiments have been added in Section 5. A number of appendices have been included for readability purposes.

2 Background

2.1 Rician distributed data

If both real and imaginary parts of a signal are corrupted with zero-mean uncorrelated Gaussian noise with with equal variance, the envelope of the magnitude signal will follow a Rician distribution. This kind of noise may be encountered in many practical situations, such as MR data, speckle [98], communication fading problems [120] and many others. The magnitude signal may be expressed

$$M = \sqrt{(A + n_1)^2 + n_2^2} \tag{1}$$

being M the magnitude image, A the original signal level if no noise is present and n_1 and n_2 uncorrelated Gaussian noise variables with zero mean and equal variance σ_n^2. The probability density function (PDF) of such an image is, as indicated, a Rician distribution [44, 45]

$$p_M(M|A, \sigma_n) = \frac{M}{\sigma_n^2} \, e^{-\frac{M^2 + A^2}{2\sigma_n^2}} \, I_0\left(\frac{AM}{\sigma_n^2}\right) \, u(M) \tag{2}$$

being $I_0(.)$ the 0th order modified Bessel function of the first kind and $u(.)$ the Heaviside step function. Although generally speaking the moments of this distribution are difficult to calculate, the even-order (non central) moments are simple polynomials as, for instance, he second order moment

$$\mu_2 = E\{M^2\} = A^2 + 2\sigma_n^2 \tag{3}$$

In the image background, where the signal-to-noise ratio is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution [45] with PDF

$$p_M(M|\sigma_n) = p_M(M|A = 0, \sigma_n) = \frac{M}{\sigma_n^2} \, e^{-\frac{M^2}{2\sigma_n^2}} \, u(M). \tag{4}$$

2.2 Parameter estimation of the Rician distribution

In this section we review a number of methods to estimate parameters in a Rician distribution. Although many approaches have been reported in the literature (see, for instance, the papers cited in the introduction) we will focus now only on those based on a stochastic model. Most of the papers cited hereafter focus on MRI processing, although some of them deal with other type of images, such as SAR [98].

The *Conventional Approach* (CA) [4] accounts for the relation between noise and signal of the second order moment in a Rician distribution, see eq. (3), so the signal

can be estimated as

$$\hat{A} = \sqrt{\langle M^2 \rangle - 2\sigma_n^2} \tag{5}$$

being $\langle M^2 \rangle$ the sample second order moment. Other approaches use the *Maximum Likelihood estimator* (ML) [80, 7, 82, 114] where the signal is estimated by maximizing the likelihood function. Koay and Basser in [48] propose an analytically exact solution assuming a Rician noise model based on image statistics and the noise level. In these three methods the variance of noise σ_n^2 is an input value, so it must be known before hand, or at least, it must be estimated. The *Expectation-Maximization* (EM) method [98, 89] provides a recursive scheme that aims at estimating the noise variance and the signal simultaneously through maximization of the expected log likelihood.

Note that the last three schemes (ML, EM and Koay's) are originally designed to estimate the signal from multiple samples. When only one image is available, the statistics must be calculated locally. In addition, in the three methods the solution is found via an iterative optimization scheme.

The different methods reported to estimate σ_n^2 from the magnitude data in Rician distributions may be roughly divided into two vast groups, namely, those that use a single magnitude image and those that use multiple images. In this paper we will mainly deal with methods in the former group.

The estimation of noise in MRI using a single image is usually done out of the background pixels, where the signal is assumed to be zero. According to eq. (3), in the areas where the signal is zero $A^2 = 0$, and then $E\{M^2\} = 2\sigma_n^2$. So a straightforward estimator —based on the method of moments— will be [7, 86, 81]

$$\widehat{\sigma_n^2} = \frac{1}{2N} \sum_{i=1}^{N} M_i^2 \tag{6}$$

being N the number of points considered for the estimation. Eq. (6) is in fact, an unbiased estimator of σ_n^2 with variance

$$\mathrm{Var}\left(\widehat{\sigma_n^2}\right) = \frac{\sigma_n^4}{N} \tag{7}$$

This estimator is also the Maximum Likelihood estimator of this parameter for the Rayleigh distribution [7].

An alternative way to estimate σ_n in non-signal areas is obtained by using the first order moment of the Rayleigh PDF. An unbiased estimator then will be [7]

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi} \frac{1}{N} \sum_{i=1}^{N} M_i} \tag{8}$$

In [114] the authors propose an estimator based on the method of moments for a Rician distribution

$$\hat{\sigma_n}^2 = \frac{1}{2} \left(\langle M^2 \rangle - \left(2\langle M^2 \rangle^2 - \langle M^4 \rangle \right)^{1/4} \right)$$

Recently Sijbers *et al.* [83, 84] proposed a new method to estimate σ_n from the image histogram. As in the background the image distribution is Rayleigh, and since the mode of this distribution equals the parameter σ_n, then $\hat{\sigma_n} = M_{\max}$, being $M_{\max}$ the mode of the histogram of the image. Noise can also be estimated using multiple images taken under identical conditions [81]. Finally, in [14] noise is estimated by using the information of several images and how they match a model.

3 Signal estimation: the LMMSE estimator

The method proposed in this paper to estimate the signal from the magnitude image is based on the Linear Minimum Mean Square Error (LMMSE) estimator [121]. Our aim is to find a closed-form estimator for a signal that follows a Rician distribution. This is in contrast to many other estimation techniques —such as ML and EM— which find the solution via an iterative optimization scheme. Closed-form solutions make estimation methods computationally more efficient than optimization-based solutions. In addition, the estimator will be based on local statistics, so one single (noisy) image —as opposed to several perfectly aligned images— will suffice to estimate the signal.

As previously stated, the moments of the Rician distribution have a non-trivial integral expression but for even-order moments, which are simple polynomials. In order to achieve a closed-form expression we will use A^2 instead of A. Consequently, all the moments to be used hereafter will be even.

3.1 LMMSE filter derivation for the general Rician model

The LMMSE estimator[2] of a parameter θ is defined [121]

$$\hat{\theta} = E\{\theta\} + \mathbf{C}_{\theta\mathbf{x}}\mathbf{C}_{\mathbf{xx}}^{-1}(\mathbf{x} - E\{\mathbf{x}\}) \tag{9}$$

being $\mathbf{x}$ the vector of available samples, $\mathbf{C}_{\mathbf{xx}}$ the covariance matrix of $\mathbf{x}$ and $\mathbf{C}_{\theta\mathbf{x}}$ the cross-covariance vector. Rewriting eq. (9) for a 2D signal with a Rician distribution

$$\widehat{A_{ij}^2} = E\{A_{ij}^2\} + \mathbf{C}_{A_{ij}^2 M_{ij}^2}\mathbf{C}_{M_{ij}^2 M_{ij}^2}^{-1}\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right) \tag{10}$$

where A_{ij} is the unknown intensity value in pixel ij and $\mathbf{M}_{ij}$ the observation vector. If the estimator is simplified to be pointwise, vectors and matrices become scalar values. Then

---

[2] As it is well known, this type of estimators model the parameter to be estimated as a sample of a random variable, the parameters of which have a known relation with the parameters of the observation [121].

$$\mathbf{C}_{M_{ij}^2 M_{ij}^2} = E\{ \left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\} \right) \left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\} \right)^T \}$$
$$= E\{ \left(M_{ij}^2 - E\{M_{ij}^2\} \right)^2 \} = E\{M_{ij}^4\} - E\{M_{ij}^2\}^2$$
$$= C_{M_{ij}^2 M_{ij}^2}$$

$$\mathbf{C}_{A_{ij}^2 M_{ij}^2} = E\{ \left(A_{ij}^2 - E\{A_{ij}^2\} \right) \left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\} \right)^T \}$$
$$= E\{ \left(A_{ij}^2 - E\{A_{ij}^2\} \right) \left(M_{ij}^2 - E\{M_{ij}^2\} \right) \}$$
$$= C_{A_{ij}^2 M_{ij}^2}$$

and making use of the model in eq. (1)

$$C_{A_{ij}^2 M_{ij}^2} = E\{A_{ij}^4\} + 2E\{A_{ij}^2\}\sigma_n^2 - E\{A_{ij}^2\}E\{M_{ij}^2\}$$

Finally, the LMMSE estimator is

$$\widehat{A_{ij}^2} = E\{A_{ij}^2\} + \frac{E\{A_{ij}^4\} + 2E\{A_{ij}^2\}\sigma_n^2 - E\{A_{ij}^2\}E\{M_{ij}^2\}}{E\{M_{ij}^4\} - E\{M_{ij}^2\}^2} \left(M_{ij}^2 - E\{M_{ij}^2\} \right) \tag{11}$$

Assuming local ergodicity, the expectation may be replaced by its sample estimator $\langle . \rangle$, that can be defined

$$\langle I_{i,j} \rangle = \frac{1}{|\eta_{i,j}|} \sum_{p \in \eta_{i,j}} I_p \tag{12}$$

with $\eta_{i,j}$ a square neighborhood around the pixel ij. This estimation may be also done using non-square weighted windows, such as Gaussian functions. Using the relations from eq. (1)

$$E\{M_{ij}^2\} = E\{A_{ij}^2\} + 2\sigma_n^2$$
$$E\{M_{ij}^4\} = E\{A_{ij}^4\} + 8\sigma_n^2 E\{A_{ij}^2\} + 8\sigma_n^4$$

and $\langle . \rangle$, the LMMSE estimator may finally be written as

$$\widehat{A_{ij}^2} = \langle M_{ij}^2 \rangle - 2\sigma_n^2 + K_{ij} \left(M_{ij}^2 - \langle M_{ij}^2 \rangle \right) \tag{13}$$

with K_{ij}

$$K_{ij} = 1 - \frac{4\sigma_n^2 \left(\langle M_{ij}^2 \rangle - \sigma_n^2 \right)}{\langle M_{ij}^4 \rangle - \langle M_{ij}^2 \rangle^2}. \tag{14}$$

Note that the σ_n^2 value must be properly estimated. This task is usually done from a selected region from the background pixels, where the signal is assumed zero. In Section 4 some new automatic methods will be presented.

Since only pointwise dependence has been considered in the filter, the extension to an arbitrary number of dimensions is straightforward by changing the estimation neighborhood. For example, for 3D images

$$\langle I \rangle = \frac{1}{|\eta_{ijk}|} \sum_{p \in \eta_{ijk}} I_p$$

Some experiments have been done in section 5.1 to illustrate the LMMSE filtering performance.

4 Estimation of noise using local statistics

The performance of the estimator previously described, as well as other methods in the literature, is directly related to the quality of the estimate of the noise variance σ_n^2. Noise estimation schemes described in section 2.2 (but the one in [83]) have some disadvantages: they need the background pixels of the image to be manually selected and the estimation is done considering that the signal is always zero in the background pixels. These two considerations make the estimate sensitive to errors and artifacts. In this section we will propose a new approach based on local statistics. We will estimate the noise using the distributions of some sample local statistics of the image, such as the sample second order moment, the sample mean and the sample variance.

Using the second order moment and the mean we will develop two estimators that can be used only when the image has a background where the distribution can be assumed Rayleigh (i.e., a non-signal background). This is the case of many MR magnitude images. Then, we will define a variance-based estimator that can be used with any kind of image that follows a Rician distribution.

4.1 Noise estimator based on the local second order moment

Taking expectations in eq. (1) for zero-mean and equal variance noise components we obtain

$$E\{M_{i,j}^2\} = E\{A_{i,j}^2\} + 2\sigma_n^2 \tag{15}$$

being $E\{I_{i,j}^2\}$ the local second order moment of an image I. It may be estimated using a neighborhood $\eta_{i,j}$ centered around the pixel under analysis giving rise to $\widehat{\mu_2}_{i,j} = \widehat{E\{I_{i,j}^2\}} = \langle I_{i,j}^2 \rangle$. For the sake of simplicity we will use a square neighborhood, as in eq. (12).

If we assume that σ_n is constant throughout the image, the effect of the noise over the local second order moment distribution will be a shift to higher values. In those regions where $A_{i,j} = 0$, it holds

$$E\{M_{i,j}^2\} = 2\sigma_n^2$$

This fact has been used in the past to estimate the noise from MRI by selecting areas in the background (see Section 2.2). In this new approach, we will not be working

with selected regions, but with the distribution of the local statistics of the whole image.

As an effect of the noise, the shape of the local second order moment distribution remains fairly unchanged in the noisy image with respect to that in the original image, but for a right-shift of the whole distribution. See, for instance, the distribution of this moment for the image in Fig. 1 when Rician noise is present; this is depicted in Fig. 2-(a).
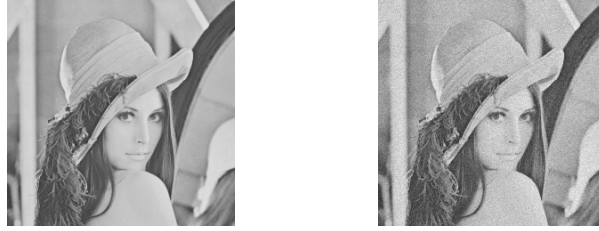

Fig. 1 Original image (left) and noisy counterpart (right) with $\sigma_n = 10$.
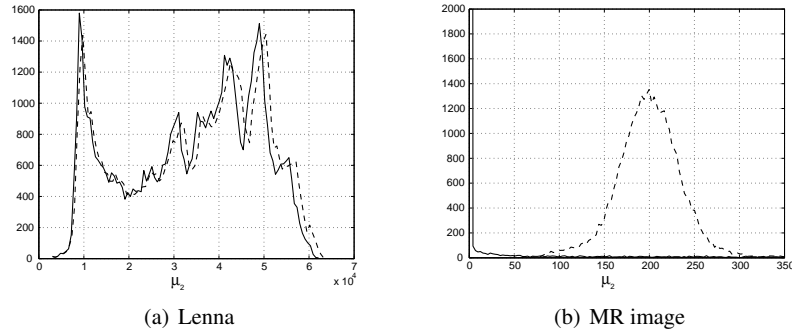
(a) Lenna   (b) MR image

Fig. 2 Distribution of the local second order moment. (a) Lenna image without noise (solid) and with Rician noise with $\sigma_n = 10$ (dashed). (b) MR image without noise (solid) and with noise (dashed), with $\sigma_n = 10$. A 7×7 window has been used.

However, in MR images —see for instance Fig. 3, taken from the BrainWeb database [53]—, due to the presence of an extensive background in which the signal is virtually zero, the distribution of the local second order moment of the original image will have a maximum in the origin, as shown in Fig. 2-(b) solid line. If the image is corrupted with Rician noise with $\sigma_n = 10$, the effect, according to eq. (15), will be a shift of the maximum from zero to $2\sigma_n^2$. This effect can be seen in Fig. 2-(b), dashed line, where the distribution of the local second order moment shows a bell shape with its maximum in 200, with in fact is $2\sigma_n^2$ when $\sigma_n = 10$.
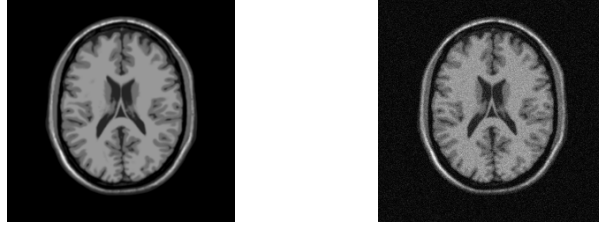
Fig. 3 Original magnitude MRI from the BrainWeb database (left) and noisy one (right) with $\sigma_n = 10$.

So, the position of this maximum may be used in order to estimate σ_n^2:

$$\widehat{\sigma_n^2} = \frac{1}{2} \arg \max \{ p(\widehat{\mu_2}) \}$$

being $p(\widehat{\mu_2})$ the distribution of the sample (local) second order moment of the noisy image. This maximum may be calculated as the mode of the distribution.

To analyze this assumption let us calculate the theoretical value of the mode. In MR images, the pixels in the background follow a Rayleigh distribution. So, let $R_i(\sigma^2)$, $i = \{1, \cdots, N\}$ be a set of independent and identically distributed (IID) Rayleigh random variables. Then [122]

$$\sum_{i=1}^{N} R_i^2(\sigma^2) \sim \gamma(N, 2\sigma^2), \ \ S = \frac{1}{N} \sum_{i=1}^{N} R_i^2(\sigma^2) \sim \gamma \left(N, \frac{2\sigma^2}{N} \right)$$

i.e., the sample local second order moment of a Rayleigh distribution follows a Gamma distribution with parameters $\alpha = N$ and $\beta = \frac{2\sigma^2}{N}$. The mode of this distribution is defined as $\text{mode}(S) = (\alpha - 1)\beta$, so, the mode of the sampling second order moment will be

$$\text{mode}(\widehat{\mu_2}_{i,j}) = \frac{N-1}{N} 2\sigma_n^2$$

If we redefine the estimator of the moment as

$$\widehat{\mu_2} = \frac{1}{N-1} \sum_{i=1}^{N} R_i^2(\sigma^2) \tag{16}$$

then

$$\text{mode}(\widehat{\mu_2}_{i,j}) = 2\sigma_n^2$$

and

$$\widehat{\sigma_n^2} = \frac{1}{2} \text{mode}(\widehat{\mu_2}_{i,j}) \tag{17}$$

To verify this hypothesis an experiment has been carried out for 40 different variances of noise. The MR image in Fig. 3 has been normalized between 0 and 1.

Local moments are calculated using 7×7 windows. For each σ_n^2 we would take the average value of 100 simulations. In Fig. 4-(a) the estimated variance is depicted versus the original one. In Fig. 4-(b), as a quality measure, we show the ratio $Q = \widehat{\sigma_n}/\sigma_n$.
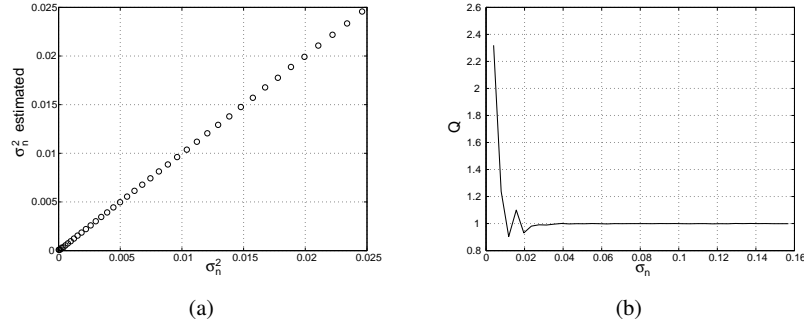
Fig. 4 Estimation of variance of noise using the mode of the local second order moment. (a) $\widehat{\sigma_n^2}$ vs. σ_n^2. (b) $\widehat{\sigma_n}/\sigma_n$ vs. σ_n

Let us now study how the window size affects the estimation. According to eq. (7) the variance of the estimator is inversely proportional to the number of points used. So, as far as ergodicity is maintained within the window, the estimation should be better for larger window sizes. However, the influence of the window size is not paramount in this new estimator, as only the maximum of the distribution is considered.
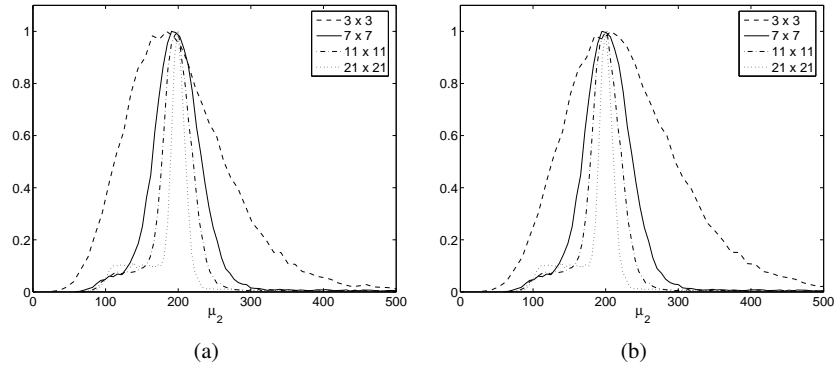
Fig. 5 (Normalized) local second order moment distribution of the noisy image with $\sigma_n = 10$ using 3×3 window (dashed), 7×7 window (solid), 11×11 window (dash-dotted) and 21×21 window (dotted). (a) Normalizing by N (b) Normalizing by $N - 1$, as in eq. (16).

As an illustration, we have calculated the local second order moment distribution of the noisy image using square windows of different sizes. The result is in Fig. 5. The effect of increasing the size of the window is that the function goes narrower, as an effect of the decrease in the variance. However, its maximum is fixed to the same value.

4.2 Estimator based on the local mean

Following a similar reasoning to the one done in the previous section, we can define a local measure based on the estimator of eq. (8). This estimator is also based on the assumption that the distribution in the background of an MR image follows a Rayleigh distribution. Due to the influence of the pixels in this background, when the local mean distribution of a non-noisy MRI is depicted, it presents a maximum around zero (see Fig. 6-(a), solid line). When the image is corrupted with Rician noise, this maximum is now shifted to a value related to σ_n (see the dotted line in Fig. 6-(a)). As the mean of a Rayleigh distribution is defined as $\mu_1 = \sigma_n\sqrt{\pi/2}$, following a similar reasoning as the one done in the previous section, we can define a new estimator

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}} \arg\max\{p(\widehat{\mu_1})\}$$

being $p(\widehat{\mu_1})$ the distribution of the sample local mean. The maximum may be calculated as the mode of the distribution:

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}} \mathrm{mode}(\widehat{\mu_1}_{i,j}) \tag{18}$$

(a)   (b)
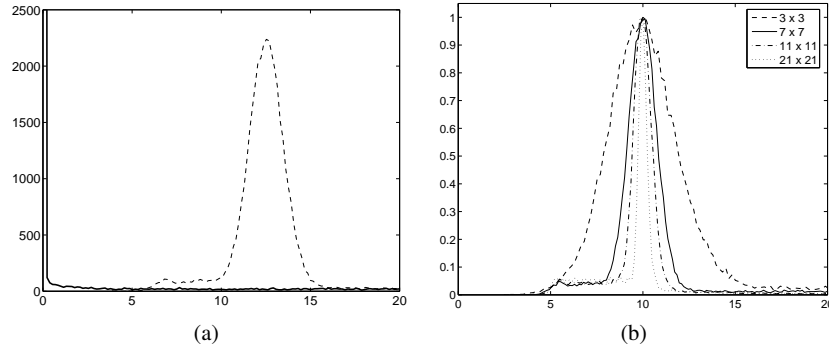
Fig. 6 (a) Sample local mean distribution of the image without noise (solid-bold) and with noise (dashed), with $\sigma_n = 10$. (b) Normalized distribution of $\sqrt{\frac{2}{\pi}}\widehat{\mu_1}_{i,j}$ of the noisy image with $\sigma_n = 10$. Window size: 3×3 (dashed), 7×7 (solid), 11×11 (dash-dotted) and 21×21 (dotted).

To study this assumption, let us analyze the PDF of the sample local mean. Let $R_i(\sigma^2), i = \{1, \cdots, N\}$ be a set of IID random variables with Rayleigh distribution and

$$S = \sum_{i=1}^{N} R_i(\sigma^2)$$

The sum of Rayleigh variables is a classical-hard-to-find problem in communications. Some approximations are usually employed, as the one in [123]

$$p_S(x) = \frac{x^{2N-1}e^{-x^2/2bN}}{2^{N-1}N^N b^N (N-1)!} \tag{19}$$

$$b = \frac{\sigma^2}{N}\left[(2N-1)!!\right]^{1/N} \tag{20}$$

which can be approximated by $b \approx \sigma^2 \frac{2}{e}$ (see appendix 7.1). The mode of the distribution is defined as $\text{mode}(S) = \arg\max_x\{p_S(x)\}$, so

$$\text{mode}(S) = \sigma_n\sqrt{\frac{2(2N-1)N}{e}}$$

which can be approximated when $N >> 1$ by $\text{mode}(S) \approx \sigma_n \frac{2N}{\sqrt{e}}$. Therefore, as the sample mean is defined $\widehat{\mu_1} = \frac{1}{N}S$ the noise estimator becomes

$$\hat{\sigma}_n \approx \frac{\sqrt{e}}{2}\text{mode}(\widehat{\mu_1}_{i,j}) \tag{21}$$

Note than in this approximation, the factor is $\frac{\sqrt{e}}{2} = 0.8244$, and in eq. (18) the factor is $\sqrt{\frac{2}{\pi}} = 0.7979$. This small difference is due to both the approximation in the PDF and to the simplifications made to obtain the mode. In what follows we will use the estimator in eq. (18), as it experimentally shows a better behavior.

To test this estimator a new experiment has been carried out for 40 different standard deviations of noise. The MR image has been normalized between 0 and 1. Local moments are calculated using 7×7 windows. For each σ_n we would take the average value of 100 simulations. In Fig. 7-(a) the estimated standard deviation is depicted vs. the original one. In Fig. 7-(b), as a quality measure, we show the ratio $\widehat{\sigma_n}/\sigma_n$.

As in the previous case, the effect of changing the size for the window used for estimation —see Fig. 6-(b)— is a change in the width of the distribution, but in any case, the maximum stays on the same value.

(a)   (b)
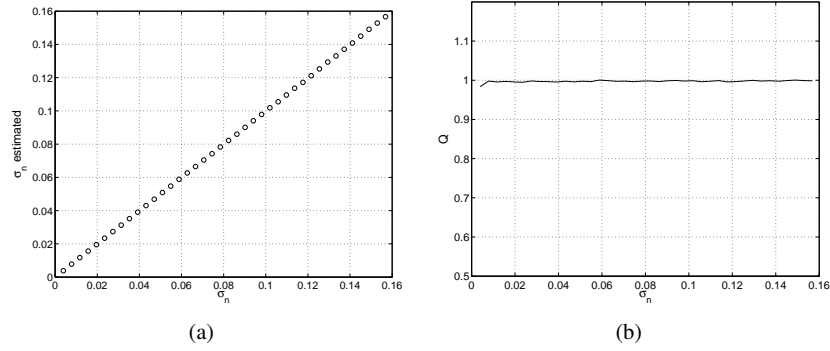
Fig. 7 Estimation of standard deviation of noise using the mode of $\widehat{Q_{E\,i,j}}$. (a) $\widehat{\sigma_n}$ vs. σ_n. (b) $Q = \widehat{\sigma_n}/\sigma_n$ vs. σ_n.

4.3 Estimation based on the local variance

The estimators previously introduced require a Rayleigh distributed area, i.e. an area in which $A_{ij} = 0$. In this section we introduce a new estimator that can be used in signals with general Rician noise. However, we will first study the case of images with zero background.

The distribution of the local variance for most non-noisy images, is a decreasing function with its maximum in zero (see appendix 7.2 for details). When corrupted with noise, the distributions present a *bell-shape distribution* whose maximum is located in a point that should be related to σ_n^2. An illustration is depicted in Fig. 8. We will take advantage of this feature for both estimators.

(a)   (b)
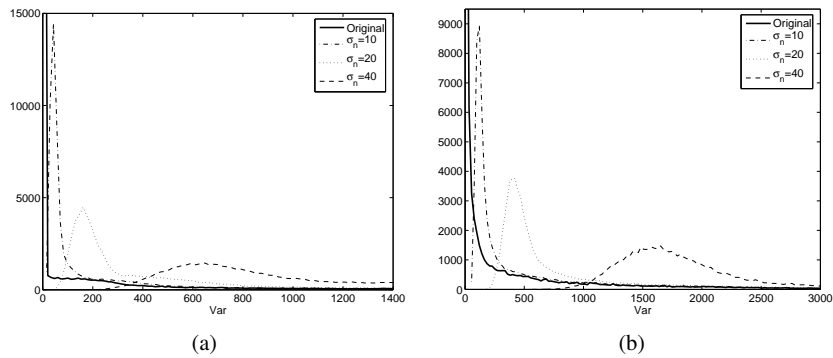
Fig. 8 Sample local variance distribution. (a) MRI with Rician noise. (b) Lenna Image with Rician noise. Original image (solid), Rician noise with $\sigma_n = 10$ (dash-dotted), with $\sigma_n = 20$ (dotted) and with $\sigma_n = 40$ (dashed).

4.3.1 Local variance in MR data

In the background of MR data the variance reduces to the variance of a Rayleigh distribution, which is defined

$$\sigma_M^2 = \sigma_n^2 \left(2 - \frac{\pi}{2} \right) \tag{22}$$

In the pixels belonging to the background then

$$\sigma_n^2 = \sigma_M^2 \left(2 - \frac{\pi}{2} \right)^{-1} \tag{23}$$

If the background of the image is properly segmented, it can be used to estimate the variance of noise using eq. (23). However, the segmentation is not necessary if we follow a philosophy as the one stated in the previous sections. If the variance is estimated locally, say $\sigma_{M_{ij}}^2$, we can define the variance of noise as

$$\widehat{\sigma}_n^2 = \left(2 - \frac{\pi}{2} \right)^{-1} \text{mode} \left(\sigma_{M_{ij}}^2 \right) \tag{24}$$

Actually, the maximum values of the distributions in Fig.8-(b) are located at in $\left(2 - \frac{\pi}{2} \right) \sigma_n^2$.

   To give a theoretical justification to this solution it would be interesting to study the distribution of the sample variance when the samples are Rayleigh random variables (if they are Gaussian, the result is well known [122]). The distribution of the sample local variance for Rayleigh random variables is (see appendix 7.3 for details)

$$p_V(x) = C_V \, e^{-\frac{xN}{2\sigma^2}} \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!} \left[\left(1 + \frac{e}{2} \right) \frac{xN}{2\sigma^2} \right]^k \tag{25}$$

being

$$C_V = \frac{N}{2\sigma^2} \frac{1}{(N-1)!} \frac{(2/e)^{N-1}}{(1+2/e)^{2N-1}}$$

In order to find the maximum independently of the value of σ, we make the change $t = \frac{xN}{2\sigma^2}$

$$f_V(t) = C_V \, e^{-t} \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!} \left[\left(1 + \frac{e}{2} \right) t \right]^k \tag{26}$$

 We have numerically calculated the maximum of this function which is

$$t_{\max} \approx 0.26 \times (N-1) \approx \left(1 - \frac{2}{e} \right) (N-1)$$

In Fig. 9 the function in eq. (26) has been depicted for $N = 25$ (dash-dotted), $N = 49$ (dashed) and $N = 81$ (solid). To avoid the dependence with N, the x-axis has been normalized by $N-1$. The maxima of the functions for the different values
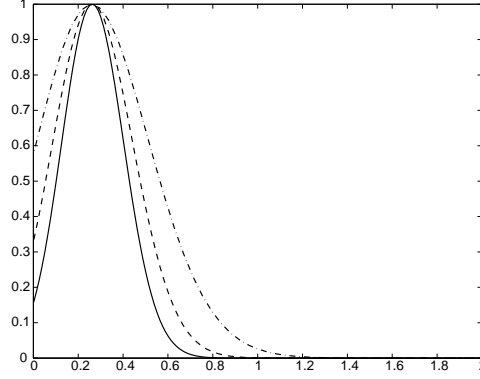
Fig. 9 (Normalized) function of eq. (26) for $N = 25$ (dash-dotted), $N = 49$ (dashed) and $N = 81$ (solid). The x-axis has been normalized by $N - 1$.

of N are all located on the same point, around 0.26. So, the mode of the distribution, after the variable change, will be

$$\text{mode}(V) = x_{\max} \approx 0.52 \frac{N-1}{N} \sigma^2 \approx \left(2 - \frac{4}{e}\right) \frac{N-1}{N} \sigma^2$$

Note that the N in the denominator is due to the fact that we have used the biased variance. Similarly as in the case of the sample mean, due to the approximation, the solution is not exactly what it should be, as $\left(2 - \frac{\pi}{2}\right) \approx 0.42$. Previously, in eq. (21), we have seen that $\sqrt{e}/2$ was in fact an approximation of $\sqrt{2/\pi}$, due to some approximation in the analysis. Note that the parameter we are using for the variance analysis is now $e/2 = 2(\sqrt{e}/2)^2$. If we use $2(\sqrt{2/\pi})^2 = 4/\pi$ instead, the numerical analysis sets the maximum in

$$t_{\max} \approx 0.21(N-1) \approx \left(1 - \frac{\pi}{4}\right)(N-1)$$

being the mode

$$\text{mode}(V) \approx 0.42 \frac{N-1}{N} \sigma^2 \approx \left(2 - \frac{\pi}{2}\right) \frac{N-1}{N} \sigma^2$$

Due to experimental results, we define the estimator as in eq. (24).

To test this estimator a new experiment has been carried out for 40 different standard deviation of noise. The MR image has been normalized between 0 and 1. Local moments are calculated using 7×7 windows. For each σ_n we would take the average value of 100 simulations. In Fig. 10-(a) the estimated standard deviation is depicted versus the original one. In Fig. 10-(b), as a quality measure, we show the ratio $\widehat{\sigma_n}/\sigma_n$.

(a)   (b)
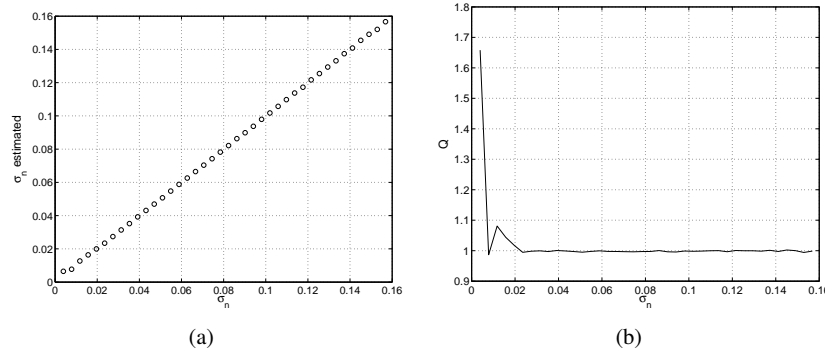
Fig. 10 Estimation of standard deviation of noise using the mode of the sample variance. (a) $\widehat{\sigma_n}$ vs. σ_n. (b) $Q = \widehat{\sigma_n}/\sigma_n$ vs. σ_n.

4.3.2 Local variance in Rician images

Let us now focus on those images with Rician noise in which the assumption of a low SNR background does not hold. As a result, the variance of noise cannot be estimated using the second order moment, nor the local mean, but it can be done using the variance.

Although most of the images we usually deal with are images that do not have a uniform background, they do have a great amount of uniform areas. By *uniform area* we mean areas of the image without borders and with soft transitions. This means that, if no texture is present, the distribution of the local variances will have its maximum value in the vicinity of zero. To illustrate this statement, see for instance the sample local distributions of Lenna and the MR (non-noisy) images in Fig 8 (solid line), where it can be seen that most values are nearly zero, as expected. This exponentially decreasing histogram may be observed in a wide range of images, from natural scenes to medical images (see appendix 7.2). An exception to this is obviously the case of pictures of complex textures, for which intensity variability is considered signal content. We must remark that this assumption does not mean that the estimation is to be done in one specific area assumed constant; the estimation will be done using the whole variance histogram and it will not require any previous selection of any area.

With this assumption in mind, we will first approximate the variance of a Rician model, to afterwards study the sample variance. The mean of a Rician distribution is given by

$$\mu_1 = E\{M\} = \sigma_n \sqrt{\frac{\pi}{2}} e^{-\frac{A^2}{4\sigma_n^2}} \left[\left(1 + \frac{A^2}{2\sigma_n^2} \right) I_0 \left(\frac{A^2}{4\sigma_n^2} \right) + \frac{A^2}{2\sigma_n^2} I_1 \left(\frac{A^2}{4\sigma_n^2} \right) \right] \quad (27)$$

In order to obtain an approximation of the behavior of the variance of the image, we will consider the asymptotic expansion of the Bessel functions [124]:

$$I_n(z) \approx \frac{e^z}{\sqrt{2\pi z}}\left(1 - \frac{4n^2 - 1}{8z}\right), \; z \to \infty \tag{28}$$

It implies that $A_{ij}^2 \gg \sigma_n^2$. After some algebra

$$\sigma_M^2 \approx \sigma_n^2 \left(1 - \frac{5}{4}\frac{\sigma_n^4}{A^2} - \cdots\right) \tag{29}$$

A natural image, like Lenna, has a distribution of $E\{A_{ij}^2\}$ as the one in Fig. 8-(a). This means that if we take local statistics $E\{A_{ij}^2\} > \sigma_n^2$ for the great majority of the pixels, and even $E\{A_{ij}^2\} \gg \sigma_n^4$. So, in those points we can say that

$$\sigma_{M_{ij}}^2 \approx \sigma_n^2 \tag{30}$$

and accordingly, if the variance of the signal is estimated locally, the variance of noise can be estimated as

$$\widehat{\sigma}_n^2 = \text{mode}(\sigma_{M_{ij}}^2) \tag{31}$$

The maximum values of the distributions in Fig.8-(a) are located at σ_n^2.

What we have done here is similar to consider the rough approximation of the noise being Gaussian when the SNR is high. This way, the sample variance has a gamma distribution [122] with mode

$$\text{mode}(\sigma_M^2) = \sigma_n^2 \frac{N-3}{N-1}$$

In Fig. 11 the local variance distribution of a noisy image has been calculated using square windows of different sizes: 5×5 (dashed), 9×9 (solid) and 21×21 (dash-dotted) applying the $\frac{N-3}{N-1}$ correction.
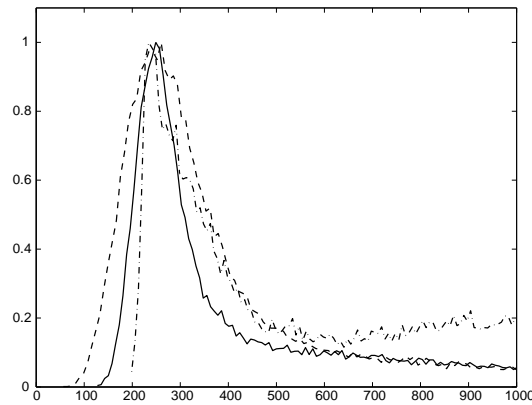
Fig. 11 Normalized local variance distribution of the Lenna image with Rician Noise ($\sigma_n = 15$) using 5×5 (dashed), a 9×9 (solid) and 21×21 (dash-dotted) windows. $\frac{N-3}{N-1}$ correction has been applied.

To completely justify this result, we need to calculate the distribution of the sample variance for the Rician model. This is not an easy task, though some approximations have been done in Appendices 7.4 and 7.5, where we show that the sample variance PDF for this model may be written as

$$p_V(t) \approx C_T e^{-t} \sum_{m=0}^{\frac{2N-1}{4}} \binom{\frac{2N-1}{4}}{m} t^m (1+N)^m \Gamma(2N-1/2-m)$$

with $t = \frac{xN}{2\sigma_n^2}$. Numerically, it is easy to proof that $t_{\max} = \frac{N-3}{2}$ So the mode of the variance distribution will be

$$\mathrm{mode}(V) = \frac{N-3}{2} \frac{2\sigma_n^2}{N} = \frac{N-3}{N} \sigma_n^2$$

Note that if the unbiased sample estimator is used, $\mathrm{mode}(V) = \frac{N-3}{N-1}\sigma_n^2$

To verify this hypothesis an experiment has been carried out for 40 different variances of noise. The Lenna image has been normalized between 0 and 1. Local variance is calculated using 7×7 windows. For each σ_n we would take the average value of 100 simulations. In Fig. 12-(a) the estimated variance is depicted versus the original one. In Fig. 12-(b), as a quality measure, we show the ratio $\hat{\sigma}_n/\sigma_n$.

(a)   (b)
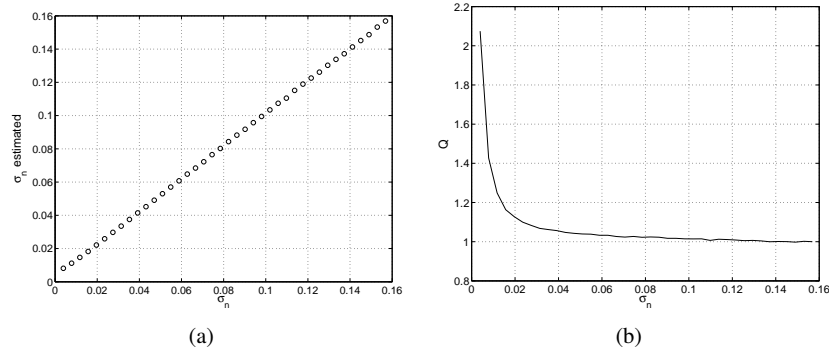
Fig. 12 Estimation of variance of noise using the mode of the local sample variance in the Lenna image. (a) $\widehat{\sigma_n}$ vs. σ_n. (b) $\widehat{\sigma_n}/\sigma_n$ vs. σ_n

4.4 Connection between the two variance models

Two models have been presented for signal estimation based on the mode of the local sample variance. The first one is based on the assumption that the image has
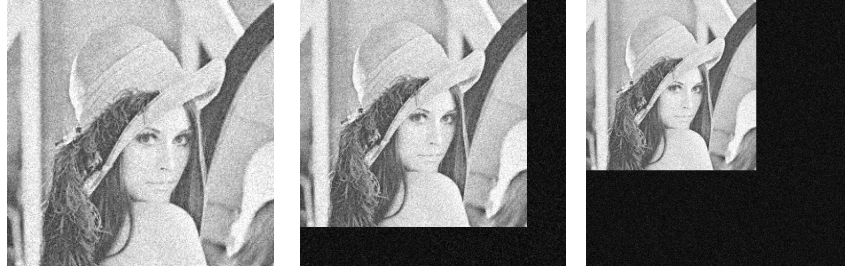
Fig. 13 Noise image with $\sigma_n = 15$ and different background size added.

Fig. 14 Normalized local variance distribution of the lenna image with Rician noise. Without extra background (solid) with background (dashed) and with a greater background (dash-dotted).

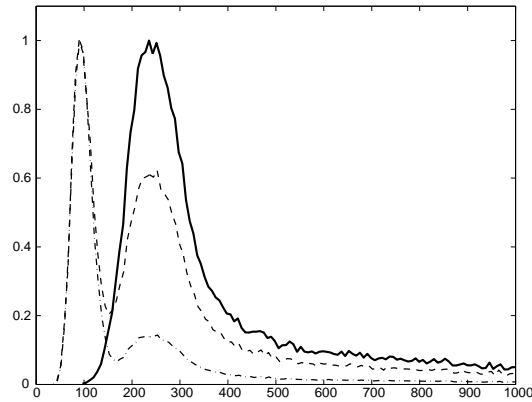a background where the Rician model tends to be Rayleigh; the second one is for images without such a background.

To study the behavior of the local variance in relation with the background, an experiment has been done. A black background has been added to the Lenna image, having a MR-like image. Rician noise with $\sigma_n = 15$ has been added (see Fig. 13). The distribution of the local variance is shown in Fig, 14. We can see that there is a maximum set around σ_n^2, as the second model stated, but in addition, a new maximum appears when the background is added. This new maximum will be set on $\sigma_n^2 \left(2 - \frac{\pi}{2}\right)$, the value predicted by the first model. So, depending on the background, the local variance distribution will be a multimodal distribution, being the local maxima placed on $\sigma_n^2 \left(2 - \frac{\pi}{2}\right)$ and σ_n^2 respectively. The global maxima will be in the mode with the highest coverage.

5 Experimental results

5.1 Synthetic experiments

Some experiments are carried out to illustrate the LMMSE filtering behavior. To be able to compare the results to a *ground truth*, we work with synthetic images artificially corrupted with noise. A magnitude MR image originally noise-free (from the BrainWeb database [53]) with 256 gray levels, is corrupted with Rician noise, see Fig. 15-(a) and 15-(b), following the model in eq. (1). The noisy image is processed using different techniques. Many methods to restore a noisy MR image have been reported. For our experiments we will select those related somehow with the method proposed in this paper. First, those stochastic schemes based on the Rician Model:

1. The *Conventional Approach* (CA) by McGibney et al. [4].
2. The *Maximum Likelihood Estimator* (ML) [80, 7, 82, 114].
3. The *Expectation-maximization Method* (EM) [98, 89].
4. The *Analytically Exact Solution*, proposed by Koay and Basser in [48].

In all the cases where the variance of noise is needed, it is manually set to its optimal value. Note that the ML and EM estimators, as well as the method by Koay and Basser are designed to work with several samples of the same image. As in the present experiment we suppose only one sample is available, the statistics are computed using local neighborhoods. In all cases 5×5 neighborhoods have been used. To compare the filters with other techniques:

5. *Adaptive Wiener Filtering* [125], using a 5×5 neighborhood. In order to achieve the best performance of the filter, σ_n is manually set to the actual value.
6. The *wavelet domain noise filter* for medical imaging proposed by Pižurica *et al.* in [115]. The best results for this experiment are achieved using $K = 5$ and 5×5 window size.

And finally, the LMMSE based schemes:

7. The *LMMSE Estimator*, as proposed in eq. (13), with σ_n manually set to the actual value. 5×5 neighborhood.
8. The *LMMSE Estimator*, as proposed in eq. (13), with σ_n automatically estimated using eq. (18). 5×5 neighborhood both for filtering and noise estimation.
9. The *LMMSE Estimator*, as proposed in eq. (13), with overestimation of noise, manually setting the standard deviation of noise to $\sigma_n + 20$.
10. Although once the image is filtered with the LMMSE estimator the output model is no longer Rician, or at least nothing assures it is, we can think in making the filter recursive. As the noise is dynamically estimated in each iteration, the filter should reach a steady state as the estimated noise gets smaller and smaller. As a result, if a proper noise estimation is done, the filter should stop modifying the input image once the noise is eliminated. We define then a recursive LMMSE (RLMMSE). For the experiments 8 and 50 iterations are considered, and a 5×5 neighborhood has been used both for filtering and noise estimation.

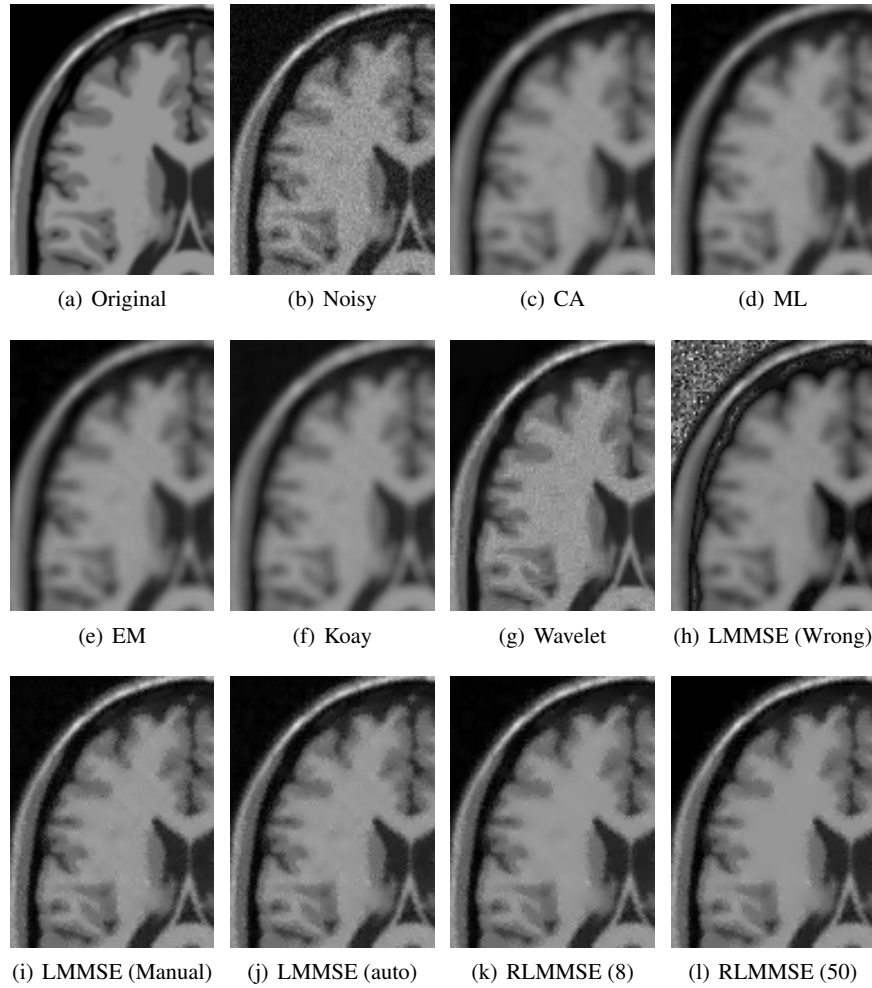|	$\sigma_n = 10$			$\sigma_n = 20$		
	SSIM	QILV	MSE	SSIM	QILV	MSE
Noisy	0.7904	0.9890	100.2940	0.5722	0.9251	395.0881
CA	0.8491	0.6430	192.1922	0.8236	0.6543	205.6214
EM	0.8685	0.6513	144.1177	0.8373	0.6342	168.0615
Koay	0.8424	0.6026	212.4360	0.8024	0.5259	290.9409
ML	0.8681	0.6516	144.2500	0.8370	0.6354	168.1712
Wiener	0.9092	0.9839	57.9197	0.8146	0.9076	161.8120
Wavelet	0.8723	0.9889	51.8368	0.7695	0.9441	132.4157
LMMSE (Overestimation)	0.8107	0.8206	229.0185	0.7465	0.7077	376.9618
LMMSE (Manual noise)	0.9168	**0.9921**	53.9731	0.8346	**0.9613**	130.5361
LMMSE (Automatic noise)	0.9177	**0.9921**	53.6904	0.8389	0.9606	128.1376
RLMMSE (8 steps)	0.9270	0.9917	**51.8197**	**0.8597**	0.9502	**122.5699**
RLMMSE (50 steps)	**0.9298**	0.9915	51.8487	0.8540	0.9429	129.5132

Table 1 Quality measures for the synthetic experiment: SSIM, QILV and MSE for the MR image with Rician noise. Two different σ_n values have been considered. In bold face the best value of each column. LMMSE-based schemes show better results in terms of noise removal and edge preservation.

To compare the restoration performance of the different methods, two quality indexes are used: the Structural Similarity (SSIM) index [126] and the Quality Index based on Local Variance (QILV) [127]. Both give a measure of the structural similarity between the ground truth and the estimated images. However, the former is more sensitive to the level of noise in the image and the latter to any possible blurring of the edges. This way we are able to assess the noise cleaning and border preserving capability of the different schemes. Both indexes are bounded; the closer to one, the better the image. In addition, the mean square error (MSE) is also calculated. To avoid any bias due to background, these three quality measures are only applied to those areas of the image inside the skull.

Table 1 shows the experimental results of the average of 100 experiments for two different values of σ_n. The best value of each column has been highlighted. Some graphical results for $\sigma_n = 10$ are shown in Fig 15.

When compared with other schemes considering a Rician noise model with proper noise estimation, the LMMSE and the RLMMSE show a better performance in terms of noise cleaning (a larger SSIM) while the edges are preserved (the QILV value gets better). However, when noise estimation is incorrectly done (as in Fig. 15-(h) for the case of noise overestimation), the noise in the background is even amplified while the inner edges are blurred.

The noise cleaning performance of the ML, EM and Koay schemes are good, but, as the QILV index points out, they cause image blurring. Consequently, image information is lost at the border and the image edges. This performance is not due to the schemes themselves, but to the fact that they are originally designed to estimate the signal from multiple samples. When only one image is available, the statistics must be calculated locally. Consequently this local estimation produces some edge smoothing, in some cases similar to the one produced by a Gaussian filter. LMMSE, although it is also based on local statistics does not show this edge-blurring behavior. This is one of the arguments for using LMMSE when only one sample is available.

(a) Original   (b) Noisy   (c) CA   (d) ML

(e) EM   (f) Koay   (g) Wavelet   (h) LMMSE (Wrong)

(i) LMMSE (Manual)   (j) LMMSE (auto)   (k) RLMMSE (8)   (l) RLMMSE (50)

Fig. 15 Experiment with synthetic noise. MRI from Brainweb (Detail). (a) Original Image. (b) Image with Rician noise with $\sigma_n = 10$. (c) Conventional approach. (d) ML estimator. (e) EM method. (f) Koay's method. (g) Wavelet Domain Noise filter. (h) LMMSE estimator with noise overestimation ($\sigma_n = 30$). (i) LMMSE estimator with $\sigma_n = 10$ (manually set). (j) LMMSE estimator automatic noise estimation. (k) Recursive LMMSE (8 iterations). (l) Recursive LMMSE (50 iterations). The LMMSE filters with correct noise estimation show the best performance, as confirmed by the numerical results in Table 1.

It is interesting to study the performance of the Wiener filtering; although it slightly blurs the image, it shows an overall good performance. Its performance is worse than the LMMSE, because the Wiener filter is based on a Gaussian noise model. This mismatch between the Rician model and the Gaussian model is not too

large in structural MRI, but becomes more important in other kind of images, like Diffusion Weighted Images (DWI).

On the other hand, the Wavelet based filter shows also a good behavior, both in edge preservation and noise cleaning, though the quantitative quality indices are a bit lower than those of the proposed schemes. In addition, the filter has a parameter K that has to be manually tuned to achieve the best results.

Finally, the RLMMSE results show a very good performance, when compared with the other schemes. There is a good balance between noise cleaning (SSIM index) and edge and structural information preservation (see QILV values). In addition, the filter shows great numerical stability: after 50 iterations the results are similar to those after 8 iterations, indicating that the filter reaches a steady state.

One main advantage of the LMMSE filter (and to some extent for the RLMMSE filter) is that the solution can be computed in one single step (or a few steps for the RLMMSE filter), making it computationally efficient for large data sets as frequently encountered in DWI. This is in contrast to the EM and ML schemes, as well as to the approach by Koay and Basser, where the solution is found by numerical optimization and thus iteratively.

A further comparison has been done with the Wavelet based filter. The brain image used in [115] has been used as ground truth, see Fig. 16. The image is artificially corrupted with Rician noise with different values of σ_n; from 1 to 30. The noisy image is filtered using the wavelet filter (with $K = 2$), the LMMSE scheme and a RLMMSE with 5 iterations and noise estimation using eq. (18). In every case, a 5×5 window has been used. The average of the quality measures of 100 experiments for each σ_n value is depicted in Fig. 17. Results show a better performance of both LMMSE schemes.
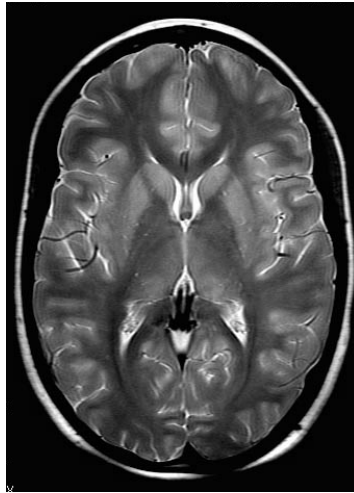

Fig. 16 Original MRI image from [115].

(a) SSIM   (b) QILV   (c) MSE
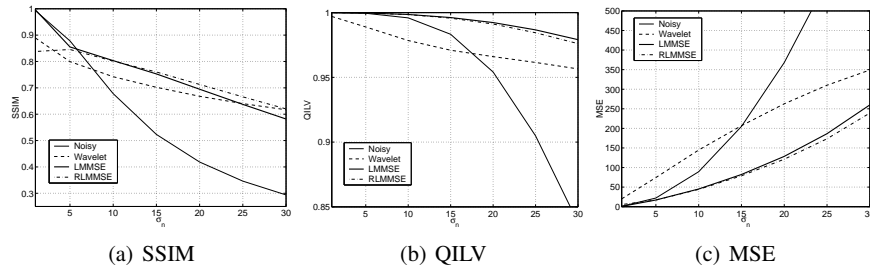
Fig. 17 Comparition between LMMSE and Wavelet-based schemes. Different values of σ_n have been considered.

A next synthetic experiment is done using a 3D ball, where Rician noise has been included (see Fig. 18). The image is filtered using a 3D version of the filter and a $5 \times 5 \times 3$ window. As no zero background is available, the noise estimation is to be done following the method in eq. (31). After 10 iterations, the result is the image in Fig. 18-right.

SSIM and QILV have been used again as quality indices. 19 iterations of the 3D RLMMSE filter have been done, with adaptive noise estimation; results are in Fig. 19. After 4 or 5 iterations the filter reaches the optimum value. If we keep on filtering, the effect will be just a slight blurring of the image.
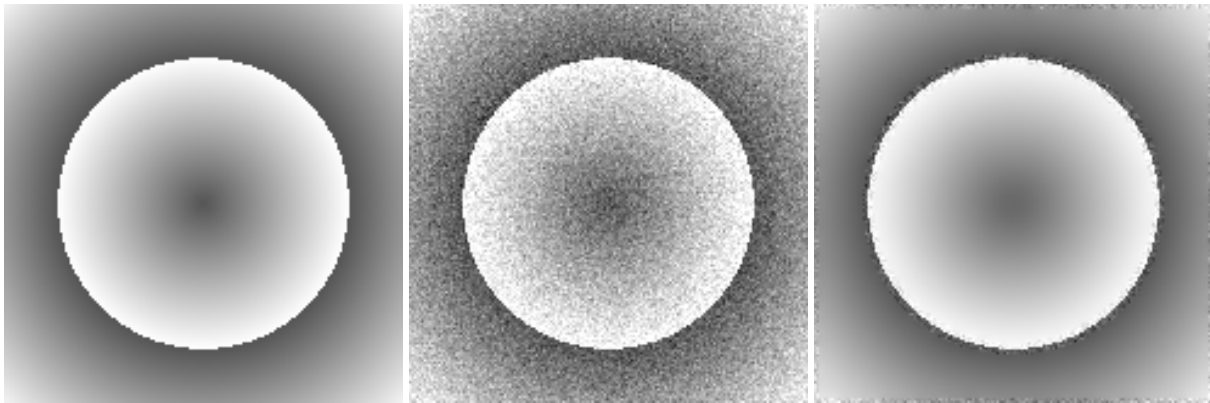
Fig. 18 One slice of a 3D volume. Original image (left), image with Rician noise (center) and filtered after 10 iterations (right).
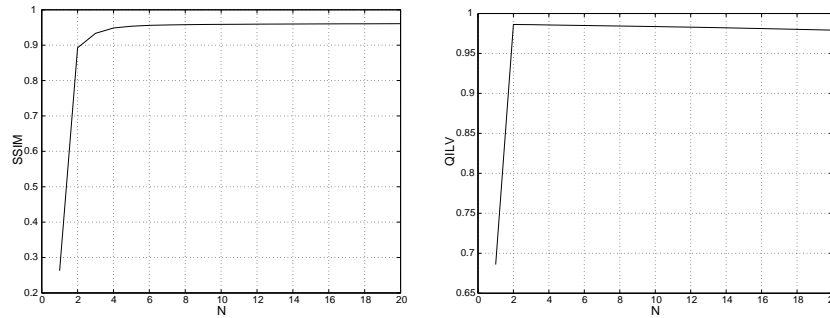
Fig. 19 SSIM index (left) and QILV index (right) of the 3D noisy ball. There is a noise cleaning as well as the structural information is maintained

5.2 Estimation using the background

The methods for noise estimation in MRI described so far have the advantage that they do not need a mask to be segmented in order to estimate the noise. So, they can be easily used in any automatic procedure. This is one of the advantages of these methods over the traditional ones described before.

If the background of the image may be isolated, the estimation can be done only over these background pixels. However, this is not a great advantage. In Fig. 20-(a) the distribution of the local second order moment is depicted, both for the whole noisy image and for the background of the image. Both distributions have almost the same shape, and, what is more important, their maxima are located at the same point. Very similar is the behavior of the local mean as shown in Fig. 20-(b).

(a) (b)
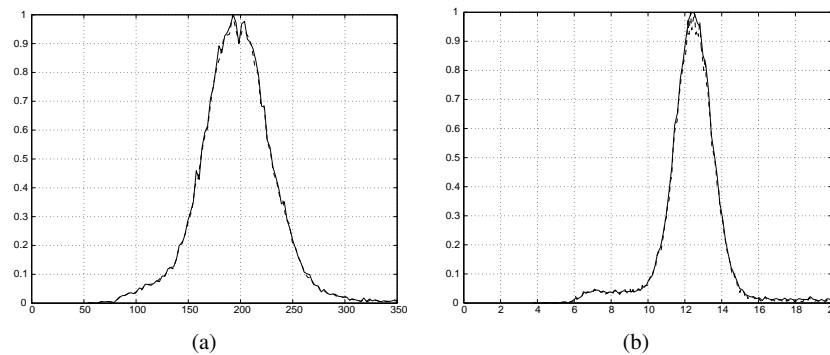
Fig. 20 (a) Distribution of the local second order moment (using 7×7 windows) of the image without noise (solid) and with rice noise ($\sigma_n = 10$): Background pixels isolated (dash-dotted) and whole image (dashed). (b) Distribution of the local mean of the noisy image with $\sigma_n = 10$ using 7×7 windows: Background pixels isolated (dash-dotted) and whole image (dashed).

(a)   (b)
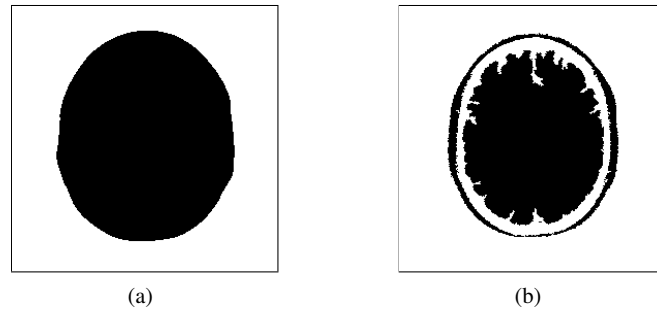
Fig. 21 Background masks used in the experiments.

In order to understand the influence of a proper background segmentation over the noise estimation a new experiment has been done. Local moments are calculated using 7×7 windows. For the noisy image $\sigma_n = 10$, and 50 simulations have been done. First we have used the mask in Fig. 21-(a) to segment the background pixels. This mask totally separates the background from the signal, and it has been manually built form the original image. The second mask is in Fig. 21-(b). It has been built from noisy data using automatic thresholding. Some of the tissue has been assigned to background. This is a common error in automatic segmentation. We will compare four different estimators:

1. The classical unbiased estimator for σ_n^2 of eq. (6).
2. The classical unbiased estimator for σ_n of eq. (8).
3. The new σ_n^2 estimator of eq. (17), based on second order moment distribution.
4. The new σ_n estimator of eq. (18), based on local mean distribution.

Results are in Fig. 22-(a) for the first mask and in Fig. 22-(b) for the second one.

When the background pixels are perfectly identified, the performance of the estimators is quite similar. All the estimations are around $\widehat{\sigma_n} = 10$. So, in this case, the new methods do not present any advantage over the classical ones. But in the second experiment, when some of the pixels of the brain are wrongly assigned to the background, the classical methods do not estimate properly, but the new ones do.

The noise estimation method based on the mode of the histogram proposed in [83] is also robust when a wrong segmentation of the background is done. However, as it is shown in Fig. 23 the estimator has a larger variance than the methods proposed in this paper.

5.3 Filtering MRI

To further verify the hypotheses proposed in the previous sections, and to have a visual idea of the behavior of the filter and noise estimators, a new experiment is

(a)   (b)
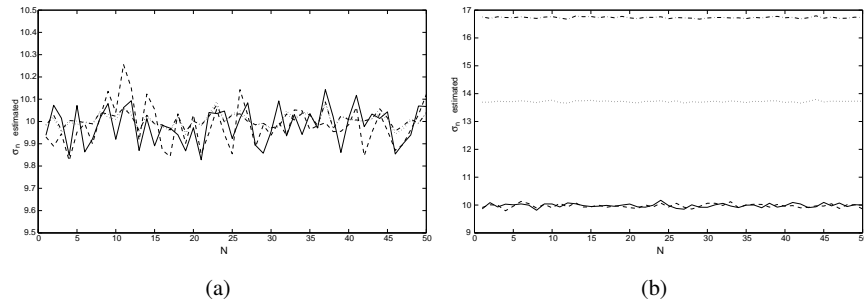
Fig. 22 Estimation of σ_n. (a) First background mask has been used: The new σ_n^2 estimator of eq. (17) based on the local second order moment (solid), new σ_n estimator of eq. (17) based on the local mean (dashed), classical unbiased estimator of eq. (6) for σ_n^2 (dash-dotted) and classical unbiased estimator of eq. (8) for σ_n (dotted). (a) Using first Background mask. (b)Using second background mask.
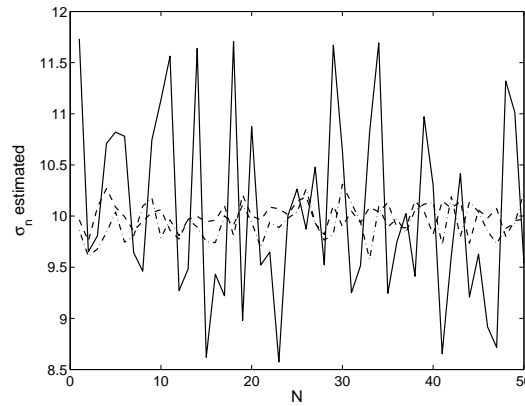
Fig. 23 Estimation of σ_n using the second background mask. Method based on the mode of the histogram (solid), new σ_n^2 estimator of eq. (17) based on the local second order moment (dashed) and new σ_n estimator of eq. (17) based on the local mean (dash-dotted).

carried out based on real data. A coronal slice from a 3D MRI volume[3] has been selected.

This original image, see Fig. 24-(a), exhibits noise. The image is filtered using the LMMSE filter and for the noise estimation the mode of the sample mean is used (eq. (18)). Result is shown in Fig. 24-(f). Some other noise-removal techniques have been used to compare with: EM estimation, the method proposed by Koay and

---

[3] Scanned in a 1.5 Tesla GE Echospeed system. Scanning Sequence: Maximum gradient amplitudes: 40 mT/M. Six images with 4 high (750 s/mm^2), and two with low (5 s/mm^2) diffusion weighting. Rectangular FOV (field of view) 220 x 165 mm. 128 x 96 scan matrix (256 x 192 image matrix). 4 mm slice thickness, 1 mm interslice distance. Receiver bandwidth 6kHz. TE (echo time) 70 ms; TR (repetition time) 80 ms (effective TR 2500 ms). Scan time 60 seconds/slice.

(a) Original   (b) EM   (c) Koay

(d) Wavelet   (e) Wiener   (f) LMMSE

(g) RLMMSE (8 steps)   (h)   RLMMSE   (100
                                              steps)
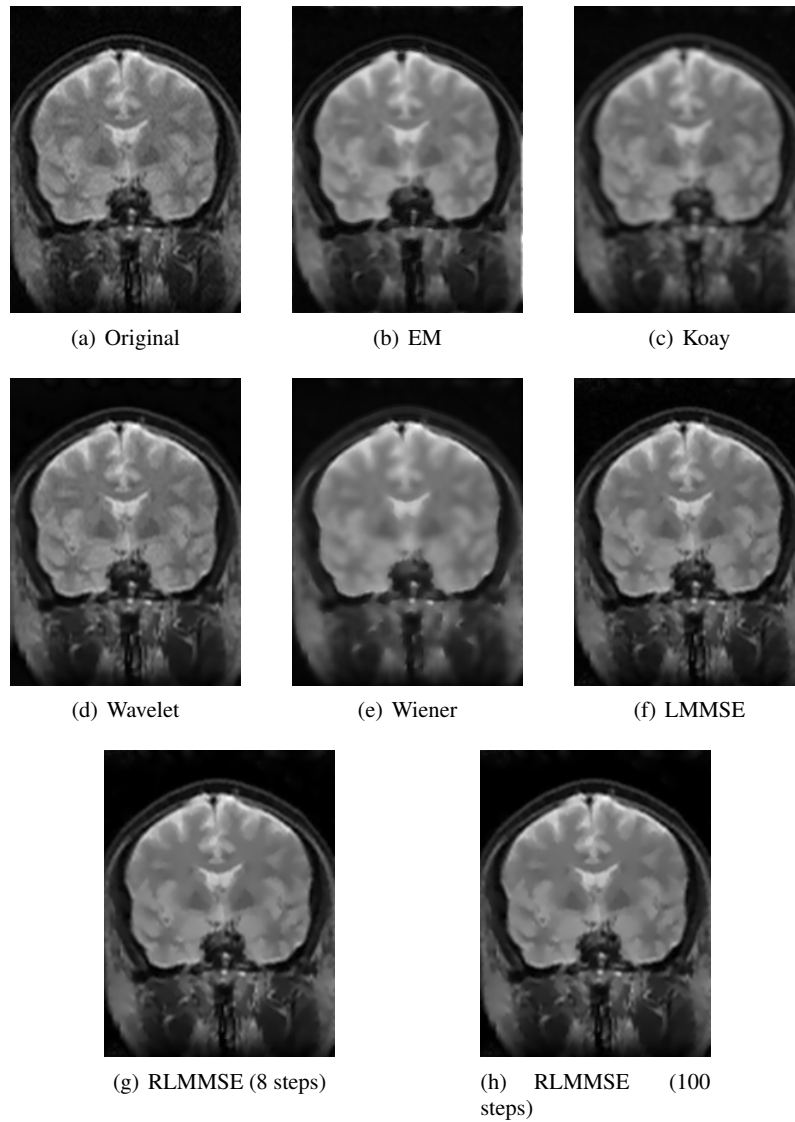
Fig. 24 Coronal slice from a 3D acquisition. (a) Original image. (b) EM method. (c) Koay's method. (d) Wavelet domain filter. (e) 2D Wiener filter. (f) LMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window. (g) RLMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window (8 iterations). (h) RLMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window (100 iterations).

Basser, the wavelet domain filter and the Wiener filter, as the latter showed a good performance in the former experiments. In all the cases adaptive noise estimation

Use	Estimator	Local Statistics	Background	Eq.
MRI	$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}(\widehat{\mu_2}_{i,j})$	$\widehat{\mu_2}_{i,j} = \frac{1}{\|\eta_{i,j}\|-1} \sum\limits_{p\in\eta_{i,j}} I_p^2$	Needed	(17)
MRI	$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\text{mode}(\widehat{\mu_1}_{i,j})$	$\widehat{\mu_1}_{i,j} = \frac{1}{\|\eta_{i,j}\|} \sum\limits_{p\in\eta_{i,j}} I_p$	Needed	(18)
MRI	$\widehat{\sigma_n^2} = \frac{2}{4-\pi}\text{mode}(\widehat{\sigma_I^2}_{i,j})$	$\widehat{\sigma_I^2}_{i,j} = \frac{1}{\|\eta_{i,j}\|-1} \sum\limits_{p\in\eta_{i,j}} \left(I_p - \widehat{\mu_1}_{i,j}\right)^2$	Needed	(24)
Any image	$\widehat{\sigma_n^2} = \text{mode}(\widehat{\sigma_I^2}_{i,j})$	$\widehat{\sigma_I^2}_{i,j} = \frac{1}{\|\eta_{i,j}\|-1} \sum\limits_{p\in\eta_{i,j}} \left(I_p - \widehat{\mu_1}_{i,j}\right)^2$	Not needed	(31)

Table 2 The different estimators of σ_n^2 and σ_n for Rician Noise proposed in the paper.

is performed using eq. (18) with a 5×5 window. Results are given in Fig. 24. We have also considered once more the possibility of making the filter recursive. Results in Fig. 24-(g) (8 iterations) shows the good behavior of this technique. Even after 100 iterations, Fig. 24-(h), due to a good noise estimation, the filter is not blurring the edges. Again, as expected, the visual results are much better for the LMMSE-based schemes, as well as the wavelet based filter, though the latter still presents a slightly noisy pattern inside the tissues. The behavior shown for the other schemes is consistent with the synthetic experiments: noise is attenuated at the cost of blurring the image.

6 Conclusions

A new filtering method based on the LMMSE estimator for Rician distributed images has been introduced, together with several noise estimation methods for that model. The filtering method has proved to be suitable for restoration in this kind of images, as it keeps the structure of the original image unaltered while suppresses most of the noise. Its performance is directly related with the goodness of the noise estimation method employed. If a good dynamic estimator is chosen, the filter may also be used recursively, showing a very good performance in noise cleaning. In addition, unlike other existing schemes also based on Rician models, the fact that a closed-form expression for the LMMSE method has been derived makes the filtering process computationally far more efficient and easier to implement.

We have also presented four different noise estimation methods, all of them based on the mode of some local statistic. A survey of these methods is on table 2. The use of the mode of the sample distribution of some local statistics (second order moment, mean and variance) makes the estimator less dependent of parameters such as the size of the estimation window, the uniformity of the background, and outliers.

The experiments done with synthetic and real images show that the combination of the LMMSE filter with the noise estimation techniques here proposed may be a very useful tool for future MRI restoration.

7 Apendices

7.1 About eq. (20)

Parameter b is defined $b = \frac{\sigma^2}{N} \left[(2N-1)!!\right]^{1/N}$, where $n!! = n(n-2)!!$, for $n \geq 2$. We can write [124]

$$\left[(2N-1)!!\right]^{1/N} = \left[\frac{2^N \Gamma\left(N+\frac{1}{2}\right)}{\sqrt{\pi}}\right]^{1/N} = \left[\frac{2^{1-N} \Gamma(2N)}{\Gamma(N)}\right]^{1/N}$$

and using the Stirling's approximation [124]

$$\Gamma(z) \approx e^{-z} z^{z-1/2} (2\pi)^{1/2} \left[1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots\right]$$

for $z \to \infty$ and $|\arg z| < \pi$ we obtain

$$\left[(2N-1)!!\right]^{1/N} \approx \left[2^{1-N} \frac{e^{-2N} (2N)^{2N-1/2} (2\pi)^{1/2} \left[1 + \frac{1}{24N} + \cdots\right]}{e^{-N} (N)^{N-1/2} (2\pi)^{1/2} \left[1 + \frac{1}{12N} + \cdots\right]}\right]^{1/N}$$

And as $N >>$ we can approximate $\left[(2N-1)!!\right]^{1/N} \approx \frac{2N}{e}$ and $b \approx \sigma^2 \frac{2}{e}$. In Fig. 25 the function $\left[(2N-1)!!\right]^{1/N}$ is depicted together with its approximation $2N/e$ for $N \in [1, 100]$.

(a)   (b)
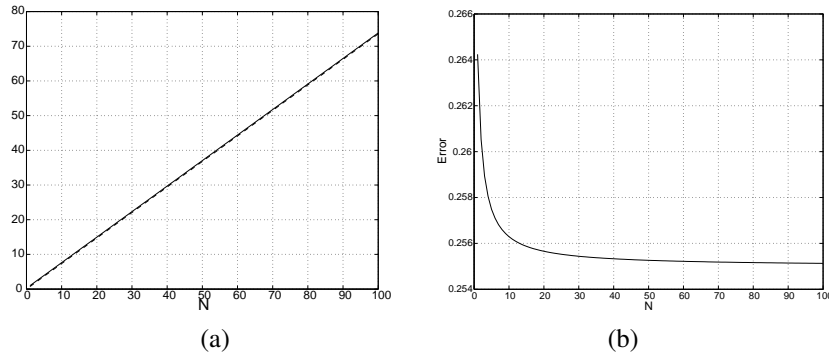
Fig. 25 (a) Function $\left[(2N-1)!!\right]^{1/N}$ (solid) and its approximation with function $2N/e$ (dashed). (b) Error=$\left[(2N-1)!!\right]^{1/N} - 2N/e$

7.2 *Local variance distribution in non-textured images*

Fig. 26 Images to be analyzed. MRI from Brainweb [53]. Radiography from Hospital Río Carrión, Palencia (Spain). Aerial picture from GOES Project Science (goes.gsfc.nasa.gov).

One of the noise estimators proposed is based on the assumption that the distribution of the local variances in a non-noisy image follows a distribution whose maximum is set on zero. To evaluate this assumption, the histogram of local variances of different kind of images has been evaluated. The images are shown in Fig. 26. Medical, meteorological and natural images have been used. The sampling local variance is calculated using 5×5 windows. The results are shown in Fig. 27. Finally, the mean of the distribution of 29 images from LIVE database [128] is presented in Fig. 28. Two cases have been considered: firstly, gray scale images and secondly RGB images where each color component have been evaluated separately.

7.3 *PDF of the Sample Local Variance for Rayleigh Random Variables*

Let $R_i(\sigma^2), i = \{1, \cdots, N\}$ be a set of random variables with Rayleigh distribution. The (biased) sample variance is defined as

$$\widehat{\mathrm{Var}}(R_i) = \frac{1}{N} \sum_{i=1}^{N} (R_i)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} R_i \right)^2$$
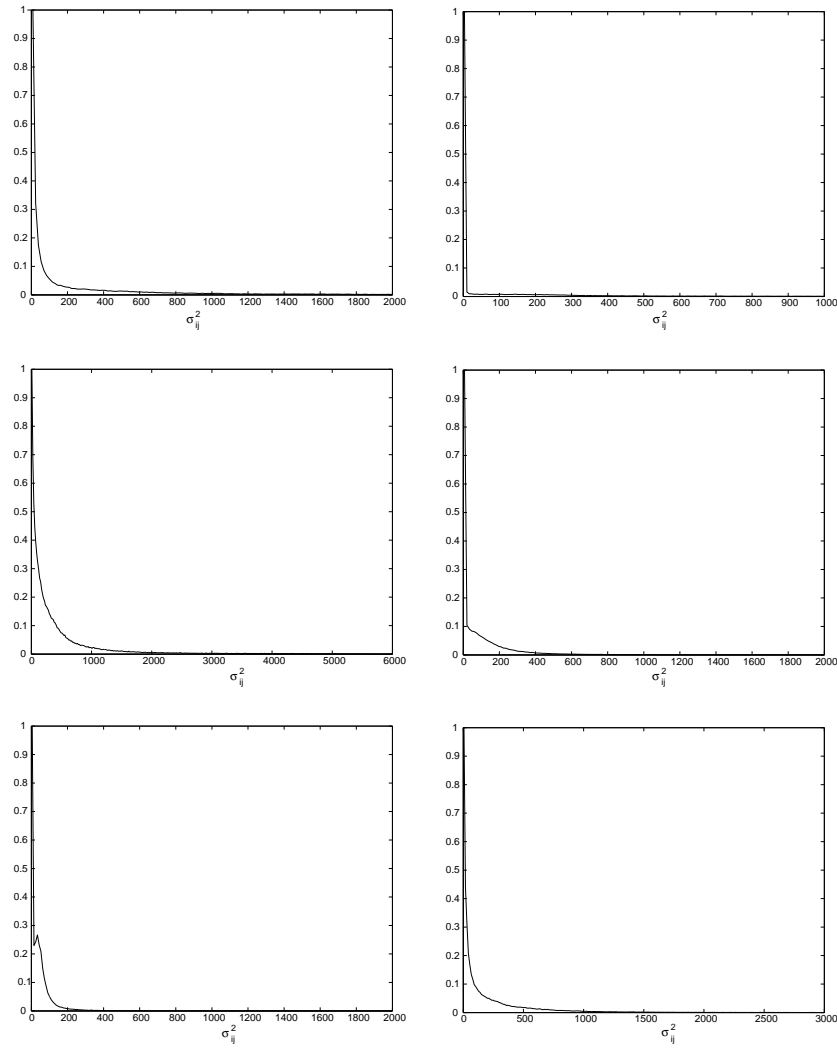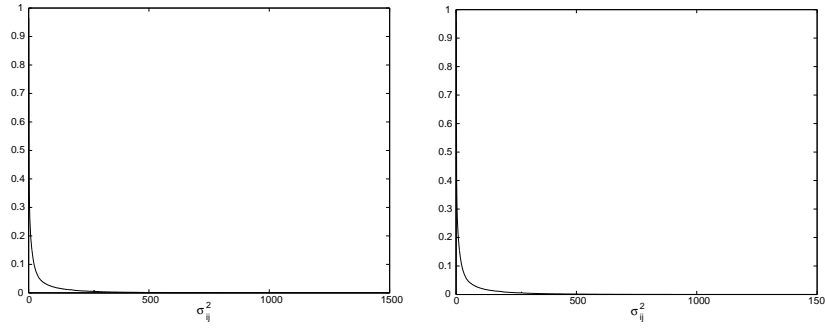
Fig. 27 Normalized distribution of local variances of the images in Fig. 26, displayed in the same order. A 5×5 window has been used.

We can define

$$S_1 = \frac{1}{N} \sum_{i=1}^{N} R_i^2(\sigma^2) \sim \chi_{2N}\left(\frac{\sigma^2}{N}\right)$$

being χ_{2N} a Chi-Square distribution, and

Fig. 28 Mean of the normalized distribution of local variances of the images from LIVE database [128]. A 5×5 window has been used. Gray scale images (left) and RGB images (right). Each color component have been evaluated separately.

$$S_2 = \frac{1}{N} \sum_{i=1}^{N} R_i(\sigma^2)$$

The sample variance can be defined as

$$V = \widehat{\mathrm{Var}}(R_i) = S_1 - (S_2)^2$$

The PDF of S_2 can be approximated [123] by

$$p_{S_2}(x) = \frac{x^{2N-1} N^N e^{-x^2 N/2b}}{2^{N-1} b^N \Gamma(N)}$$

According to the fundamental theorem [122] the PDF of $(S_2)^2$ is

$$p_{(S_2)^2}(x) = \frac{x^{N-1} N^N e^{-xN/2b}}{2^N b^N \Gamma(N)}$$

$$(S_2)^2 \sim \chi_{2N}\left(\frac{b}{N}\right)$$

We can approximate (see appendix 7.1) $b \approx \sigma^2 \frac{2}{e}$, so

$$(S_2)^2 \sim \chi_{2N}\left(\frac{\sigma^2}{N} \frac{2}{e}\right)$$

The distribution of the sample variance is then the distribution of the difference of two Chi-Square random variables. According to [45] this distribution, for $x \geq 0$ and the same (even) number of degrees of freedom is

$$p_X(x) = \frac{1}{2\sigma_1^2} e^{-\frac{x}{2\sigma_1^2}} \frac{1}{(N-1)!} \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^N \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!}$$

$$\times \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)^{N-1-k} \left(\frac{x}{2\sigma_1^2}\right)^k$$

Making $\sigma_1^2 = \frac{\sigma^2}{N}, \sigma_2^2 = \frac{\sigma^2}{N}\frac{2}{e}$ we can write

$$p_V(x) = C_V \, e^{-\frac{xN}{2\sigma^2}} \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!} \left[\left(1 + \frac{e}{2}\right) \frac{xN}{2\sigma^2}\right]^k$$

being

$$C_V = \frac{N}{2\sigma^2} \frac{1}{(N-1)!} \frac{(2/e)^{N-1}}{(1 + 2/e)^{2N-1}}$$

7.4 Difference of Non-central Chi-Square Random Variables

Let X_1 and X_2 be two random variables with Non-central Chi-square distributions, such as

$$X_1 \sim \chi^2_{K_1}(x\alpha_1; \lambda_1) \quad X_2 \sim \chi^2_{K_2}(x\alpha_2; \lambda_2)$$

with PDFs

$$p_i(x) = f_i(x\alpha_i; K_i, \lambda_i) = \frac{e^{-\frac{x\alpha_i + \lambda_i}{2}}}{2} \left(\frac{x\alpha_1}{\lambda_i}\right)^{\frac{K_i-2}{4}} I_{\frac{K_i}{2}-2}\left(\sqrt{x\alpha_i\lambda_i}\right) u(x) \quad (32)$$

being $I_n(z)$ the modified Bessel function of the first kind, which can be rewritten using ascending series [124] as

$$I_n(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}z\right)^{n+2k}}{k!\Gamma(n+k+1)}$$

Eq. (32) can be accordingly rewritten as

$$p_i(x) = e^{-\frac{x\alpha_i + \lambda_i}{2}} \sum_{m=0}^{\infty} \left(\frac{1}{2}\right)^{\frac{K_i}{2}+2m} \frac{(x\alpha_1)^{\frac{K_i-2}{2}+2m} \lambda_1^m}{m!\Gamma\left(\frac{K_i}{2}+m\right)} u(x)$$

Let us define the variable V as $V = X_1 - X_2$. For $x \geq 0$ we can write

$$p_V(x) = \int_0^{\infty} p_1(x+y)p_2(y)dy$$

If using Newton's generalized binomial theorem, the final PDF can be written as

$$p_V(x) = e^{-\frac{x\alpha_1 + \lambda_1 + \lambda_2}{2}} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \left(\frac{1}{2}\right)^{n+m+p+1} \lambda_1^m \lambda_2^n \binom{\frac{K_1}{2} - 1 + m}{p}$$

$$\frac{\alpha_1^{K_1/2-1+m} \alpha_2^{K_2/2-1+n}}{(\alpha_1 + \alpha_2)^{\frac{K_1+K_2}{2}-1+m+n-p}} \frac{\Gamma\left(\frac{K_1+K_2}{2} - 1 + m + n - p\right)}{m! n! \Gamma\left(\frac{K_1}{2} + m\right) \Gamma\left(\frac{K_2}{2} + n\right)} x^p \quad (33)$$

7.5 Sample variance in Rician distributed data

Let R_i, $i = \{1, \cdots, N\}$ be a set of random variables with Rician distribution and

$$S_1 = \frac{1}{N} \sum_{i=1}^{N} R_i^2 \quad S_2 = \frac{1}{N} \sum_{i=1}^{N} R_i$$

The sample variance can be defined as $V = S_1 - (S_2)^2$ The PDF of S_1 is [120]

$$p_{S_1}(x) = M_1 x^{\frac{N-1}{2}} e^{-xN/2\sigma^2} I_{N-1}\left(\frac{\sqrt{xN}A}{\sigma^2}\right)$$

with M_1 a constant, and $A^2 = \sum_i A_i^2$. The sum of Rician distribution is, as in the Rayleigh case, a classical problem in communications. We can use the approximation in [120]. From here the PDF of $(S_2)^2$ would be

$$p_{S_2^2}(x) = M_2 x^{\frac{N-1}{2}} e^{-xN/2c_2^2} I_{N-1}\left(\frac{\sqrt{xN}b}{c_1 c_2}\right)$$

with M_2 a constant, and c_1, c_2 and b parameters related with N and the Signal to Noise ratio. Both PDF may be expressed as Non-central Chi-Square distributions:

$$S_1 \sim \chi_{2N}^2\left(\frac{xN}{\sigma^2}, \frac{A^2}{\sigma^2}\right) \quad S_2^2 \sim \chi_{2N}^2\left(\frac{xN}{c_2^2}, \frac{b^2}{c_1^2}\right)$$

The subtraction of two Non-central Chi-Square random variables has been studied in appendix 7.4. Accordingly, the PDF of the sample variance will be like the one in eq. (33). However, this long equation may have some problems when trying to numerically determine its maximum, specially when the argument of the Modified Bessel function is not small.

Experimentally, a good approximation has been obtained, without any dependence on the parameters c_1, c_2, A and b.

$$p(t) \approx C_T e^{-t} \sum_{m=0}^{\frac{2N-1}{4}} \binom{\frac{2N-1}{4}}{m} t^m (1+N)^m \Gamma\left(2N - 1/2 - m\right) \qquad (34)$$

with $t = \frac{xN}{2\sigma_n^2}$.

To show the behavior of the approximation, some data has been generated following a Rician model. A constant image with value 20 has been corrupted with Rician noise with $\sigma_n = 6$. The distribution of some local statistics have been calculated using 3×3, 5×5 and 7×7 square windows. The results are shown in Fig. 29. In solid line the actual data distribution is depicted, and the theoretical distribution is in dashed line.
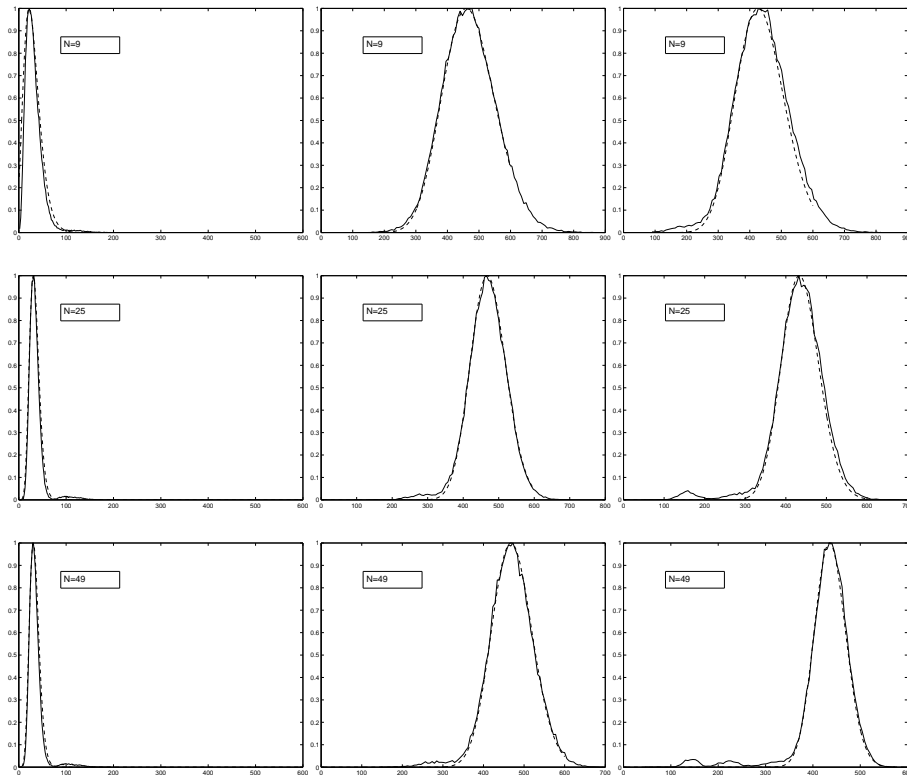
Fig. 29 Rician distributed data. Theoretical PDF vs. real data distribution. Left: Local Sample variance. Center: Local second order moment. Right: Square local mean. Solid line: data distribution. Dashed line: theoretical PDF.

Acknowledgements

Noise estimation in single and multiple coil MR data

Santiago Aja-Fernández, Antonio Tristán-Vega, Carlos Alberola-López*

Abstract Noise estimation is a challenging task in Magnetic Resonance Imaging (MRI), with applications in quality assessment, filtering or diffusion tensor estimation. Main noise estimators based on the Rician model are revisited and classified in this paper, and new useful methods are proposed. Additionally, all the surveyed estimators are extended to the non–central Chi model, which applies to multiple coil MRI and some important Parallel Imaging algorithms for accelerated acquisitions. The proposed new noise estimation procedures, based on the distribution of local moments, show better performance in terms of smaller variance and unbiased estimation over a wide range of experiments, with the additional advantage of not needing to explicitly segment the background of the image.

1 Introduction

Noise in the **k**-space in Magnetic Resonance (MR) data from each coil is assumed to be a zero-mean uncorrelated Gaussian process with equal variance in both the real and imaginary parts. As a result, in single coil systems magnitude data in the spatial domain is modeled using a Rician distribution [44, 5]. In the same way, the composite signal in coils systems with multiple channels may be modeled as non–central Chi distributed [27] if no subsampling of the **k**-space is assumed. The acquisition rate can be increased with parallel MRI (pMRI) techniques via subsampled acquisitions of the **k**-space data. In these cases, reconstruction methods have to be used in order to suppress the aliasing and underlying artifacts created by the subsampling. Dominant among these are SENSE [32] and GRAPPA [33], reviews of which can be found in [30, 31]. From a statistical point of view, such a reconstruction will affect

---

the stationarity of the noise in the reconstructed data, i.e. the spatial distribution of the noise across the image [43]. As a result, the variance of noise may vary for different image locations. Moreover it may also vary from one coil to another. However, under the assumption of a nearly homogeneous variance [46], the data may be considered to follow a general non–central Chi distribution. This distribution reduces to a Rician if SENSE is used.

Noise in MR data, either from multiple or single coil acquisitions, is known to affect the visual quality of the MR images and different processing techniques, such as segmentation, registration or tensor estimation in Diffusion Tensor MRI (DT-MRI) [6]. Accordingly, the estimated noise power gives a measure of the quality of the data. This estimation can be used to measure the Signal-to-Noise Ratio (SNR) and as an input parameter in MR processing algorithms. Many filtering methods to *improve* SNR in MRI need an estimated value for σ_n^2: as the conventional approach [4], maximum likelihood based methods [80, 7, 114], expectation maximization formulations with Rician noise assumptions [98], Linear Minimum Mean Square Error (LMMSE) based schemes [6, 8, 129, 9] and unbiased non-local mean schemes [88, 130, 11, 13]. New techniques for DTI tensor estimation [110, 17], segmentation methods based on the Rician distribution and fiber orientation estimators [55] also depend upon an estimated σ_n^2 value.

The aims of this paper are (1) to review and classify different approaches to estimate noise in Rician magnitude MR images; (2) to propose new methods to estimate noise in magnitude images; (3) to extend the Rician-based estimators to the non-central Chi model; and (4) to propose a method to estimate noise in complex MR images, if they are available. The advantage of this last estimation is that no background assumption is needed. A survey of all the methods may be found in Table 2. The methods are tested and compared using synthetic and real data.

2 Theory

2.1 Statistical noise model in single and multiple coil MR signal

For a single–coil acquisition the complex spatial MR data is typically modeled as a complex Gaussian process

$$C(\mathbf{x}) = A(\mathbf{x}) + n(\sigma_n^2), \tag{1}$$

with $C(\mathbf{x})$ the complex spatial signal, $A(\mathbf{x})$ the original signal if no noise is present and $n(\sigma_n^2) = n_r(\mathbf{x}; \sigma_n^2) + j n_i(\mathbf{x}; \sigma_n^2)$ complex uncorrelated Gaussian noise with zero mean and variance σ_n^2. The magnitude signal $M(\mathbf{x})$ is the Rician distributed envelope of the complex signal [5]

$$M(\mathbf{x}) = \sqrt{|A(\mathbf{x}) + n(\sigma_n^2)|^2}, \tag{2}$$

$E\{X\}$	Expectation of random variable X
σ_X^2	Variance of random variable X
$\langle M(\mathbf{x}) \rangle$	(Global) Sample mean of image $M(\mathbf{x})$
	$\langle M(\mathbf{x}) \rangle = \frac{1}{|\Omega|} \sum\limits_{\mathbf{x} \in \Omega} M(\mathbf{x})$
$\langle M(\mathbf{x}) \rangle_{\mathbf{x}}$	Local sample local mean of image $M(\mathbf{x})$
	$\langle M(\mathbf{x}) \rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|} \sum\limits_{\mathbf{p} \in \eta(\mathbf{x})} M(\mathbf{p})$
	($\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$)
$\text{Var}(M(\mathbf{x}))_{\mathbf{x}}$	Sample local variance of $M(\mathbf{x})$
	$\text{Var}(M(\mathbf{x}))_{\mathbf{x}} = \langle M^2(\mathbf{x}) \rangle_{\mathbf{x}} - \langle M(\mathbf{x}) \rangle_{\mathbf{x}}^2$
$M(\mathbf{x_B})$	Background area of image $M(\mathbf{x})$
	$\mathbf{x_B} = \mathbf{x}|A(\mathbf{x}) = 0$
$M(\mathbf{x_R})$	$M(\mathbf{x})$ in the region $\mathbf{R}$
	$\mathbf{x_R} \in \mathbf{R}$
$\text{mode}\{I(\mathbf{x})\}$	Mode of the distribution of $I(\mathbf{x})$
	$\text{mode}\{I(\mathbf{x})\} = \arg\max_{\mathbf{x}}\{p_I(\mathbf{x})\}$
$\widehat{a}$	Estimator of parameter a

Table 1 Notation

with probability density function (PDF) [44]:

$$p_M(M|A, \sigma_n) = \frac{M}{\sigma_n^2}\, e^{-\frac{M^2+A^2}{2\sigma_n^2}}\, I_0\left(\frac{AM}{\sigma_n^2}\right)\, u(M), \qquad (3)$$

with $I_0(.)$ the 0th order modified Bessel function of the first kind, $u(.)$ the Heaviside step function. In the image background, where the SNR is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution with PDF

$$p_M(M|\sigma_n) = \frac{M}{\sigma_n^2}\, e^{-\frac{M^2}{2\sigma_n^2}}\, u(M). \qquad (4)$$

In a multiple–coil MR acquisition system the acquired signal in each signal may be modeled in the complex spatial domain as the original signal corrupted with complex additive Gaussian noise, with zero mean and equal variance. Thus, the complex signal in coil l (for $l = 1, 2, \ldots, L$) can be expressed as

$$C_l(\mathbf{x}) = A_l(\mathbf{x}) + n_l(\sigma_n^2), \qquad (5)$$

with $n_l(\sigma_n^2) = n_{l_r}(\mathbf{x}; \sigma_n^2) + jn_{l_i}(\mathbf{x}; \sigma_n^2)$ complex Gaussian noise with zero mean and variance σ_n^2 and $A_l(\mathbf{x})$ the original signal in coil l if no noise is present. If no subsampling is done in the $\mathbf{k}$-space, the composite magnitude image may be obtained using methods such as the sum-of-squares (SoS) [27, 38, 48]:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |C_l(\mathbf{x})|^2}, \tag{6}$$

Defining $A_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |A_l(\mathbf{x})|^2}$, and assuming the noise components to be identically and independently distributed, the envelope of the magnitude signal M_L will follow a non–central Chi distribution with PDF [27]:

$$p_{M_L}(M_L|A_L, \sigma_n, L) = \frac{A_L^{1-L}}{\sigma_n^2} M_L^L e^{-\frac{M_L^2 + A_L^2}{2\sigma_n^2}} I_{L-1}\left(\frac{A_L M_L}{\sigma_n^2}\right) u(M_L), \tag{7}$$

which reduces to the Rician distribution [5] for $L = 1$. In the background, this PDF simplifies to a central Chi distribution with PDF:

$$p_{M_L}(M_L|\sigma_n, L) = \frac{2^{1-L}}{\Gamma(L)} \frac{M_L^{2L-1}}{\sigma_n^{2L}} e^{-\frac{M_L^2}{2\sigma_n^2}} u(M_L), \tag{8}$$

which reduces to Rayleigh for $L = 1$.

This statistical model is the usual one for the magnitude signal in phased array coils and parallel imaging assuming that no subsampling is done in the **k**-space, and the image is reconstructed using the SoS method. However, one of the aims of parallel imaging is precisely to accelerate the acquisition process by sub–sampling **k**-space data in each coil. From a statistical point of view, such a reconstruction will affect the stationarity of the noise in the reconstructed data, i.e. the spatial distribution of the noise across the image [43]. Noise distributions in SENSE and GRAPPA are known to vary across pixels. As a result the statistics of the composite magnitude signal are not strictly stationary. In addition, when reconstructed with GRAPPA σ_n^2 may also vary from coil to coil and then, eq. (7) does not exactly hold, since it assumes that σ_n^2 is the same for every coil. In practical situations, if the variance of noise is homogeneous enough across pixels and coils, data are usually considered to follow a non-central Chi distribution if reconstructed with GRAPPA and SoS [46].

2.2 Noise estimation in the Complex Domain

Noise estimation in MR is usually done over the (composite) magnitude image, since it is the usual output of the scanning process. However, if data in the complex spatial domain are available, the estimation may be easier done in that domain. As stated before, reconstructed data in each coil $C_l(\mathbf{x})$, both for single or multiple acquisition, are corrupted with Gaussian noise with zero mean and variance σ_n in the real and imaginary parts. Noise estimation is carried out assuming that the noise is uncorrelated and with identical variance in each pixel and it will be reduced to the estimation of the variance in a well-known Gaussian problem over one of the components. Many solutions have been proposed, such as methods based on

wavelets [131, 132], singular-value-decomposition [133] or fuzzy logic [134]. Alternatively, fast and simple solutions based on blockwise operations have also been reported [125, 135]. In [108] a fast and accurate method is presented (see Table 1 for notation)

$$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(C_{l_j}(\mathbf{x}))_{\mathbf{x}}\}, \tag{9}$$

with $C_{l_j}(\mathbf{x})$ either the real or the imaginary component of the complex data $C_l(\mathbf{x})$ in l^{th} coil. Although the estimation benefits from the existence of a uniform background, it is not necessary to have one. The only requirement for the image is not to be a texture [108], which is the case for MR. So, this estimator will always be valid and accurate in the MR case. Note that although the estimator is defined over the complex spatial domain, it could also be defined over the $\mathbf{k}$-space data, where the Gaussian model also holds.

If no complex data are available, noise estimation is carried out over the composite magnitude image, assuming again that the noise is uncorrelated and with identical variance in each pixel. Methods performing such estimation from magnitude data may roughly be divided into two groups: (i) approaches estimating the noise variance using a single magnitude image and (ii) approaches using multiple images. In this paper we will focus on the former. Noise estimation using a single image is usually based on background intensities, where the true signal amplitude should vanish and the Rayleigh or central Chi assumptions hold. Some of these techniques require a previous background segmentation.

2.3 Noise estimation assuming Rician and Rayleigh distributions

The main noise estimators assuming Rician distributed MRI are (see Table 1 for notation):

1. *Estimators that need prior segmentation of the background area:* based on the mean and the second order moment of Rayleigh data [86, 81]:

$$\widehat{\sigma_n^2} = \frac{1}{2}\langle M^2(\mathbf{x_B})\rangle \tag{10}$$

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\langle M(\mathbf{x_B})\rangle. \tag{11}$$

   where $M(\mathbf{x_B})$ is the (segmented) background of image $M(\mathbf{x})$. Eq. (10) is the Maximum Likelihood estimator for the Rayleigh distribution [7]. Note that both estimators use all –and only– the pixels in the background to perform the estimation. Accordingly, a previous segmentation of such an area is needed.
2. *Estimators based on the background that do not need prior segmentation.* These estimators assume a Rayleigh PDF in the background of the image. The estimation is done without segmentation, by taking the maximum value of some local distribution. Different approaches are proposed:

a. Using the mode of some parameter related to σ_n, as the maximum of the histogram of the image [37, 136, 83],

$$\widehat{\sigma_n} = \text{mode}\{M(\mathbf{x})\}. \tag{12}$$

or sample local moments of the image [8]

$$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\{\langle M^2(\mathbf{x})\rangle_{\mathbf{x}}\} \tag{13}$$

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\text{mode}\{\langle M(\mathbf{x})\rangle_{\mathbf{x}}\} \tag{14}$$

$$\widehat{\sigma_n^2} = \frac{2}{4-\pi}\text{mode}\{\text{Var}(M(\mathbf{x}))_{\mathbf{x}}\} \tag{15}$$

where $\langle M(\mathbf{x})\rangle_{\mathbf{x}}$ is the sample local mean of image $M(\mathbf{x})$, as opposed to $\langle M(\mathbf{x})\rangle$ which is the *global* sample mean. Taking the square root will reduce the dynamic range of the second order moment. From [8] it is easy to derive that the unbiased estimator will be:

$$\widehat{\sigma_n} = \sqrt{\frac{N}{2N-1}}\text{mode}\left\{\sqrt{\langle M^2(\mathbf{x})\rangle_{\mathbf{x}}}\right\} \tag{16}$$

with N the number of points used for the sample moment estimation, i.e $N = |\eta(\mathbf{x})|$. In [8] estimators based on local moments were proven to be more robust than those based on the image histogram, mainly due to the fact that the distributions of this moments are much less spread giving a better mode estimation.

b. *Methods that fit a Rayleigh-related distribution to the histogram of the data.* Brummer *et al.* in [37] uses a least squares fitting of the Rayleigh distribution with the histogram of the data in the background

$$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{l=l_0}^{l_c}\left(h_M(l) - K\frac{l}{\sigma^2}e^{-\frac{l^2}{2\sigma^2}}\right)^2 \tag{17}$$

where h_M is the (partial) histogram of the data, l_0 and l_c are the lower and upper bounds of the histogram and K is an amplitude factor. The upper bound is usually set to $l_c = 2\sigma_n$ to cut off the influence of the Rician data. An initial estimate of this parameter is needed. Chang in [137] makes the fit using a Gaussian smoothing of the histogram of the image:

$$\widehat{\sigma_n} = \arg\max_{\sigma}\frac{1}{ns}\sum_{i=0}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\sigma-x_i}{s}\right)^2} \tag{18}$$

with n the sample size, $s = 1.06\sigma_0 n^{1/5}$ the smoothing width and σ_0 the standard deviation of the smoothing kernel.

Using the distributions for the sample moments derived in [8], it is possible to use the least squares fitting over the histograms of the sample moments instead of using the histogram of the image. Accordingly, the following new estimators are proposed (for details see Appendix 6.1)

$$\widehat{\sigma_n} = \arg\min_{\sigma,K} \sum_{l=l_0}^{l_c} \left(h_{<M^2>}(l) - K \frac{l^{N-1}N^N}{(2\sigma^2)^N \Gamma(N)} e^{-\frac{lN}{2\sigma^2}} \right)^2 \qquad (19)$$

$$\widehat{\sigma_n} = \arg\min_{\sigma,K} \sum_{l=l_0}^{l_c} \left(h_{<M>}(l) - K \frac{l^{2N-1}N^N}{2^{N-1}b^N \Gamma(N)} e^{-\frac{l^2N}{2b}} \right)^2 \qquad (20)$$

$$\widehat{\sigma_n} = \arg\min_{\sigma,K} \sum_{l=l_0}^{l_c} \left(h_{\sqrt{<M^2>}}(l) - K \frac{l^{2N-1}N^N}{2^{N-1}\sigma^{2N} \Gamma(N)} e^{-\frac{l^2N}{2\sigma^2}} \right)^2 , \quad (21)$$

where $h_{<M^2>}$, $h_{<M>}$ and $h_{\sqrt{<M^2>}}$ are the histograms of the sample second order moment, the sample mean, and the square root of the second order moment, respectively. All the moments are local and computed in a neighborhood. N is the number of points used for the local sample moment estimation and $b \approx \sigma^2 \frac{\pi}{4}$.

c. *Maximum likelihood estimation.* Using the joint PDF of the histogram of Rayleigh data, [84] proposes a maximum likelihood (ML) estimation of σ_n

$$\widehat{\sigma_n} = \arg\min_{\sigma} \left[N_k \log \left(e^{-\frac{l_0^2}{2\sigma^2}} - e^{-\frac{l_k^2}{2\sigma^2}} \right) - \sum_{i=1}^{k} n_i \log \left(e^{-\frac{l_{i-1}^2}{2\sigma^2}} - e^{-\frac{l_i^2}{2\sigma^2}} \right) \right]$$
$$(22)$$

with $l_i \ i = 0, \cdots, K$ the boundaries of the histogram bins, n_i the number of observations in bin $[l_{i-1}, l_i]$ and $N_k = \sum_{i=1}^{k} n_i$. A method to select the number of bins is also provided.

If the ML estimation is done over the second order moment of the Rayleigh data, a new estimator may be defined (see Appendix 6.1 for details)

$$\widehat{\sigma_n} = \arg\min_{\sigma} \left[N_k \log \left(\Gamma \left(N+1, l_0 \frac{N}{2\sigma^2} \right) - \Gamma \left(N+1, l_k \frac{N}{2\sigma^2} \right) \right) \right.$$
$$\left. - \sum_{i=1}^{k} n_i \log \left(\Gamma \left(N+1, l_{i-1} \frac{N}{2\sigma^2} \right) - \Gamma \left(N+1, l_i \frac{N}{2\sigma^2} \right) \right) \right] (23)$$

Note that the histogram now considered is the second order moment. For a more compact dynamic range, the square root of this moment may be used instead:

$$\widehat{\sigma_n} = \arg\min_\sigma \left[N_k \log \left(\varGamma \left(N, \frac{l_0^2 N}{2\sigma^2} \right) - \varGamma \left(N, \frac{l_k^2 N}{2\sigma^2} \right) \right) \right.$$

$$\left. - \sum_{i=1}^{k} n_i \log \left(\varGamma \left(N, \frac{l_{i-1}^2 N}{2\sigma^2} \right) - \varGamma \left(N, \frac{l_i^2 N}{2\sigma^2} \right) \right) \right]. \qquad (24)$$

3. *Estimator based on Rician data.* Assuming that all the data are Rician distributed and no Rayleigh background is present, and the SNR is high enough ($A/\sigma_n > 5$), the following estimator may be defined [8]

$$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(M(\mathbf{x}))_\mathbf{x}\}. \qquad (25)$$

Note that it is similar to the estimation of the variance of Gaussian additive noise [108]. If the image has Rayleigh background, a selection inside the tissue area with no background, say $\mathbf{R}$, must be considered. This region must be roughly defined using a thresholding method, such as the ones proposed in [138, 139]. The estimator may be redefined as

$$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(M(\mathbf{x_R}))_\mathbf{x}\}, \ \mathbf{x_R} \in \mathbf{R} \qquad (26)$$

where $M(\mathbf{x_R})$ are image values inside the region $\mathbf{R}$. Since this method is based on an approach of the sample variance, the estimation is less accurate than the Rayleigh-based methods. Thus, it is advised only to be used when no Rayleigh areas are available.

A survey of the methods together with the labels that will be used in the experiments may be found in Table 2.

The different noise estimators presented in this section have different performance under different environments. Estimators Bk-M2 and Bk-M1, eqs. (10)-(11), have smaller variances because more points are used in the estimation; they also have the added advantage of avoiding the calculation of the mode or any optimization process. However, they also present some drawbacks. First, the background must be segmented, usually with manual human intervention, making the method difficult to integrate in automatic procedures. What is more, if the segmentation is automatically done, and some error is committed, these two estimators may be highly biased, as reported in [8]. The estimation also assumes that the original background is homogeneous, and signal-free. This assumption makes the estimate sensitive to errors and ghosting artifacts.

Estimators that do not need prior segmentation of the background are more robust to errors, artifacts and small traces of signal in the background [8, 84]. Due to these sources of error, it is advisable to use estimators in eqs. (12)-(24) even when an automatic segmentation of the background is feasible. A robust estimation method would involve a rough automatic segmentation of the Rayleigh area –using for instance some thresholding method, as the ones proposed in [138, 139]– and a subsequent estimation using one of eqs. (12)-(24) over the thresholded data.

To finish, we would like to stress the importance of the Rician-based estimator in eq. (25)-(26) when dealing with MR data without background. Note that MRI is not only devoted to brain imaging; in other areas only tissue is scanned, and therefore no estimator based on background assumptions can be used.

2.4 Noise estimation assuming non–central Chi distributions

Many methods have been proposed for noise estimation in MR data assuming a Rician distribution. In this section the main ones will be extended to the non–central Chi model. Note that this model is suitable for multiple coil imaging without sub-sampling and it can be also extended to GRAPPA reconstructed pMRI.

Two different noise measures will be estimated: σ_n^2, i.e. the noise variance in each coil and $\sigma_{nL}^2 = L\sigma_n^2$, with L the number of coils. Noise estimators for this model may be defined as (see Table 1 for notation, see Appendix 6.2 for the moments of Chi distribution and see Appendix 6.1 for derivation):

1. *Estimators that need prior segmentation of the background area* [27, 46]:

$$\widehat{\sigma_{nL}^2} = \frac{1}{2}\langle M_L^2(\mathbf{x_B})\rangle \tag{27}$$

$$\widehat{\sigma_{nL}} = \frac{1}{\sqrt{2}}\langle M_L(\mathbf{x_B})\rangle \frac{\sqrt{L}\Gamma(L)}{\Gamma\left(L+\frac{1}{2}\right)} \tag{28}$$

Estimator in eq. (27) is the Maximum Likelihood estimator for the Central Chi distribution.

2. *Estimators based on the background that do not need prior segmentation.* These estimators assume a central Chi PDF in the background of the image.

   a. Using the mode of some local parameter related to σ_n, as the maximum of the histogram of the image

$$\widehat{\sigma_n} = \frac{1}{\sqrt{2L-1}}\text{mode}\{M_L(\mathbf{x})\}. \tag{29}$$

Using the sample local moments of the image:

$$\widehat{\sigma_{nL}^2} = \frac{1}{2}\text{mode}\{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\} \tag{30}$$

$$\widehat{\sigma_{nL}} = \frac{1}{\sqrt{2}}\text{mode}\{\langle M_L(\mathbf{x})\rangle_{\mathbf{x}}\} \frac{\sqrt{L}\Gamma(L)}{\Gamma\left(L+\frac{1}{2}\right)} \tag{31}$$

$$\widehat{\sigma_n^2} = \left(2L - \frac{2\Gamma^2\left(L+\frac{1}{2}\right)}{\Gamma^2(L)}\right)^{-1}\text{mode}\{\text{Var}(M_L(\mathbf{x}))_{\mathbf{x}}\} \tag{32}$$

b. *Methods that fit a central Chi-related distribution to the histogram of the data.*
First using least squares we fit a central-Chi distribution with the histogram of
the data in the background

$$\widehat{\sigma_n} = \arg \min_{\sigma, K} \sum_{m=m_0}^{m_c} \left(h_M(m) - K \frac{2^{1-L}}{\Gamma(L)} \frac{m^{2L-1}}{\sigma^{2L}} e^{-\frac{m^2}{2\sigma^2}} \right)^2 \tag{33}$$

where h_M is the (partial) histogram of the data, m_0 and m_c are the lower and
upper bounds of the histogram and K is an amplitude factor. The upper bound
is set to $m_c = 2\sigma_n$ to cut off the influence of the non–central Chi data.
Using the distributions for the sample moment, it is possible to use the least
square fitting over the histograms of the second order sample moment instead
of using the histogram of the image. Accordingly, the following new estimator
is proposed (see Appendix 6.1 for details)

$$\widehat{\sigma_n} = \arg \min_{\sigma, K} \sum_{m=m_0}^{m_c} \left(h_{<M_L^2>}(l) - K \frac{m^{NL-1} N^{NL}}{(2\sigma^2)^{NL} \Gamma(NL)} e^{-\frac{mN}{2\sigma^2}} \right)^2 \tag{34}$$

$h_{<M_L^2>}$ being the histogram of the sample second order moment and N the
number of points used for the sample moment estimation.
c. *Maximum likelihood estimation.* Using the joint PDF of the histogram of cen-
tral Chi data, a maximum likelihood estimation of σ_n is proposed (see Ap-
pendix 6.1 for details)

$$\widehat{\sigma_n} = \arg \min_{\sigma} \left[N_k \log \left(\Gamma \left(L, \frac{m_0^2}{2\sigma^2} \right) - \Gamma \left(L, \frac{m_c^2}{2\sigma^2} \right) \right) \right.$$
$$\left. - \sum_{i=1}^{K} n_i \log \left(\Gamma \left(L, \frac{m_{i-1}^2}{2\sigma^2} \right) - \Gamma \left(L, \frac{m_i^2}{2\sigma^2} \right) \right) \right] \tag{35}$$

with m_i $i = 0, \cdots, K$ the boundaries of the histogram bins, n_i the number of
observations in bin $[m_{i-1}, m_i]$ and $N_k = \sum_{i=1}^{K} n_i$.
If the ML estimation is done over the second order moment of the data, a new
estimator may be defined (see Appendix 6.1 for details)

$$\widehat{\sigma_n} = \arg \min_{\sigma} \left[N_k \log \left(\Gamma \left(NL, m_0 \frac{N}{2\sigma^2} \right) - \Gamma \left(NL, m_k \frac{N}{2\sigma^2} \right) \right) \right.$$
$$\left. - \sum_{i=1}^{K} n_i \log \left(\Gamma \left(NL, m_{i-1} \frac{N}{2\sigma^2} \right) - \Gamma \left(NL, m_i \frac{N}{2\sigma^2} \right) \right) \right] \tag{36}$$

3. *Estimator assuming that no background is present.* If the data is non–central Chi
distributed and the SNR is high enough, i.e. $A_L >> \sigma_n$ (see simplification of
the variance in Appendix 6.2):

$$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(M_l(\mathbf{x}))_\mathbf{x}\} \tag{37}$$

Note that it is again similar to the estimation of the variance of Gaussian additive noise [108] and also similar to the estimation of σ_n^2 for Rician noise under the same high SNR assumption.

A survey of the methods together with the labels that will used in the experiments may be found in Table 2.

Noise can also be estimated using other methods, like multiple images or the method of moments [114] or the methods porposed by Koay *et al.* [48, 60], but they will not be considered in this paper.

3 Materials and Methods

Synthetic experiments for noise estimation were carried out using a 2D synthetic slice from a BrainWeb MR volume [53], see Fig. 1, with intensity values in $[0 - 255]$. The average intensity value for the White Matter is 158, for the Gray Matter is 105, for the cerebrospinal fluid 36 and 0 for the background. The image has been used for noise estimation out of Rician and non–central Chi data, therefore it has been corrupted with noise in three different ways, which leads to three different experiments:

1. Artificially corrupting with uncorrelated Rician noise with σ_n ranging in $[5-30]$, which mean SNR values in $[5.3 - 31.6]$ for White Matter and in $[2.5 - 21]$ for Gray Matter. Estimators based on the Rayleigh background that do not need prior segmentation have been tested with this method.
2. Simulating a 8-coil system, 8 images have been corrupted with independent Gaussian noise and the composite image is created using SoS. This way, the resulting image will follow a non–central Chi distribution. Different values of σ_{nL} ranging in $[5 - 30]$ are considered. Estimators based on the central Chi background that do not need prior segmentation have been tested with this method.
3. Due to interpolation of the data in the scanner, the initial assumption of uncorrelated noise does not usually hold in real MR data. Although the Rician and Rayleigh assumptions are still valid for single coil data –note that the averaging of N Gaussians with zero mean and variance σ^2 is another Gaussian with zero mean and variance σ^2/N,– estimators based on moments may fail. To test the estimators we generate correlated noise processes using a 3×3 averaging window prior to adding them to the complex MR data and to generating the magnitude image. Again, σ_n range in $[5-30]$, and the estimators used in the first experiment are tested again.

In all the experiments, for every σ_n value, the mean and the variance of 1000 experiments for each estimation method are considered. The whole estimation is done over the final magnitude image. The local moments have been calculated using 7×7 neighborhoods, and histograms have been calculated over 1000 bins for all cases. For the sake of comparison, no optimization in the election of the bins has been used in any method.

	Assump.	Method	Refs.	Label
1	Rayleigh	$\widehat{\sigma_n^2} = \frac{1}{2}\langle M^2(\mathbf{x_B})\rangle$	[86, 81]	Bk M2
	Rayleigh	$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\langle M(\mathbf{x_B})\rangle$	[81]	Bk M1
2-a	Rayleigh	$\widehat{\sigma_n} = \text{mode}\{M(\mathbf{x})\}$	[37, 83]	Max H
	Rayleigh	$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\{\langle M^2(\mathbf{x})\rangle_\mathbf{x}\}$	[8]	Mode M2
	Rayleigh	$\widehat{\sigma_n} = \sqrt{\frac{N}{2N-1}}\text{mode}\left\{\sqrt{\langle M^2(\mathbf{x})\rangle_\mathbf{x}}\right\}$	New	Mode M2(sq)
	Rayleigh	$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\text{mode}\{\langle M(\mathbf{x})\rangle_\mathbf{x}\}$	[8]	Mode M1
	Rayleigh	$\widehat{\sigma_n^2} = \frac{2}{4-\pi}\text{mode}\{\text{Var}(M(\mathbf{x}))_\mathbf{x}\}$	[8]	Mode V1
2-b	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{l=l_0}^{l_c}\left(h_M(l) - K\frac{l}{\sigma^2}e^{-\frac{l^2}{2\sigma^2}}\right)^2$	[37]	Brummer
	Rayleigh	$\widehat{\sigma_n} = \arg\max_{\sigma}\frac{1}{nh}\sum_{i=0}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\sigma-x_i}{h}\right)^2}$	[137]	Chang
	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{l=l_0}^{l_c}\left(h_{<M^2>}(l) - K\frac{l^{N-1}N^N}{(2\sigma^2)^N\Gamma(N)}e^{-\frac{lN}{2\sigma^2}}\right)^2$	New	LS-M2
	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{l=l_0}^{l_c}\left(h_{<M>}(l) - K\frac{l^{2N-1}N^N}{2^{N-1}b^N\Gamma(N)}e^{-\frac{l^2N}{2b}}\right)^2$	New	LS-M1
	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{l=l_0}^{l_c}\left(h_{\sqrt{<M^2>}}(l) - K\frac{l^{2N-1}N^N}{2^{N-1}\sigma^{2N}\Gamma(N)}e^{-\frac{l^2N}{2\sigma^2}}\right)^2$	New	LS-M2 (sq)
2-c	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma}\left[N_k\log\left(e^{-\frac{l_0^2}{2\sigma^2}} - e^{-\frac{l_k^2}{2\sigma^2}}\right) - \sum_{i=1}^k n_i\log\left(e^{-\frac{l_{i-1}^2}{2\sigma^2}} - e^{-\frac{l_i^2}{2\sigma^2}}\right)\right]$	[84]	Sijbers
	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma}\left[N_k\log\left(\Gamma\left(N+1,l_0\frac{N}{2\sigma^2}\right) - \Gamma\left(N+1,l_k\frac{N}{2\sigma^2}\right)\right) - \sum_{i=1}^k n_i\log\left(\Gamma\left(N+1,l_{i-1}\frac{N}{2\sigma^2}\right) - \Gamma\left(N+1,l_i\frac{N}{2\sigma^2}\right)\right)\right]$	New	ML-M2
	Rayleigh	$\widehat{\sigma_n} = \arg\min_{\sigma}\left[N_k\log\left(\Gamma\left(N,\frac{l_0^2N}{2\sigma^2}\right) - \Gamma\left(N,\frac{l_k^2N}{2\sigma^2}\right)\right) - \sum_{i=1}^k n_i\log\left(\Gamma\left(N,\frac{l_{i-1}^2N}{2\sigma^2}\right) - \Gamma\left(N,\frac{l_i^2N}{2\sigma^2}\right)\right)\right]$	New	ML-M2(sq)
3	Rician	$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(M(\mathbf{x}))_\mathbf{x}\}$	[8]	Mode Vr
4	central χ	$\widehat{\sigma_{nL}^2} = \frac{1}{2}\langle M_L^2(\mathbf{x_B})\rangle$	[27, 46]	Bk M2–χ
	central χ	$\widehat{\sigma_{nL}} = \frac{1}{\sqrt{2}}\langle M_L(\mathbf{x_B})\rangle\frac{\sqrt{L}\Gamma(L)}{\Gamma(L+\frac{1}{2})}$	[46]	Bk M1–χ
5-a	central χ	$\widehat{\sigma_n} = \frac{1}{\sqrt{2L-1}}\text{mode}\{M_L(\mathbf{x})\}$	New	Max H–χ
	central χ	$\widehat{\sigma_{nL}^2} = \frac{1}{2}\text{mode}\{\langle M_L^2(\mathbf{x})\rangle_\mathbf{x}\}$	New	Mode M2–χ
	central χ	$\widehat{\sigma_{nL}} = \frac{1}{\sqrt{2}}\text{mode}\{\langle M_L(\mathbf{x})\rangle_\mathbf{x}\}\frac{\sqrt{L}\Gamma(L)}{\Gamma(L+\frac{1}{2})}$	New	Mode M1–χ
	central χ	$\widehat{\sigma_n^2} = \left(2L - \frac{2\Gamma^2(L+\frac{1}{2})}{\Gamma^2(L)}\right)^{-1}\text{mode}\{\text{Var}(M_L(\mathbf{x}))_\mathbf{x}\}$	New	Mode V1–χ
5-b	central χ	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{m=m_0}^{m_c}\left(h_M(m) - K\frac{2^{1-L}}{\Gamma(L)}\frac{m^{2L-1}}{\sigma^{2L}}e^{-\frac{m^2}{2\sigma^2}}\right)^2$	New	LS-H–χ
	central χ	$\widehat{\sigma_n} = \arg\min_{\sigma,K}\sum_{m=m_0}^{m_c}\left(h_{<M_L^2>}(l) - K\frac{m^{NL-1}N^{NL}}{(2\sigma^2)^{NL}\Gamma(NL)}e^{-\frac{mN}{2\sigma^2}}\right)^2$	New	LS-M2–χ
5-c	central χ	$\widehat{\sigma_n} = \arg\min_{\sigma}\left[N_k\log\left(\Gamma\left(L,\frac{m_0^2}{2\sigma^2}\right) - \Gamma\left(L,\frac{m_c^2}{2\sigma^2}\right)\right) - \sum_{i=1}^K n_i\log\left(\Gamma\left(L,\frac{m_{i-1}^2}{2\sigma^2}\right) - \Gamma\left(L,\frac{m_i^2}{2\sigma^2}\right)\right)\right]$	New	ML-H–χ
	central χ	$\widehat{\sigma_n} = \arg\min_{\sigma}\left[N_k\log\left(\Gamma\left(NL,m_0\frac{N}{2\sigma^2}\right) - \Gamma\left(NL,m_k\frac{N}{2\sigma^2}\right)\right) - \sum_{i=1}^K n_i\log\left(\Gamma\left(NL,m_{i-1}\frac{N}{2\sigma^2}\right) - \Gamma\left(NL,m_i\frac{N}{2\sigma^2}\right)\right)\right]$	New	ML-M2–χ
6	non-central χ	$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(M_l(\mathbf{x}))_\mathbf{x}\}$	New	Mode Vn
7	Gaussian	$\widehat{\sigma_n^2} = \text{mode}\{\text{Var}(C_{l_j}(\mathbf{x}))_\mathbf{x}\}$	[108]	Mode Vg

Table 2 Survey of noise estimators for single and multiple coil MR data. Note that estimators in boxes 1 and 4 require background segmentation.

For estimation over real data, two different data sets have been considered. First, one slice of the baseline from a single coil DTI acquisition, scanned in a 1.5 Tesla GE system, slick thickness: 5 mm (real data set 1, RD1); and second a multiple coil acquisition, from an 8 coil GE Signa 1.5 Tesla EXCITE 11m4 scanner, FSE Pulse Sequence, TR=500 ms, TE= 13.8 ms, matrix size= 256×256, FOV=20×20 cm, slick thickness= 5 mm (real data set 2, RD2).

RD1 will be used to compare estimators that require a background segmentation with those that do not need it. To that end, the Rayleigh noisy background is automatically segmented. Noise estimation is done using different estimators over the segmented background and over the whole image.

RD2 will be used to test the estimation assumptions over the complex spatial domain and over Rician and non-central Chi composite images.
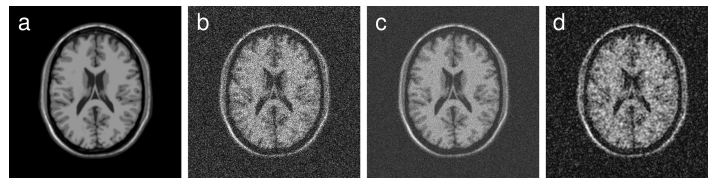
Fig. 1 Synthetic MR image from Brainweb used for the noise estimation experiments. (a) Original image. (b) Image with uncorrelated Rician noise. (c) Noisy image following a non–central Chi distribution, simulating a 8–coils system with no subsampling. (d) Image with correlated Rician noise.
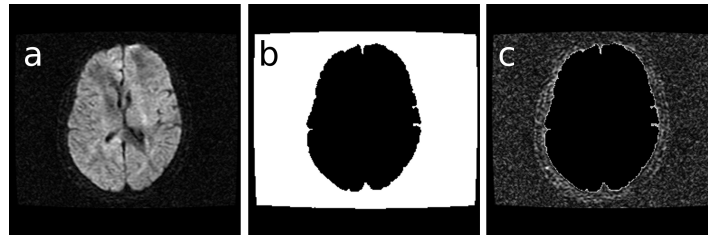
Fig. 2 Real Data set 1 (RD1). Slice from a single coil DTI-MR acquisition. (a) Original image. (b) Automatic background segmentation mask. (c) Segmented background.
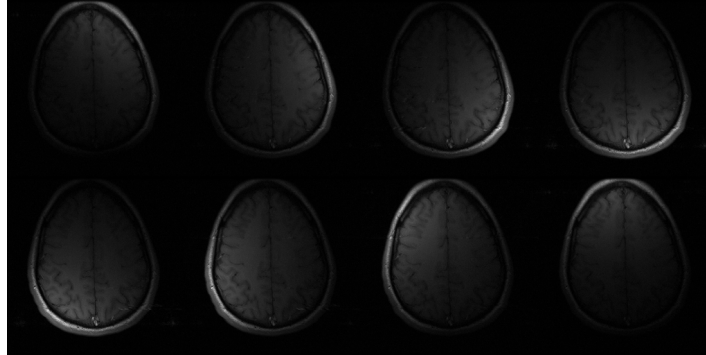
Fig. 3 Real Data set 2 (RD2). Slice from an 8-coil adquisition.
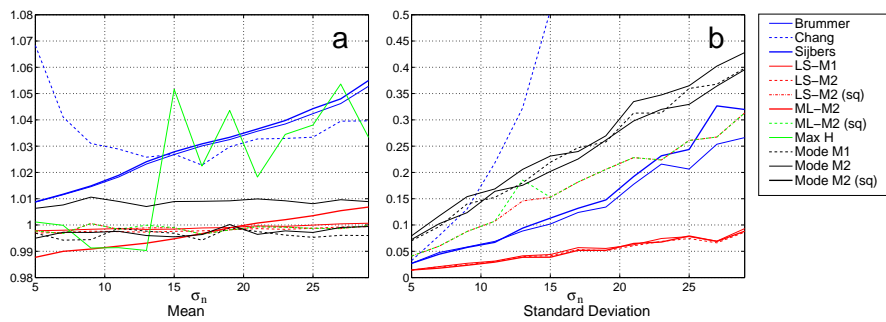
Fig. 4 Comparison of different noise estimators for Rician magnitude MR data assuming uncorrelated noise. 1000 experiments are considered for each sigma value. (a) Mean of the estimated value divided by the actual value. (b) Standard deviation of the experiments. (Standard deviation for 'Max H' is out of scale).

4 Results

4.1 Synthetic experiments for noise Estimation

We will focus **first** on those methods based on the Rayleigh background that do not need prior segmentation. Results for a wide range of σ_n values are shown in Fig. 4 (see Table 2 for reference). In Fig 4-(a), the mean of the 1000 experiments divided by the actual value of σ_n is depicted. Accordingly, the closer to 1, the better the estimation. In Fig 4-(b), the standard deviation of the experiments is shown; the lower the value, the better the estimation.

From the results it can be seen that although all the methods accurately estimate the noise –the values are in a small range between 0.99 and 1.06– the ones based on moments (local mean and local second order moment) have a smaller bias for a wider range of values. Max-H is the method with the weakest performance, due to

the difficulty to accurately estimate the mode out of a wide distribution. Methods based on the mode of moments (Mode-M1, Mode-M2) show very good average behavior, although its variance is slightly larger than that of other methods, due to the non linear nature of the mode operator. The proposed new methods (LS-M1, LS-M2 y ML-M2) show the better performance in terms of unbiased estimation and lower variance.
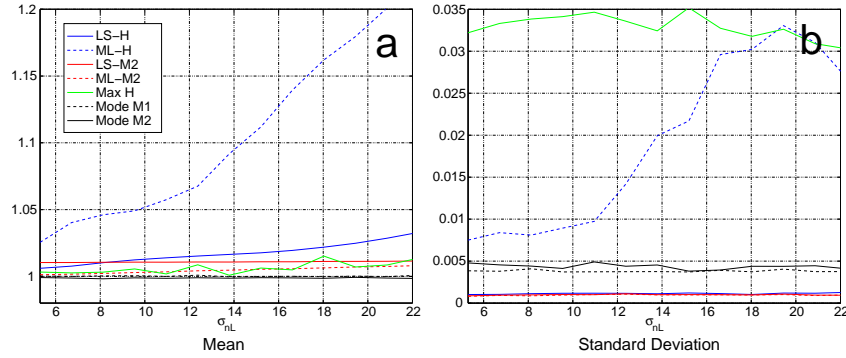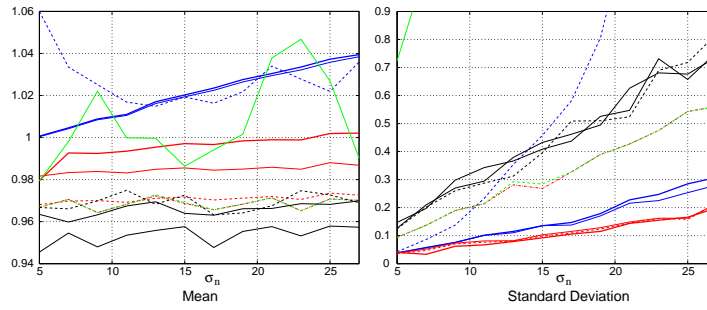

Fig. 5 Comparison of different noise estimators for multiple coil MR data assuming uncorrelated noise. 1000 experiments are considered for each sigma value, $L = 8$ coils and no subsampling of the **k**-space. (a) Mean of the estimated value divided by the actual value. (b) Standard deviation of the experiments.

The **second** synthetic experiment focuses on estimation over a non–central Chi distributed image, and we will use the estimators based on the the central Chi background that do not need prior segmentation. Labels for each method can be found in Table 2. Results are shown in Fig. 5: in Fig 5-(a), the mean of the 1000 experiments divided by the actual value of $\sigma_n L$ is depicted and in Fig 5-(b), the standard deviation of the experiments is shown. The interpretation is the same as the previous experiment.

From the results it can be seen that although most of the methods accurately estimate the noise –the values are in a small range between 0.99 and 1.04 for all methods but ML-H– the ones based on moments have again a smaller bias for a wider range of values. Max-H and ML-H are the methods with the weakest performance. It is probably due, as in the Rician case, to the use of the histogram of the image. Methods dealing with the histogram of the second order moments show a more robust behavior.

Results for the third experiment –estimation out of correlated Rician MR data– are now in Fig. 6. While Brummer's, Chang's and Sijbers' methods do not show any significant change, the moments-based methods show a small negative bias and a larger variance, as expected. Despite the worsening variance, LS-M1, LS-M2 and ML-M2 are still by far the most accurate estimators.

To conclude we must say that from the synthetic experiments, methods showing the best performance are those based on the second order moment, mainly ML- and

Fig. 6 Comparison of different noise estimators for MR data with correlated noise. 1000 experiments are considered for each sigma value. Left: mean of the estimated value divided by the actual value. Right: Standard deviation of the experiments. (See Legend in Fig. 4).

LS–based methods, both for Rician and non–central Chi distributions. However, the methods based on the mode are the only ones that do not rely on an optimization method to achieve the final solution. So, although their variance is slightly larger than other methods, they present a simple and fast alternative for noise estimation in MR.

4.2 Estimation on real data: the background problem

The behavior of the estimators that require a segmented background is analyzed using the RD1. The image in Fig. 2 has an artificial zero background that must be removed before estimation. Noisy background is automatically segmented and the estimation is done using different estimators over the segmented background and over the whole image. Results are on Table 3. Labels for each method can be found in Table 2. The estimated values are compared with each other as there is no ground truth to refer to.

Since the segmented background is not totally homogeneous –note that part of the skull has been added to the automatically segmented area,– estimators that require segmentation (Bk-M2 and Bk-M1) are clearly biased when compared with the other estimators. Note that the other estimators do not show a great variation when using only the background of the whole image. ML and LS based estimators are the ones with less influence from the segmentation.

4.3 Estimation on real data: the complex domain

As a final experiment, the noise estimation will be done over the complex data of RD2, using the method described in Section 2.2, and compared to the magnitude estimators. For the magnitude image two different sets are considered:

Method	Segmented Bkg	Whole Image
Bk M2	29.59	—
Bk M1	27.55	—
Max H	21.98	22.00
Mode M1	23.85	22.34
Mode M2	22.89	22.90
Mode V1	21.83	21.39
Brummer	23.52	23.55
LS M1	23.64	23.64
LS M2	23.41	23.41
Sijbers	23.59	23.61
ML-M2	23.99	24.01

Table 3 σ_n estimation using different noise estimation over real single coil MR data.

1. The composite magnitude image, as the SoS of the complex data. One non–central Chi image is generated.
2. 8 Rician magnitude images, one from each coil.

The advantage of estimating in the complex domain is that no background assumption is needed. Results for σ_n estimation are on Table 4 (for the Rician the average of the estimation for each coil is considered). Note that the estimation done in the complex domain is consistent with the estimation done over the magnitude image. Estimator Max-H overestimates the level of noise for the non-central Chi experiment, but the rest of the estimators give a similar value of σ_n. Then, we can conclude that the estimation over the complex data is equivalent to the estimation over the magnitude image. The former will always be more accurate though complex data are not always available.

Data	Method	$\widehat{\sigma_n}$
Complex Image (Real)	Mode Vg	$1.41 \ 10^{-2}$
Complex Image (Imag)	Mode Vg	$1.41 \ 10^{-2}$
Magnitude (Non–central Chi)	Max H	$2.02 \ 10^{-2}$
	Mode M1	$1.56 \ 10^{-2}$
	Mode M2	$1.42 \ 10^{-2}$
	Mode V1	$1.41 \ 10^{-2}$
Magnitude (Rician)	Max H	$1.42 \ 10^{-2}$
	Mode M1	$1.41 \ 10^{-2}$
	Mode M2	$1.41 \ 10^{-2}$
	Mode V1	$1.42 \ 10^{-2}$

Table 4 Different σ_n estimation over the real data set 2. Complex data and magnitude are considered.

5 Conclusions

Some important topics about noise estimation in MR data have been studied. The main noise estimation methods out of Rician magnitude MR data are revisited and classified according to the model they rely on. Some new estimators based on Least Squares and Maximum Likelihood approaches are proposed. These new methods show a more robust behavior when compared with the existing ones. All the reviewed estimators, together with the new ones, have been extended to the important case of non–central Chi distributed MR data.

Estimators based on a Rayleigh (or central Chi) background which do not require its segmentation are more robust to this issue, but may benefit as well of a previous rough segmentation. Finally, estimators based on Rician (or non–central Chi) data generally have higher variance, but benefit from not needing a background to be present in the image, which may be a requisite in some MRI data sets.

In addition, a method to estimate the noise from the data in the spatial complex domain is also provided. Note that the estimation of noise in the complex domain may benefit of the great existing knowledge about estimation in Gaussian data in literature. On the other hand, an estimation based on the mode operator, as the ones proposed, does not requires any segmentation or background assumption, and it does not rely on any optimization method to achieve the final solution. Thus, although the estimation variance for this method may be slightly larger than for others, we think it is a simple and fast alternative for noise estimation that could be directly implemented in the scanning software.

6 Appendices

6.1 About the noise estimators

Equations related to the noise estimators are derived. Let $R_i(\sigma^2)$, $i = \{1, \cdots, N\}$ be a set of random variables with Rayleigh distribution and let $X_i(L, \sigma^2)$ be a set of independent and identically distributed (IID) central Chi variables. Then

$$S_1 = \frac{1}{N} \sum_{i=1}^{N} R_i(\sigma^2) \qquad p_{S_1}(x) = \frac{x^{2N-1} N^N}{2^{N-1} b^N \Gamma(N)} e^{-\frac{x^2 N}{2b}} u(x)$$

$$S_2 = \frac{1}{N} \sum_{i=1}^{N} R_i^2(\sigma^2) \qquad p_{S_2}(x) = \frac{x^{N-1} N^N}{(2\sigma^2)^N \Gamma(N)} e^{-\frac{x N}{2\sigma^2}} u(x)$$

$$S_{\sqrt{2}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} R_i^2(\sigma^2)} \; p_{S_{\sqrt{2}}}(x) = \frac{x^{2N-1} N^N}{2^{N-1} \sigma^{2N} \Gamma(N)} e^{-\frac{x^2 N}{2\sigma^2}} u(x)$$

$$T_2 = \frac{1}{N} \sum_{i=1}^{N} X_i^2(L, \sigma^2) \; p_{T_2}(x) = \frac{x^{NL-1}}{\Gamma(NL)} \left(\frac{N}{2\sigma^2}\right)^{NL} e^{-\frac{x N}{2\sigma^2}} u(x)$$

In [8] estimators derived from the first three distributions are defined. The mode of the last distribution is

$$\text{mode}\{T_2\} = \frac{NL - 1}{N} 2\sigma^2$$

An estimator may be defined using the second order moment (and assuming $NL >> 1$) as

$$\widehat{\sigma_n^2} = \frac{1}{2L} \text{mode}\{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\}.$$

The joint PDF of the histogram of a distribution $p(x; \mathbf{a})$ is given by [84]

$$p(n_i|\mathbf{a}, l_i) = \frac{N_k!}{\prod_{i=1}^{K} n_i!} \prod_{i=1}^{K} p_i^{n_i}(\mathbf{a})$$

with p_i the probability that an observation falls in the range $[l_{i-1}, l_i]$:

$$p_i(\mathbf{a}) = \frac{\int_{l_{i-1}}^{l_i} p(y; \mathbf{a})d(y)}{\sum_{i=1}^{K} \int_{l_{i-1}}^{l_i} p(y; \mathbf{a})d(y)}$$

For S_2 we can define

$$p_i(\sigma, N) = \frac{\Gamma\left(N+1, l_{i-1}\frac{N}{2\sigma^2}\right) - \Gamma\left(N+1, l_i\frac{N}{2\sigma^2}\right)}{\Gamma\left(N+1, l_0\frac{N}{2\sigma^2}\right) - \Gamma\left(N+1, l_k\frac{N}{2\sigma^2}\right)} \tag{38}$$

being $\Gamma(n, x)$ the (upper) incomplete Gamma function $\Gamma(n, x) = \int_x^{\infty} t^{n-1}e^{-t}dt$. For the second order moment of the central Chi distribution T_2

$$p_i(\sigma, N) = \frac{\Gamma\left(NL, m_{i-1}\frac{N}{2\sigma^2}\right) - \Gamma\left(NL, m_i\frac{N}{2\sigma^2}\right)}{\Gamma\left(NL, m_0\frac{N}{2\sigma^2}\right) - \Gamma\left(NL, m_k\frac{N}{2\sigma^2}\right)} \tag{39}$$

If the same reasoning is done over the histogram of the background of the parallel image the probability is

$$p_i(\sigma, N) = \frac{\Gamma\left(L, \frac{m_{i-1}^2}{2\sigma^2}\right) - \Gamma\left(L, \frac{m_i^2}{2\sigma^2}\right)}{\Gamma\left(L, \frac{m_0^2}{2\sigma^2}\right) - \Gamma\left(L, \frac{m_k^2}{2\sigma^2}\right)}. \tag{40}$$

6.2 Moments of the Central and non–central Chi distributions

For a central Chi distribution the main moments are

$$E\{M_L\} = \sigma_n \sqrt{2} \frac{\Gamma\left(L + \frac{1}{2}\right)}{\Gamma(L)} \tag{41}$$

$$E\{M_L^2\} = 2\sigma_n^2 L \tag{42}$$

$$\sigma_{M_L}^2 = \sigma_n^2 \left(2L - 2 \left[\frac{\Gamma^2\left(L + \frac{1}{2}\right)}{\Gamma^2(L)} \right] \right) \tag{43}$$

For the non–central Chi distribution the second order moment and the mean are

$$E\{M_L^2\} = 2\sigma_n^2 L + A_L^2 \tag{44}$$

$$E\{M_L\} = \sqrt{2}\sigma_n \frac{\Gamma\left(L + \frac{1}{2}\right)}{\Gamma(L)} \, _1F_1\left(-\frac{1}{2}, L, -\frac{A_L^2}{2\sigma_n^2}\right) \tag{45}$$

$_1F_1(a, b, z)$ being the Confluent Hypergeometric Function of the first kind. If $x = \frac{A_L}{\sigma_n}$, assuming that x is large (high SNR assumption), and taking into account that

$$_1F_1(b - a, b, z) = e^z \, _1F_1(a, b, -z)$$

$$_1F_1(a, b, z) = \frac{\Gamma(b)}{\Gamma(a)} e^z z^a - b \left(1 - \frac{(a-1)(a-b)}{z} + O\left(\frac{1}{z^2}\right) \right)$$

we can make the approximation

$$E\{M_L\} = \sqrt{2}\sigma_n \frac{\Gamma\left(L + \frac{1}{2}\right)}{\Gamma(L)} e^{-\frac{x^2}{2}} \, _1F_1\left(L + \frac{1}{2}, L, \frac{x^2}{2}\right)$$

$$= \sqrt{2}\sigma_n \frac{\Gamma\left(L + \frac{1}{2}\right)}{\Gamma(L)} e^{-\frac{x^2}{2}} e^{\frac{x^2}{2}} \frac{\Gamma(L)}{\Gamma\left(L + \frac{1}{2}\right)} \frac{x}{\sqrt{2}} \left(1 + \frac{L - \frac{1}{2}}{x^2} + O\left(\frac{1}{x^4}\right) \right)$$

$$= \sigma_n x \left(1 + \frac{L - \frac{1}{2}}{x^2} + O\left(\frac{1}{x^4}\right) \right)$$

Therefore, the variance for high SNR may be approximated as

$$\sigma_{M_L}^2 = \sigma_n^2 \left(2L + x^2 - x^2 \left(1 + \frac{L - \frac{1}{2}}{x^2} + O\left(\frac{1}{x^4}\right) \right)^2 \right)$$

$$= \sigma_n^2 \left(1 + O\left(\frac{1}{x^4}\right) \right) \tag{46}$$

Acknowledgments

Effective Noise Estimation and Filtering from Correlated Multiple-Coil MR data

Santiago Aja-Fernández, Véronique Brion, Antonio Tristán-Vega*

Abstract Modern MRI protocols based on multiple-coil acquisitions have carried on a new attention to noise and signal statistical modeling, as long as most of the existing techniques for data processing are model-based. In particular, non-accelerated multiple-coil and GRAPPA acquisitions have brought noncentral-χ (nc-χ) statistics into stake as a suitable substitute for traditional Rician distributions. However, this model is only valid when the signals received by each coil are roughly uncorrelated. The recent literature on this topic suggests this is often not the case, so that nc-χ statistics are in principle not adequate. Fortunately, such model can be adapted through the definition of a set of effective parameters, namely an effective noise power (greater than the actual power of thermal noise in the RF receiver) and an effective number of coils (smaller than the actual number of RF receiving coils in the system). The implications of these artifacts in practical algorithms have not been discussed elsewhere. In the present paper we aim studying their actual impact, and suggesting practical rules to cope with them. We define the main noise parameters in this context, introducing a new expression for the effective variance of noise which is of capital importance for the two image processing problems studied: first, we propose a new method to estimate the effective variance of noise from the composite magnitude signal of MR data when correlations are assumed. Second, we adapt several model-based image denoising techniques to the correlated case, using the noise estimation techniques proposed. We show, through a number of experiments with both synthetic, phantom, and *in vivo* data, that neglecting the correlated nature of noise in multiple-coil systems implies important errors even in the simplest cases. At the same time, the proper statistical characterization of noise through effective parameters drives to improved accuracy (both qualitatively and quantitatively) for both of the problems studied.

---

1 Introduction

Noise in MR data is statistically modeled attending to the scanner coil architecture. For a single–coil acquisition, the complex spatial MR data is typically assumed to be a complex Gaussian process, where the real and imaginary parts of the original signal are corrupted with uncorrelated Gaussian noise with zero mean and equal variance. Thus, the magnitude signal is the Rician distributed envelope of the complex signal [5]. Most of the noise estimation techniques, as well as filtering methods for MRI relay on this Rician model. However, it does not hold in every case. In fact, for modern coil architectures, the Rician model is often not valid. A mismatch between the actual data and the model can affect not only visual inspection but also processing techniques such as segmentation, registration or tensor estimation in Diffusion Tensor MRI (DT-MRI) [6, 14].

If multiple–coil MR acquisition systems are considered, noise in each receiving coil in the **k**–space can be also modeled as a complex stationary Additive Gaussian Noise process, with zero mean and equal variance [36]. Assuming the noise components to be identically and independently distributed, the **k**-space fully sampled, and the composite magnitude signal (CMS) obtained using sum-of-squares (SoS) [27], the CMS will follow a non-central chi (nc-χ) distribution [27]. The noise variance will be the same for all image points in both the **k**–space and **x**–space domains, i.e. the noise in the image may be considered spatial-stationary.

The nc-χ distribution is becoming more and more common to model the behavior or signal and noise in multiple-coil systems, see for instance [60, 10] and it is also commonly used as the model for GRAPPA reconstructed data [43, 46, 42]. However, the CMS will behave as a nc-χ only if all the coils in the scanner have the same variance of noise, and there is no correlation between them. In practical cases, correlation between coils exists, and therefore the *standard* nc-χ model may not be valid.

In [49] authors showed that if multiple coils and correlated noise are considered, the data does not strictly follow a nc-χ. However, for practical purposes, it can be modeled as such, but taking into account two effects:

1. *Effective parameters* must be considered. This means that, due to the correlation, the distribution is very similar to a nc-χ but considering a smaller number of coils and a greater variance of noise.
2. The effective parameters will also depend on the signal, and hence on the position within the image. As a result, there will be different variance of noise in different areas of the image. Therefore, the pattern of noise will be spatially variant, and the noise becomes nonstationary.

Similar conclusions were raised in [42] for GRAPPA-reconstructed data. In the GRAPPA case, the effect of the noise variance varying for different image locations is increased due to the reconstruction process.

Due to both effects, methods traditionally used that rely on the nc-χ will only be valid if they are *trimmed* to the non-stationarity and effective parameters are used. Explaining how the adjustment is to be done is the main purpose of this paper. For

the sake of illustration we will focus on two of the most popular problems dealing with noise in MR: noise estimation and noise filtering. Many papers can be found about these topics in the literature along the last 5 years, see for instance [6, 43, 47, 88, 12, 91, 84].

Noise estimation is a particularly interesting topic, since most of the existing methods are based on the assumption of an homogeneous distribution of noise across the image, that can easily be measured in the background area. In this new scenario, we will first study *what to estimate* and how to do it under the restrictions posed by the model. Different cases will be studied and guidelines for estimation will be set. Although some particular estimators are chosen, the results here presented may be extended to more complex estimators, such as the ones presented in [47].

In the second part of the paper, as an illustrative example, we adapt some well-known **image denoising** techniques to the correlated case, using the noise estimation techniques described before. Many of the most popular filtering algorithms are based on the statistical characterization of noise, underlying the assumption that it can be considered as a stationary process. We show that the use of effective values and the nc-χ models will avoid the reformulation of the methods from scratch.

2 Background: Statistical Model in Correlated Multiple–Coil MR signals

In [49], authors proposed an alternative model for MR signals from multiple-coil systems. In this section, we will go over that model, its implications over the final CMS and the way to apply the main equations in filtering and noise estimation problems.

Noise in multiple coils systems, if the **k**-space is fully sampled and SoS is used to recover the CMS, is usually assumed to follow a nc-χ model [27, 46, 47, 107] with parameters L (number of coils) and σ_n^2 (variance of noise in each coil) and with Probability Density Function (PDF):

$$p_{M_L}(M_L|A_L, \sigma_n, L) = \frac{A_L^{1-L}}{\sigma_n^2} M_L^L e^{-\frac{M_L^2 + A_L^2}{2\sigma_n^2}} I_{L-1}\left(\frac{A_L M_L}{\sigma_n^2}\right) u(M_L), \quad (1)$$

with $A_L^2(\mathbf{x}) = \sum_{l=1}^{L} |A_l(\mathbf{x})|^2$ and $A_l(\mathbf{x})$ the original complex signal in each coil, $I_L(.)$ the L-th order modified Bessel function of the first kind, and $u(.)$ the Heaviside step function. In the background, this PDF simplifies to a central χ (c-χ).

The CMS will only show nc-χ statistics if the variance of noise is the same for all coils, and no correlation exists between them. However, it is well known that in phased array systems noise correlations do exist [38, 39, 40, 41] and they can seriously affect the statistical distribution of data, especially for modern machinery with a large number of receiving antennae.

In the general case, assuming L coils, the covariance matrix between them, $\mathbf{\Sigma}^2$, is an arbitrary, symmetric, positive definite matrix, where the off-diagonal elements stand for the correlations between each pair of coils. In this case, the actual PDF is not strictly a nc-χ. However, since correlations affect the number of Degrees of Freedom (DoF), in [49] it was shown that such a model is a good approximation of the actual distribution but parameters L and σ_n^2 must be replaced by their effective values:

$$L_{\text{eff}}(\mathbf{x}) = \frac{A_T^2(\mathbf{x})\,\text{tr}\left(\mathbf{\Sigma}^2\right) + \left(\text{tr}\left(\mathbf{\Sigma}^2\right)\right)^2}{\mathbf{A}^*(\mathbf{x})\mathbf{\Sigma}^2\mathbf{A}(\mathbf{x}) + ||\mathbf{\Sigma}^2||_F^2}; \tag{2}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\text{tr}\left(\mathbf{\Sigma}^2\right)}{L_{\text{eff}}(\mathbf{x})}, \tag{3}$$

with $||.||_F$ the Frobenius norm and $\mathbf{A}(\mathbf{x}) = [A_1(\mathbf{x}), A_2(\mathbf{x}), \cdots, A_L(\mathbf{x})]^T$. Similar formulation is proposed for GRAPPA reconstructed data in [42].

From these equations, some conclusions can be extracted to be taken into account in future methods:

1. The effective variance of noise will increase due to the correlations between coils.
2. The effective number of coils will be reduced. Due to the correlation, the distribution will be similar to another system with fewer coils.
3. Both effective values will depend on the position, $\mathbf{x}$. Therefore, the variance of noise will not be the same for each pixel in the image. Its distribution will be non-stationary.
4. The product $L_{\text{eff}}(\mathbf{x}) \cdot \sigma_{\text{eff}}^2(\mathbf{x})$ is a constant value, $\text{tr}\left(\mathbf{\Sigma}^2\right)$, that does not depend on $\mathbf{x}$.

For practical purposes, some simplifications can be made over these two equations. If we assume (1) that the variance of noise $\sigma_i^2 = \sigma_n^2$ is the same for every coil, and (2) that $A_i = A_j$ for all i, j, and using the following covariance matrix:

$$\mathbf{\Sigma}^2 = \sigma_n^2 \cdot \begin{pmatrix} 1 & \rho_{12}^2 & \cdots & \rho_{1L}^2 \\ \rho_{21}^2 & 1 & \cdots & \rho_{2L}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{L1}^2 & \rho_{L2}^2 & \cdots & 1 \end{pmatrix} \tag{4}$$

the effective values may be simplified to:

$$L_{\text{eff}}(\mathbf{x}) = L\left(1 + (L-1)\frac{A_T^2(\mathbf{x})\,\langle\rho^2\rangle + L\,\sigma_n^2\,\langle\rho^4\rangle}{A_T^2(\mathbf{x}) + L\,\sigma_n^2}\right)^{-1} \tag{5}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \sigma_n^2\left(1 + (L-1)\frac{A_T^2(\mathbf{x})\,\langle\rho^2\rangle + L\,\sigma_n^2\,\langle\rho^4\rangle}{A_T^2(\mathbf{x}) + L\,\sigma_n^2}\right), \tag{6}$$

with $\langle.\rangle$ the sample mean operator, so that:

$$\langle \rho^n \rangle = \frac{1}{L(L-1)} \sum_{i \neq j} \rho_{ij}^n$$

If $A_T^2(\mathbf{x})\langle \rho^2 \rangle$ is comparable to $L\,\sigma_n^2\,\langle \rho^4 \rangle$, i.e. the SNR is low, the effective parameters will depend on the position $\mathbf{x}$. Therefore the data is no longer stationary, the noise power varying along with $A_L^2(\mathbf{x})$. Two extreme cases can be considered:

1. In the background, where no signal is present and hence SNR=0, the effective values are:

$$L_{\text{eff},B} = \frac{L}{1 + \langle \rho^4 \rangle(L-1)} \tag{7}$$

$$\sigma_{\text{eff},B}^2 = \sigma_n^2\,(1 + \langle \rho^4 \rangle(L-1)). \tag{8}$$

2. For high SNR areas, say $\frac{A_T^2}{\sigma_n^2} \to \infty$:

$$L_{\text{eff},S} = \frac{L}{1 + \langle \rho^2 \rangle(L-1)} \tag{9}$$

$$\sigma_{\text{eff},S}^2 = \sigma_n^2\,(1 + \langle \rho^2 \rangle(L-1)). \tag{10}$$

These two cases give respectively the lower and upper bounds of σ_{eff}^2 within the image (vice-versa for L_{eff}).

The model here presented is far from the *standard* nc-χ generally used, and clearly very far from the Rician model. As stated before, this mismatch between model and data will render most of the existing estimation/filtering algorithms inaccurate. See eq. (8) and (10), for instance: the variance of noise in the background and in the signal areas will be different. Even considering only these two values, if the estimation of noise is done using the background (as it has been done traditionally), there will be a bias if used over the signal areas.

All in all, the model here presented poses some problems over practical noise filtering and noise estimation methods:

- The noise is non-stationary, i.e., the variance of noise will vary across the image. The existing noise estimation methods rely on the assumption that the noise is the same for the whole image.
- Even for the simplest case, the value of noise in the background of the image and within the signal areas are different.
- Although the number of coils is usually known, the only data generally available for filtering and noise estimation is the CMS, $M_L(\mathbf{x})$, so that the equations of the model cannot be directly applied. Effective values have also to be estimated.
- Most of the filtering methods in literature also assume an uniform pattern of noise across the image. When applied to multiple-coil data, they have to be properly adapted.

In the following sections we propose some practical solutions to solve these problems.

$E\{X\}$	Expectation of random variable X
σ_X^2	Variance of random variable X
$\langle \psi \rangle$	Sample mean of set ψ_i
	$\langle \psi \rangle = \frac{1}{|\Omega|} \sum_{i \in \Omega} \psi_i$
$M(\mathbf{x_B})$	Background area of image $M(\mathbf{x})$
$M(\mathbf{x_S})$	Signal area of image $M(\mathbf{x})$
$\langle M(\mathbf{x}) \rangle$	(Global) sample local mean of image $M(\mathbf{x})$
$\langle M(\mathbf{x}) \rangle_{\mathbf{x}}$	Local sample local mean of image $M(\mathbf{x})$
	$\langle M(\mathbf{x}) \rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} M(\mathbf{p})$
	($\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$)
$\text{Var}(M(\mathbf{x}))$	(Global) Sample variance of $M(\mathbf{x})$
$\text{Var}(M(\mathbf{x}))_{\mathbf{x}}$	Sample local variance of $M(\mathbf{x})$
	$\text{Var}(M(\mathbf{x}))_{\mathbf{x}} = \langle M^2(\mathbf{x}) \rangle_{\mathbf{x}} - \langle M(\mathbf{x}) \rangle_{\mathbf{x}}^2$
$\text{mode}\{I(\mathbf{x})\}$	Mode of the distribution of $I(\mathbf{x})$
	$\text{mode}\{I(\mathbf{x})\} = \arg\max_{\mathbf{x}}\{p_I(\mathbf{x})\}$
$\widehat{a}$	Estimator of parameter a

Table 1 Notation

3 Noise estimation

Most noise estimation methods for multiple-coils systems are based on the assumption that the signal follows a nc-χ distribution, and accordingly, the background area of the CMS is described by a c-χ. In [47], a set of noise estimators based on different moments of the nc-χ and c-χ were presented, all of them relying on the assumption of uncorrelated coils. According to the results summarized in the previous section, when correlations exist, the same model can be used, but effective parameters must be taken into account. The main problem that arises is that the effective values depend on the position $\mathbf{x}$, i.e., the parameters vary across the image, and the methods proposed for estimation may be no longer valid. In this section, we present a method to overcome this limitation.

3.1 Estimation of σ_{nL}^2

We will define a noise estimator based on the second order moment of the background of the image. Following the philosophies proposed in [47] other estimators may be used, like those based on the sample mean. The second order moment of the background data, where no signal is present, is defined as

$$\text{E}\{M_L^2\} = 2\sigma_n^2 L. \tag{11}$$

If effective parameters $L_{\text{eff}}(\mathbf{x})$ and $\sigma^2_{\text{eff}}(\mathbf{x})$ are taken into account we can rewrite it as

$$
\begin{aligned}
\text{E}\{M_L^2(\mathbf{x})\} &= 2\,\sigma^2_{\text{eff}}(\mathbf{x})\,L_{\text{eff}}(\mathbf{x}) \\
&= 2\,\text{tr}\left(\mathbf{\Sigma}^2\right) \\
&= \sum_{i=1}^{L}\sigma_i^2 \\
&= 2\,L\,\langle\sigma_i^2\rangle
\end{aligned}
$$

Note that although both effective parameters are $\mathbf{x}$-dependent, the product is not. We can write

$$
\begin{aligned}
L_{\text{eff}}(\mathbf{x})\cdot\sigma^2_{\text{eff}}(\mathbf{x}) &= L\cdot\langle\sigma_i^2\rangle \\
&\triangleq \sigma^2_{n\,L}
\end{aligned}
\tag{12}
$$

In the particular case in which the variance of noise is equal in each coil $\sigma_n^2 = \langle\sigma_i^2\rangle$ and

$$
\sigma^2_{n\,L} = L\cdot\sigma_n^2
$$

So, regardless of the correlation between coils, the second order moment does not depend on $\mathbf{x}$:

$$
\text{E}\{M_L^2(\mathbf{x})\} = 2\sigma^2_{n\,L}.
$$

Following the noise estimation philosophy in [8, 47] we can define a noise estimator based on the local sample estimation of the second order moment:

$$
\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|}\sum_{\mathbf{p}\in\eta(\mathbf{x})}M_L^2(\mathbf{p}),
$$

with $\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$. $\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}$ is known to follow a Gamma distribution [47, 45] whose mode is

$$
\text{mode}\left\{\langle M_L^2\rangle_{\mathbf{x}}\right\} = 2\sigma^2_{\text{eff}}(\mathbf{x})\frac{|\eta(\mathbf{x})|L_{\text{eff}}(\mathbf{x})-1}{|\eta(\mathbf{x})|}\approx 2\sigma^2_{n\,L}
$$

when $|\eta(\mathbf{x})|L_{\text{eff}} >> 1$. The estimator is then defined as

$$
\widehat{\sigma^2_{n\,L}} = \frac{1}{2}\text{mode}\left\{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\right\}
\tag{13}
$$

This estimator does not require a previous segmentation of the background, due to the mode operator. A simplest estimator may also be defined, but segmentation of the background region is needed:

$$
\widehat{\sigma^2_{n\,L}} = \frac{1}{2}\langle M_L^2(\mathbf{x_B})\rangle
\tag{14}
$$

Following the notation in [47], $M_L(\mathbf{x_B})$ are the background pixels of the CMS.

The same method can be extrapolated to GRAPPA reconstructed data. However, note that the product $L_{\text{eff}}(\mathbf{x}) \cdot \sigma^2_{\text{eff}}(\mathbf{x}$ is not a constant value [42], but

$$ L_{\text{eff}}(\mathbf{x}) \cdot \sigma^2_{\text{eff}}(\mathbf{x} = \text{tr}\left(\mathbf{C}^2_X(\mathbf{x})\right), $$

with $\mathbf{C}^2_X(\mathbf{x})$ the covariance matrix of the interpolated data for each spatial location. Thus, a different value will arise for each $\mathbf{x}$. The advantage of using the product is that $\text{tr}\left(\mathbf{C}^2_X(\mathbf{x})\right)$ does not depend on signal values, but only on the reconstruction coefficients and on the original covariance matrix. A method for GRAPPA noise estimation was proposed in [63].

3.2 Estimation of effective values

Although many methods and applications based on the nc-χ use only the $\sigma^2_{n_L}$ value (see next section), there are other situations in which the effective values of noise and number of coils are needed. Note that the effective values will now be $\mathbf{x}$-dependent.

According to eqs. (2) and (3), the effective values depend on the actual value of the signal in each coil, and on the covariance matrix. Even for the most simplified versions of the formulae, eqs. (5) and (6), the values will still depend on the correlation coefficients between coils. For the estimation we will consider that the only data available is the CMS $M_L(\mathbf{x})$ and the number of coils L.

To estimate the effective values over the background, we will use the variance of the c-χ distribution:

$$ \sigma^2_{M_L} = \sigma^2_n \left(2L - 2 \left[\frac{\Gamma^2\left(L + \frac{1}{2}\right)}{\Gamma^2(L)} \right] \right) $$

Let $\widehat{\sigma^2_{n,B}}$ be the estimated variance in the background area. We can simply estimate it as the sample variance of the segmented background:

$$ \widehat{\sigma^2_{n,B}} = \text{Var}(M_L(\mathbf{x_B})). \tag{15} $$

To avoid the segmentation of the background, we can use the mode of the histogram of the variance:

$$ \widehat{\sigma^2_{n,B}} = \text{mode}\left\{\text{Var}(M_L(\mathbf{x_B}))_{\mathbf{x}}\right\} \tag{16} $$

For more detail the signal and background distributions of the variance in MR data, check [107].

Regardless the method chosen, we assume that the effective values are constant through the background and thus we can estimate them through the following iterative process:

$$\left(\sigma^2_{\text{eff},B}\right)_{i+1} = \frac{\widehat{\sigma^2_{n,B}}}{2\left(L_{\text{eff},B}\right)_i - 2\left[\dfrac{\Gamma^2\left(\left(L_{\text{eff},B}\right)_i+\frac{1}{2}\right)}{\Gamma^2\left(\left(L_{\text{eff},B}\right)_i\right)}\right]} \tag{17}$$

$$\left(L_{\text{eff},B}\right)_{i+1} = \frac{\widehat{\sigma^2_{nL}}}{\left(\sigma^2_{\text{eff},B}\right)_{i+1}} \tag{18}$$

with $\left(L_{\text{eff},B}\right)_0 = L$ the initialization value. The term $\widehat{\sigma^2_{nL}}$ is estimated as indicated in the previous section.

The effective values when SNR$\to \infty$ can be calculated over the signal area, where the variance is defined

$$\sigma^2_{M_L} = \sigma^2_n\left(1 + O\left(\frac{1}{M_L^4}\right)\right) \approx \sigma^2_n$$

Then, we can define

$$\widehat{\sigma^2}_{\text{eff},S} = \text{mode}\left\{\text{Var}(M_L(\mathbf{x_S}))_{\mathbf{x}}\right\} \tag{19}$$

$$L_{\text{eff},S} = \frac{\widehat{\sigma^2_{nL}}}{\widehat{\sigma^2}_{\text{eff},S}} \tag{20}$$

with $\text{Var}(M_L(\mathbf{x_S}))_{\mathbf{x}}$ the sample local variance of the signal area of the image.

The effective values $\sigma^2_{\text{eff},B}$ and $\sigma^2_{\text{eff},S}$ give the lower and upper bounds of the actual $\sigma_{\text{eff}}(\mathbf{x})^2$ across the image. Using the simplified version of the effective variance of noise in eq. (10):

$$\sigma^2_{\text{eff}}(\mathbf{x}) = \sigma^2_n\left(1 + (L-1)\frac{A_T(\mathbf{x})^2\langle\rho^2\rangle + L\,\sigma^2_n\langle\rho^4\rangle}{A_T(\mathbf{x})^2 + L\,\sigma^2_n}\right)$$

$$= \sigma^2_n\left(1 + (L-1)(1-\phi_n(\mathbf{x}))\langle\rho^2\rangle + (L-1)\phi_n(\mathbf{x})\langle\rho^4\rangle\right)$$

with

$$\phi_n(\mathbf{x}) = \frac{L\sigma^2_n}{A_T^2(\mathbf{x}) + L\sigma^2_n} = \frac{1}{\text{SNR}^2(\mathbf{x}) + 1}, \tag{21}$$

and $\text{SNR}^2(\mathbf{x}) = \frac{A_T^2(\mathbf{x})}{L\sigma^2_n}$. After some algebra we can write

$$\sigma^2_{\text{eff}}(\mathbf{x}) = (1 - \phi_n(\mathbf{x}))\cdot\sigma^2_{\text{eff},S} + \phi_n(\mathbf{x})\cdot\sigma^2_{\text{eff},B} \tag{22}$$

Since $0 \leq \phi_n(\mathbf{x}) \leq 1$, then $\sigma^2_{\text{eff}}(\mathbf{x}) \in [\sigma^2_{\text{eff},B}, \sigma^2_{\text{eff},S}]$. A rough estimation of $\phi_n(\mathbf{x})$ can be done using the sample second order moment (although more complex estimation could also be done):

$$\widehat{\phi_n}(\mathbf{x}) = \frac{\widehat{\sigma_{nL}^2}}{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - \widehat{\sigma_{nL}^2}} \tag{23}$$

As a final result, the effective values across the image may be estimated as:

$$\widehat{\sigma^2}_{\text{eff}}(\mathbf{x}) = (1 - \widehat{\phi_n}(\mathbf{x})) \cdot \widehat{\sigma^2}_{\text{eff},S} + \widehat{\phi_n}(\mathbf{x}) \cdot \widehat{\sigma^2}_{\text{eff},B} \tag{24}$$

$$\widehat{L}_{\text{eff}}(\mathbf{x}) = \frac{\widehat{\sigma_{nL}^2}}{\widehat{\sigma^2}_{\text{eff}}(\mathbf{x})} \tag{25}$$

4 Noise filtering

As a practical example of the utility of the different noise parameters estimated in the previous section, some noise-reduction filters for MR are considered. Many of the schemes proposed in the last decade are based on the assumption of an underlying Rician distribution. Lately, some attempts to extend these methods to parallel imaging and multiple coils have been made. Regardless of the filter or method, since most of them are based on a signal and noise statistical model, they rely on the accurate estimation of some noise-related parameters.

On the other hand, it will not be necessary to totally reformulate the filtering techniques in order to cope with correlation between coils. By using *effective values* for number of coils and variance of noise, we are implicitly *decorrelating* the noise between coils, and therefore the nc-χ model can be used directly.

4.1 Conventional Approach and NLM

The extension of some noise filtering methods for MR data from Rician to a nc-χ model is straightforward. As shown before, if this extension takes into account the effective values for noise and number of coils, it may be valid also when correlation between coils exists. The simplest model based method for Rician noise was proposed by *McGibney et al.* in [4]. The so-called *conventional approach* (CA) is basically an averaging of the squared signal with bias removal, assuming a Rician distribution of the data. The extension to nc-χ is:

$$\widehat{A_L}(\mathbf{x}) = \sqrt{\max(\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - 2\sigma_{nL}^2, 0)} \tag{26}$$

Since the parameter σ_{nL}^2 does not depend on the position, it can be estimated using eq. (13).

Despite the simplicity of the method, the *philosophy* of bias removal over the second order moment has been applied to more complex filters, such as nonlocal means (NLM) schemes [88, 140, 13]. If $\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}$ is replaced by the NLM of $M_L^2(\mathbf{x})$, the

result will be a NLM filter for nc-χ data, similar to the unbiased schemes proposed in [12, 130, 11].

$$\widehat{A_L}(\mathbf{x}) = \sqrt{\max(\text{NLM}(M_L^2(\mathbf{x})) - 2\sigma_{nL}^2, 0)} \qquad (27)$$

A similar philosophy for other filtering schemes may be found in [91, 73, 141].

The advantage of working with the second order moment is that the noise does not depend on the position within the image. As long as we work with σ_{nL}^2, the scenario is alike the non-correlated case.

What is more, if the second order moment is used for noise estimation –as in eq. (13). and eq.(14)–, note that $E\{M_L^2(\mathbf{x_B})\}$ (the second order moment of the background) follows the identity:

$$E\{M_L^2(\mathbf{x_B})\} = \begin{cases} 2\sigma_n^2 & \text{Rayleigh data} \\ 2\sigma_n^2 L & \text{c-}\chi \text{ data} \\ 2\sigma_{\text{eff}}^2 L_{\text{eff}} = 2\,\text{tr}\left(\mathbf{\Sigma}^2\right) & \text{Correlated c-}\chi \text{ data} \end{cases}$$

Whatever the case, Rician, correlated or no-correlated nc-χ, the signal estimator will be exactly the same:

$$\widehat{A_L}(\mathbf{x}) = \sqrt{E\{M_L^2(\mathbf{x})\} - E\{M_L^2(\mathbf{x_B})\}} \qquad (28)$$

and eq. (26) and eq. (27) are therefore practical implementations of eq. (28). The main advantage of this feature is that all the software and methods relying on eq. (28), initially designed for Rician noise, will also be valid for other kinds of noise. As a final recall, this is only valid if the second order moment is used for noise estimation. If any other method is used, the estimator may differ for different kinds of noise.

4.2 LMMSE estimator

A Linear Minimum Mean Square Error (LMMSE) estimator for the Rician model was proposed in [6, 8]. This filter has been extensively used due to its simplicity, speed, and robustness, see for instance [142, 143, 144]. Some modifications may also be found in [62, 9]. The extension to uncorrelated nc-χ data is straightforward, as presented in [10]:

$$\widehat{A_L^2(\mathbf{x})} = \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - 2L\sigma_n^2 + K_L(\mathbf{x})\left(M_L^2(\mathbf{x}) - \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\right), \qquad (29)$$

where $K_L(\mathbf{x})$ is defined as

$$K_L(\mathbf{x}) = 1 - \frac{4\sigma_n^2\left(\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - L\sigma_n^2\right)}{\langle M_L^4(\mathbf{x})\rangle_{\mathbf{x}} - \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}^2}. \qquad (30)$$

Unlike the filters in the previous section, the one here defined depends not only on $L\sigma_n^2$, but also on σ_n^2. If no correlation between coils exists, we can easily estimate this parameter as $\sigma_n^2 = \sigma_{nL}^2/L$. However, once correlations are present, this value would underestimate the actual level of noise. Thus, in order to cope with the correlation, effective parameters must be used:
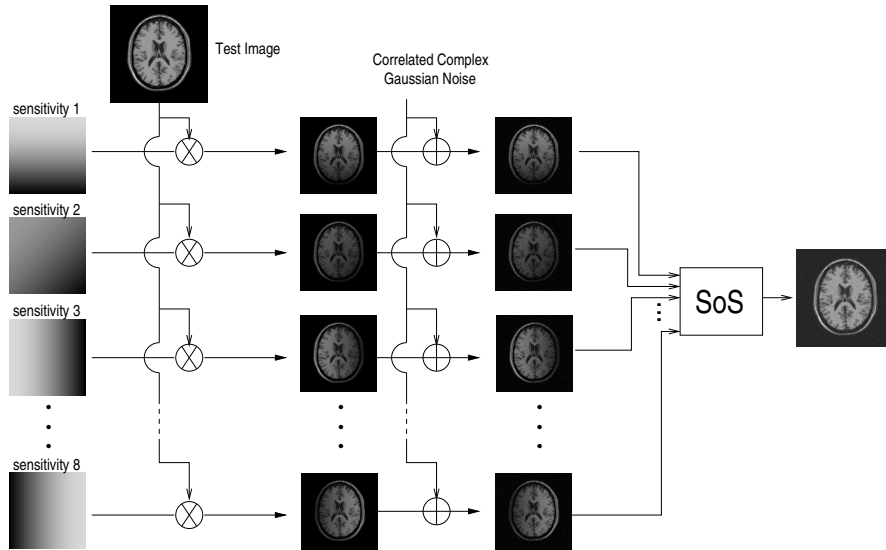
$$\widehat{A_L^2(\mathbf{x})} = \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - 2\sigma_{nL}^2 + K_L(\mathbf{x})\left(M_L^2(\mathbf{x}) - \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\right), \qquad (31)$$

$$K_L(\mathbf{x}) = 1 - \frac{4\sigma_{\mathrm{eff}}^2(\mathbf{x})\left(\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - \sigma_{nL}^2\right)}{\langle M_L^4(\mathbf{x})\rangle_{\mathbf{x}} - \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}^2}. \qquad (32)$$
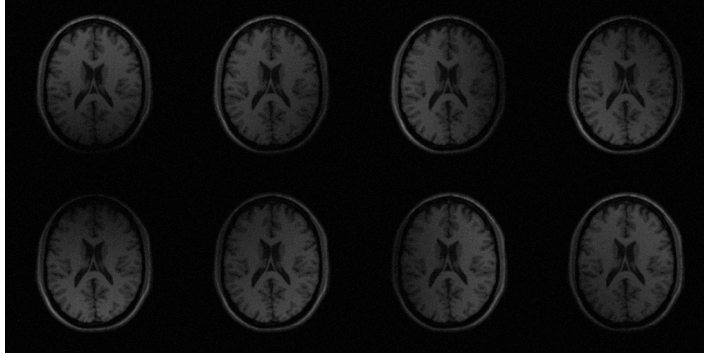
In this case, not only σ_{nL}^2 but $\sigma_{\mathrm{eff}}^2(\mathbf{x})$ has to be estimated. Using eq. (24) $K_L(\mathbf{x})$ becomes:

$$K_L(\mathbf{x}) = 1 - 4\frac{\left(\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}} - 2\sigma_{nL}^2\right)\sigma_{\mathrm{eff},S}^2 + \sigma_{nL}^2\sigma_{\mathrm{eff},B}^2}{\langle M_L^4(\mathbf{x})\rangle_{\mathbf{x}} - \langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}^2}.$$

5 Materials and methods

Fig. 1 Synthetic phantom generation. From a single slice from BrainWeb, the sensitivity maps are simulated and correlated noise is added to every coil. The CMS is obtained by SoS.

Fig. 2 Synthetic MR image from Brainweb used for the noise estimation experiments. An 8-coil system is simulated.

For the sake of validation, the following experiments are considered:

1. Noise estimation: To test the noise estimation methods proposed, synthetic experiments are considered. A synthetic phantom is built using different levels of noise and different correlations between coils, following the scheme in Fig. 1. The starting point is a 2D synthetic slice $S_0(\mathbf{x})$ from a BrainWeb MR volume [53], with intensity values in $[0 - 255]$. The average intensity value for the White Matter is 158, for the Gray Matter is 105, for the cerebrospinal fluid 36 and 0 for the background. An 8-coil system is simulated using an artificial sensitivity map coded for each coil so that $A_L^2(\mathbf{x}) = \sum_l |A_l|^2 = S_0^2(\mathbf{x})$. The signal in each coil is corrupted with complex Gaussian noise with variance σ_n^2 (both in the real and imaginary parts) ranging in $[5 - 30]$ and a correlation coefficient between coils ρ^2 ranging in $[0.01, 0.4]$. The CMS is reconstructed from the noisy data in each coil using SoS. 100 realizations are done for each pair $[\sigma_n^2, \rho^2]$.

From the phantom, assuming we just know the CMS $M_L(\mathbf{x})$ and the number of coils L, the following parameters are estimated:

1. $\widehat{\sigma^2}_{nL}$ using both eq. (13) and eq. (14).
2. Effective noise in the background: $\widehat{\sigma^2}_{\text{eff},B}$ using eq. (17).
3. Effective noise in the signal areas, $\widehat{\sigma^2}_{\text{eff},S}$ using eq. (19).
4. Global value $\widehat{\sigma^2}_{\text{eff}}(\mathbf{x})$ from eq. (24).

All the sample local moments are calculated using 7×7 neighborhoods. For the first three cases, in which the estimated values are scalar, the quality measure is defined as the ratio:

$$Q = \frac{\text{estimated value}}{\text{theoretical value}}.$$

Accordingly, the closer to one, the better the estimator. For the last case, since $\widehat{\sigma^2}_{\text{eff}}(\mathbf{x})$ is an image the average of the quality measure for every $\mathbf{x}$ is considered. Finally, for each pair $[\sigma_n^2, \rho^2]$ the average of the 100 experiments is considered.

To test the influence of the assumption $A_i = A_j$ used to define the estimators, the previous experiment is repeated, but considering *uniform* sensitivity maps. This way, the signal in each coil (before adding noise) is the same. Results are compared with the previous experiment.

2. Noise filtering: In the second experiment, the influence of the correlation over filtering will be studied. To that end, the phantom generated in the first experiment is used. We will consider different values for the original variance of noise, σ_n^2, and ρ^2 ranging in $[0.01, 0.3]$. In order to study only the error due to the filtering model, for the experiments the actual values of the effective parameters will be used, calculated from the complex data. First, we will compare the following LMMSE estimators:

1. cχ-LMMSE: The nc-χ version of the estimator assuming correlations between coils, proposed in eq. (31).
2. χ-LMMSE: The nc-χ version without correlation, proposed in [10]. For this case we will consider $\sigma_n^2 = \sigma_{n_L}^2 / L$.
3. LMMSE: The Rician version proposed in [6, 8], assuming $\sigma_n^2 = \sigma_{n_L}^2$. This assumption is based on the estimation of σ_n^2. If Rician data is assumed (rather that nc-χ) and the estimation of σ_n^2 is done using the second order moment, what we are really estimating is $\sigma_{n_L}^2$, see [107].

Secondly, the different nc-χ based filters proposed in Section 4 are compared among them for different σ_n and ρ^2 parameters:

1. cχ-LMMSE: The nc-χ LMMSE proposed in eq. (31).
2. NLM: The NLM in eq. (27). The fast implementation proposed in [13] is used.
3. CA: The conventional approach in eq. (26).
4. Wavelet: In order to compare with a Wavelet-based filter, the implementation proposed in [115] by Pižurica is used.

To compare the restoration performance of the different methods, two quality indexes are used: the Structural Similarity (SSIM) index [126] and the Quality Index based on Local Variance (QILV) [127]. Both give a measure of the structural similarity between the ground truth and the estimated images. However, the former is more sensitive to the level of noise in the image and the latter to any possible blurring of the edges. Both indexes are bounded; the closer to one, the better the image. We will use both as a measure of the noise *removed* from the image. In addition, the QILV will also act as an indicator of the edge blurring: due to its nature, any minimal smoothing of the edges will cause the index to drop down. Finally, the square root of the mean square error (sMSE) is also calculated. These three quality measures are only being applied to those areas of the original image greater than zero; this way the background is not taken into account when evaluating the quality of each method. The average of 15 experiments is considered for each σ_n, each ρ^2 and each filter.

3. *In vivo* data: As a final experiment, we provide an illustrative in vivo example with real data. 20 identical and independent repetitions of a DWI data set with one baseline image and 15 gradient directions were successively taken from a healthy
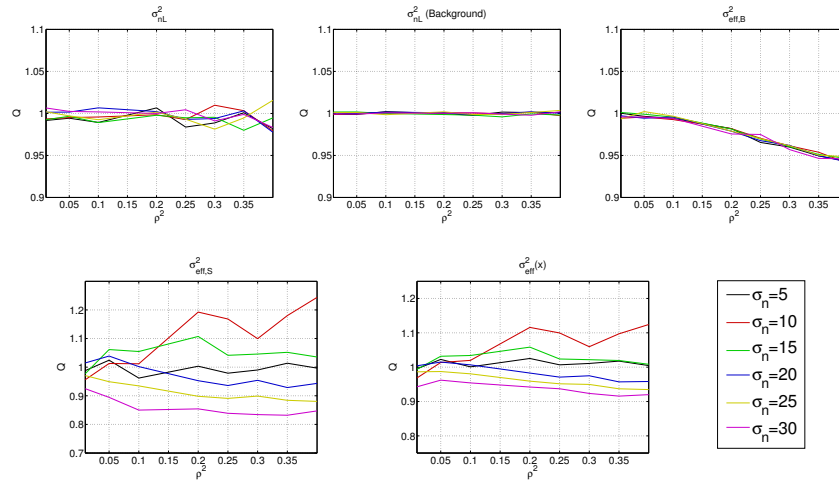
Fig. 3 Estimation of noise-related parameters for the synthetic phantom. Variable sensitivity maps are used.

volunteer. To keep a reasonable acquisition time, one axial slice was collected in a 3T, 8-coil (without down-sampling) Philips scanner using a diffusion parameter $b = 1,200s/mm^2$. To obtain a *silver standard*, we have averaged the 20 repetitions in each of the eight coils, so that we obtain eight *clean* channels which are further combined using SoS to compound the final signal[2]. This averaged image is compared to one realization and the different filtering techniques. In addition, visual comparison will also be used.

6 Results

Results for the **noise estimation** experiments with the synthetic phantom are on fig. 3 (variable sensitivity maps) and fig. 4 (homogeneous sensitivity maps). In both cases, the estimators proposed for σ_{nL}^2 show a very good behavior. The one based on the segmented background, as expected, is the one with the more consistent estimation. However, the one based on the mode, shows also a very good performance and coherence for multiple σ_n and ρ^2 values. In addition, the latter does not require a prior segmentation of the background.

An accurate estimation of σ_{nL}^2 is the keystone for a good estimation of the effective parameters, if needed, and assures and unbiased filtering for the methods proposed in the paper. In addition, since σ_{nL}^2 is constant across the image and re-

---

[2] Due to phase-shifts between the different realizations, the average has had been done over the Rician envelope of each coil, using a Rician-MLE scheme [82]
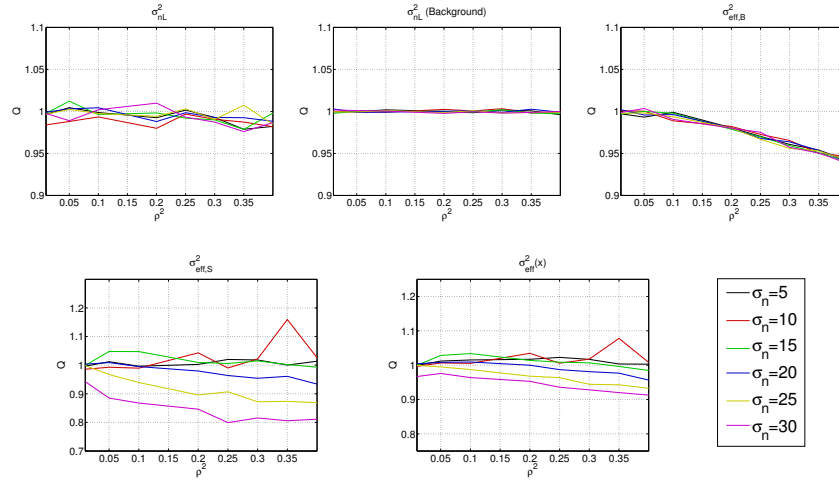
Fig. 4 Estimation of noise-related parameters for the synthetic phantom. Uniform sensitivity maps are used.

gardless the correlation level ρ^2, the estimation can be properly done for a wide range of values.

The estimated effective value of noise for the different cases, shows also a good behavior, but with greater error. First, for $\sigma^2_{\text{eff},B}$ (variance of noise in the background) there is an underestimation as ρ^2 grows. The same underestimation is done for all the σ_n. This is due to the mismatch between the real probabilistic model of the data and the nc-χ assumed. As shown in [49], as ρ^2 grows, also does the error in the model.

On the other hand, the estimation of $\sigma^2_{\text{eff},S}$ (noise in the signal areas) worsens as the noise grows. There is also a slight decay with ρ^2, but less noticeable. The effect this time is related to the large SNR assumption: as σ^2_n grows, the assumption differs from reality. In addition, to estimate this parameter, the assumption $A_i = A_j$ was also made. The error shown in fig. 3, where the signals are different form coil to coil is also larger than the coil in fig. 4. According to this, there is also a source of error introduced by this assumption.

Finally, the global error when considering $\sigma^2_{\text{eff}}(\mathbf{x})$ lies within acceptable values in both cases. This values will be enough for filtering algorithms. For the sake of illustration, one particular case of $\sigma^2_{\text{eff}}(\mathbf{x})$ is shown in fig.5.

Quantitative results of the **filtering experiment** with synthetic data are on fig. 6 and fig. 7. For visual comparison check fig. 8. The first comparison is done for the LMMSE-based schemes, with two different levels of noise. In both cases, the SSIM index is better for the Rician case. This can be explained by the nature of the index itself: the SSIM index measures the amount of noise removed, but it does not take into account any blurring of the edges. According to the theoretical model, the Rician LMMSE uses a overestimated level of noise, and hence over-smooths
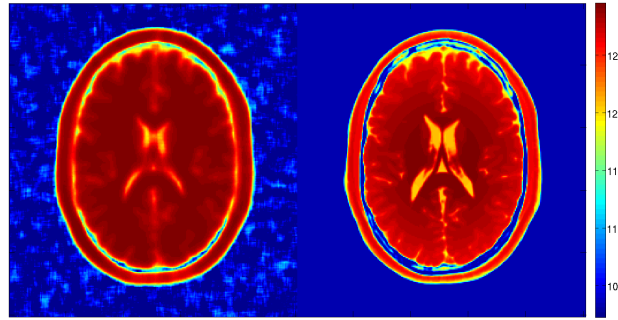
Fig. 5 Estimated effective standard deviation of noise and actual value (estimated from simulated values). $\sigma_n = 10$ and $\rho^2 = 0.1$.
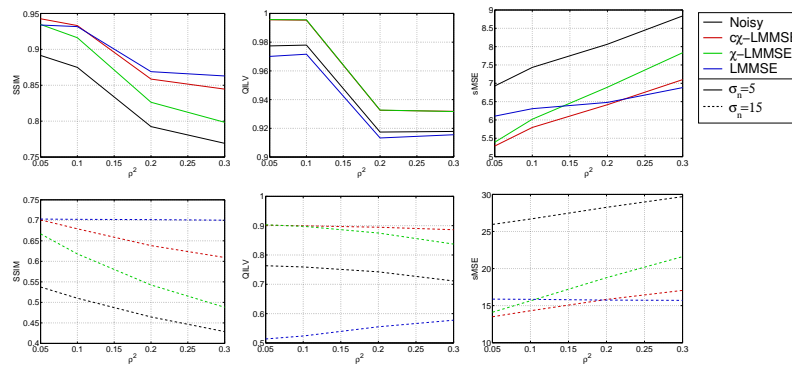

**Fig. 6** Quality measures for the LMMSE-based schemes filtering experiment.

the image. Drawing our attention to the QILV index, which is sensitive to blurring of the edges, the value drops for the Rician LMMSE. The blurring can be seen in the example in Fig. 8-(g). On the other hand, the χ-LMMSE (without considering correlations between coils) implicitly uses an underestimated value of noise. As a result, its filtering capability is smaller than the cχ-LMMSE. This can be seen in smaller SSIM values for the χ-LMMSE, smaller or equal values of QILV and greater sMSE. Visual inspection of Fig. 8-(e) and 8-(f) shows that the signal area of the image is noisier for χ-LMMSE.

Secondly, different filtering schemes are compared in Fig. 7, also for the synthetic data. In all the cases, the NLM proposed and the cχ-LMMSE are the ones showing a greater noise reduction (SSIM values) while keeping the edges and structures (QILV values). Visually, the NLM scheme is the one showing the better results for these experiments. The advantage of the cχ-LMMSE is that it is computationally simpler than NLM (though the implementation proposed in [13] considerably improves the performance and speed of the method).

Finally, *in vivo* results are depicted in Fig. 9. Only 4 DWIs of the 15 are shown for illustration. The different methods perform as with synthetic data. Wavelet based
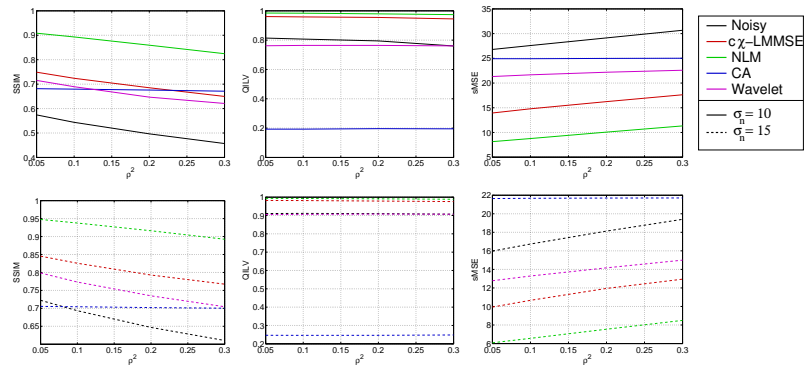
Fig. 7 Quality measures for different noise filtering schemes experiment.

(a) Original   (b) Noisy   (c) CA   (d) Wavelet

(e) cχ-LMMSE   (f) χ-LMMSE   (g) LMMSE   (h) NLM
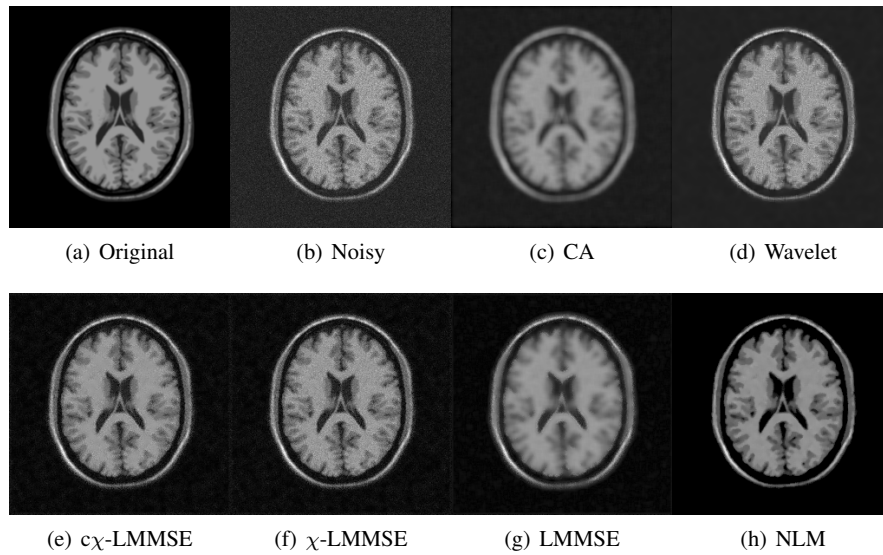
Fig. 8 Visual comparison: Synthetic slice from BrainWeb with different filtering methods. Parameters: 8 coils, $\sigma_n = 15$, $\rho = 0.15$.

filter, as optimized for Rician data, overfilters the image, blurring structures. The cχ-LMMSE scheme removes noise keeping structures, although there is a slight blur in the DWIs when compared to the averaged image. The χ-LMMSE, as expected, does not filter the image, since it is not considering effective values for the variance of noise. Finally, the NLM scheme removes all the noise, but the resulting data looks a bit unrealistic (mainly in the DWI).
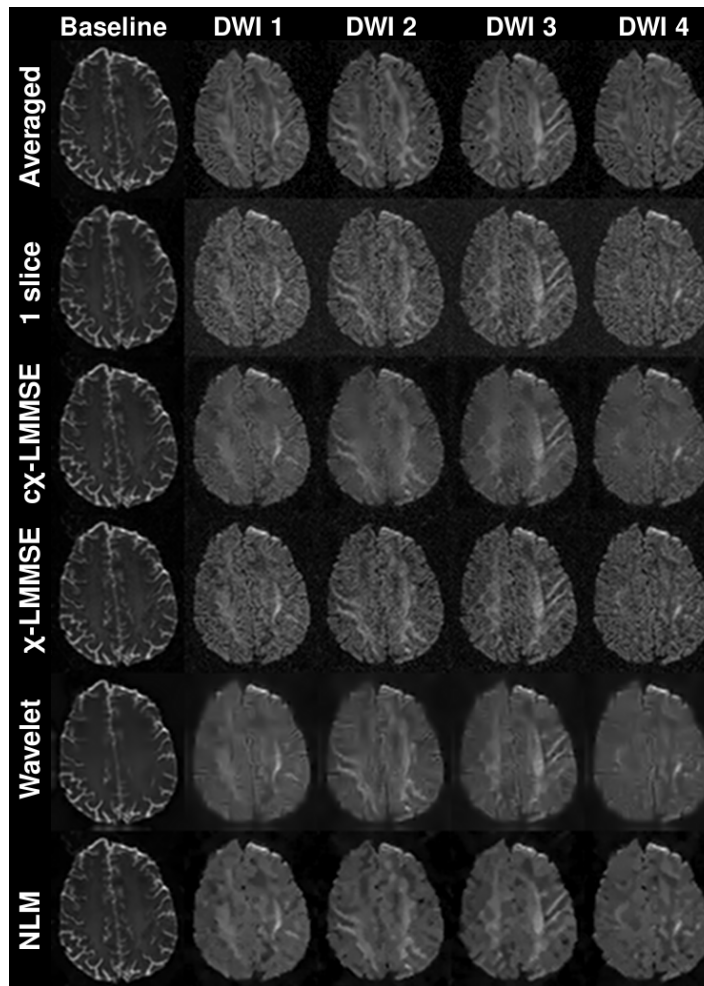
Fig. 9 Visual comparison: Filtering of baseline and 4 DWI from DWI data. Real data from a 3T Philips scanner with 8 coils.

7 Discussion and Conclusions

The practical implications of the correlations between coils in multiple-coil systems have been studied. It is well known that in phased array systems noise correlations do exist between coils [38, 39, 40, 41]. The effects of these noise correlations have been usually left aside, justified by a minimal *visual* effect over the final image. However, currently, the use of MRI data is not reduced to visual inspection. Tensor estimation, filtering, segmentation, feature extraction, and many automatic and semi-automatic procedures rely on a well-defined statistical model of the signal and noise. It is in these cases that correlations may seriously effect the data, since they

may alter the underlying model. As previously shown in [49], the correlations between coils modify the data statistics, so that the final CMS is similar to another one with a reduced number of coils and a greater variance of noise. This modification will be especially significant for modern machinery with a large number of receiving antennae.

From a practical point of view, the effect of the correlations between coils in multiple-coil systems makes the probabilistic model for the CMS differ from a nc-χ model, which is the distribution usually assumed for multiple–coil data. Luckily, a nc-χ model can still be assumed with a small error if effective parameters are considered for the variance of noise and the number of coils. Thus, main processing techniques, such as filtering or noise estimation, that are valid for a nc-χ model, will still be valid in correlated systems, as long as they are reformulated to cope with effective values. In this paper, that *reformulation* has been done for two practical cases: noise estimation and model-based noise filtering.

Once multiple coils and correlations are considered, it is not obvious which is the *level of noise* of the image. As a result of the correlation (or the reconstruction procedure if GRAPPA is considered), the final noise in the CMS is non-stationary, i.e., the variance of noise will vary across the image. Even for the simplest case, the effective value of noise in the background of the image and within the signal areas will be different. It would imply an augmentation of the complexity of many model based algorithms. However, the main parameter to have into account is no longer the original variance of noise σ_n^2 of its effective version $\widehat{\sigma^2}_{\text{eff}}(\mathbf{x})$, but σ_{nL}^2. This parameter is constant through the image, even when the effective variance of noise is not. This feature constitutes an advantage that should be used when designing processing techniques. As previously shown in the experiments, σ_{nL}^2, for being a single value for the whole image, is an easy parameter to estimate, unlike the effective parameters, whose estimation is more convoluted and less precise.

Using noise-filtering as illustration, we have also shown that when working with model-based methods, it is necessary to adapt this methods to the new model of noise. The correlated case generalizes the non-correlated one. For example, a proper selection of the magnitude being estimated allows to use the well-known NLM scheme *as it is* through eq. (28), without any further modification that increases its complexity.

In addition, experimental results show that in practical cases, such as filtering, a proper model will improve the performance. As an illustration, we have chosen the LMMSE scheme, for being an estimator totally based on the underlying model of signal and noise. Thus, it is particularly sensitive to mismatches between the model and the data. In the experiments carried out, we can see that assuming Rician data will implicitly mean overestimating the noise. This results in over-smoothed data, blurring of the edges, and spoiling of fine structures. On the other hand, if the effective values are not used, underestimation of the variance arises even when the nc-χ model is assumed. As a result, the method has a poor denoising performance.

Finally, we want to recall that the study here done is based on a SoS reconstruction of the CMS. Although it is initially valid for correlated multiple-coil fully-sampled data, it could be easily extended to GRAPPA, assuming a new covariance

matrix derived from the reconstruction process. The main difference when GRAPPA is considered is that the product $L_{\text{eff}}(\mathbf{x}) \cdot \sigma^2_{\text{eff}}(\mathbf{x})$ is not a constant, but a different value for each point in the image, and therefore it must be estimated using an *ad-hoc* method. Furthermore, if other kind of reconstructions different from SoS are used, the whole study must be redone, although the main implications remain. The study is also not valid for all those parallel methods that carry out the reconstruction in the $\mathbf{x}$-space (like SENSE), since the nc-χ distribution is not guaranteed.

In addition, the filters and noise estimators proposed in this paper are not intended to be ultimate. Our purpose here was to show how proper corrections must be done over well known algorithms when dealing with correlations between coils, and the advantage in doing so. Many other filtering methods can be proposed, and surely they will overcome the ones in this paper. More accurate noise estimation methods for the effective variance of noise can also be developed, based for example on those proposed in [47]. However, our aim was mainly to clarify what we are talking about when we say *noise* in this framework, and how to deal with correlated multiple coil systems from a signal processing point of view.

Acknowledgements

About the background distribution in MR data: a local variance study

Santiago Aja-Fernández, Gonzalo Vegas-Sánchez-Ferrero, Antonio Tristán-Vega*

Abstract A model for the distribution of the sample local variance of Magnetic Resonance data is proposed. It is based on a bimodal Gamma distribution, whose maxima are related to the signal and background areas of the image. The model is valid for single– and multiple–coil systems. The proposed distribution allows us to characterize some signal/background properties in MR data. As an example, the model is used to study the effect of the background size over noise estimation techniques and a method to test the validity of background-based noise estimators is presented.

1 Introduction

Noise is known to be one of the main sources of quality deterioration in Magnetic Resonance (MR) data. Not only visual inspection but also processing techniques such as segmentation, registration or tensor estimation in Diffusion Tensor MRI (DT-MRI) will be affected or biased due to the presence of noise [4, 5, 6]. In High Angular Resolution Diffusion Imaging (HARDI) for instance, to achieve a speedup in the acquisition time the temporal averaging is reduced; as a consequence, the noise power is increased proportionally to the square root of the speedup.

Noise in the **k**-space in MR data is usually assumed to be a zero-mean uncorrelated complex Gaussian process in each scanner coil, with equal variance in both the real and imaginary parts. As a result, in single coil systems, magnitude data in the spatial domain is modeled using a Rician distribution [44, 5], which reduces to a Rayleigh in the background. In the same way, the composite signal in coils systems

---

with multiple channels under some assumptions may be modeled as non–central Chi distributed [27, 46]. In the background, this distribution reduces to a central Chi.

The characterization of background and signal areas in MR data is an important source of information that has been used in a myriad of applications, such as segmentation [138], noise estimation [37, 47, 84], quantification of geometric distortions in tissues [145] and nonstationarity tests of the data [146].

In the current paper, a study of the relation between signal areas and background is carried out. It is based on the properties of the distribution of the sample local variance (SLV) in both regions, since it keeps important information of the structural content in the image. The distribution of the SLV has been previously used in the image processing and analysis field as a way to study the image structure [147], to estimate the level of noise [148, 108, 8], to predict spatial patterns of textures [149] or as the keystone for image quality assessment methods [127].

The particular features of signal and noise in MR data make the SLV a powerful source of information about noise, background and structure. A study of the properties of the SLV in MR data is carried out in the paper for both single– and multiple–coil systems. The probability distribution of the SLV is well known for certain kind of data [122, 150]; the SLV of Gaussian Random variables behaves like a Gamma distribution (or Chi-Square for normalized variables). However, for other parent distributions, this PDF can only be approximated [151, 152, 150, 153]. A preliminary study of SLV for single–coil MR data can be found in [8]. In this paper, a simplified model based on a bimodal Gamma distribution is proposed for MR data.

As a practical application of the proposed approach, a method to test the validity of noise estimation techniques is presented. Many different ways to accurately estimate the noise power from MR data have been proposed in literature. Most of them rely on a statistical model for the signal/background areas. The assumption of a homogeneous noisy background is the keystone of these methods, and also one of their main drawbacks: the models need a certain amount of background pixels to perform the estimation. Conversely, new scanning techniques and software eliminate the most part of the noisy background, drastically reducing the number of points available for estimation. Thus, many estimation techniques may be no longer valid. The method here proposed, based on the properties of the distribution of the previously introduced SLV, will check whether the number of available background pixels is enough to carry out the estimation with a particular method.

2 Theory

2.1 Statistical noise model in single and multiple coil MR signal

For a single–coil acquisition the complex spatial MR data is typically modeled as a complex Gaussian process, where the real and imaginary parts of the original signal

are corrupted with uncorrelated Gaussian noise with zero mean and equal variance σ_n^2. The magnitude signal $M(\mathbf{x})$ is the Rician distributed envelope of the complex signal [5] with probability density function (PDF) [44]:

$$p_M(M|A,\sigma_n) = \frac{M}{\sigma_n^2} \; e^{-\frac{M^2+A^2}{2\sigma_n^2}} \; I_0\left(\frac{AM}{\sigma_n^2}\right) \; u(M), \tag{1}$$

with $A = A(\mathbf{x})$ the original signal if no noise is present, $I_0(.)$ the 0th order modified Bessel function of the first kind and $u(.)$ the Heaviside step function. For high Signal to Noise Ratio (SNR), the Rician distribution approaches a Gaussian distribution.

In the image background, where the SNR is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution with PDF

$$p_M(M|\sigma_n) = \frac{M}{\sigma_n^2} \; e^{-\frac{M^2}{2\sigma_n^2}} \; u(M). \tag{2}$$

In a multiple–coil MR acquisition system the acquired signal in the complex spatial domain in each coil may be also modeled as the original signal corrupted with complex additive Gaussian noise, with zero mean and equal variance σ_n^2. If no subsampling is done in the $\mathbf{k}$-space, the composite magnitude image may be obtained using methods such as the sum-of-squares (SoS) [27, 38]. Assuming the noise components to be independent and identically distributed (IID), the envelope of the magnitude signal $M_L(\mathbf{x})$ will follow a non–central Chi distribution with PDF [27]:

$$p_{M_L}(M_L|A_L,\sigma_n,L) = \frac{A_L^{1-L}}{\sigma_n^2} M_L^L e^{-\frac{M_L^2+A_L^2}{2\sigma_n^2}} I_{L-1}\left(\frac{A_L M_L}{\sigma_n^2}\right) u(M_L), \tag{3}$$

with L the number of coils and $A_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |A_l(\mathbf{x})|^2}$ the original signal. Eq. (3) reduces to the Rician distribution for $L = 1$. In the background, this PDF simplifies to a central Chi distribution with PDF:

$$p_{M_L}(M_L|\sigma_n,L) = \frac{2^{1-L}}{\Gamma(L)} \frac{M_L^{2L-1}}{\sigma_n^{2L}} e^{-\frac{M_L^2}{2\sigma_n^2}} u(M_L), \tag{4}$$

which reduces to Rayleigh for $L = 1$.

The latter statistical model is the usual one for the composite magnitude signal in phased array coils and parallel imaging assuming that no subsampling is done in the $\mathbf{k}$-space, and the image is reconstructed using the SoS method. However, one of the aims of parallel imaging is precisely to accelerate the acquisition process by sub–sampling $\mathbf{k}$-space data in each coil. In these cases, reconstruction methods have to be used in order to suppress the aliasing and underlying artifacts created by the subsampling. Dominant among such methods are SENSE [32] and GRAPPA [33], reviews of which can be found in [30, 31]. From a statistical point of view, the reconstruction will affect the stationarity of the noise in the reconstructed data, i.e. the spatial distribution of the noise across the image [43]. As a result the statistics of the composite magnitude signal are not strictly stationary. In addition, when re-

constructed with GRAPPA, σ_n^2 may also vary from coil to coil and therefore eq. (3) does not exactly hold, since it assumes that σ_n^2 is the same for every coil. In practical situations, if the variance of noise is homogeneous enough across pixels and coils, data are usually considered to follow a non-central Chi distribution if reconstructed with GRAPPA and SoS [46].

2.2 A model for the sample local variance of MR data

The (biased) SLV of an image $M(\mathbf{x})$ is defined as

$$\widehat{\text{Var}}(M(\mathbf{x})) = \frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} M^2(\mathbf{p}) - \left(\frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} M(\mathbf{p}) \right)^2 \quad (5)$$

with $\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$. For the sake of simplicity in notation, let us denote $|\eta(\mathbf{x})| = N$. We define the random variable $V = \widehat{\text{Var}}(M(\mathbf{x}))$, whose moments are studied in Appendix 6.1. Moments of the distributions used hereafter may be found in Appendix 6.2.

Let us first consider the Rayleigh/Rice MR model. The distribution of the SLV for Rayleigh data may be accurately approximated[2] by the distribution of the difference of two dependent Chi-Square random variables with the same degrees of freedom [8, 45]. For $x \geq 0$ and an even number of degrees of freedom the PDF becomes

$$p_V(x) = \frac{|x|^{N-1} \exp\left(-\frac{1}{4}\alpha^+ x\right)}{\Gamma(N) \left[2\sigma_1^2 \sigma_2^2 (1-\rho^2)\gamma^-\right]^N} \sum_{k=0}^{N-1} \frac{\Gamma(N+k)}{\Gamma(k+1)\Gamma(N-k)} \left(\frac{2}{\gamma^- |x|}\right)^k \quad (6)$$

with

$$\gamma^- = \frac{\left[(\sigma_2^2 - \sigma_1^2)^2 + 4\sigma_1^2 \sigma_2^2 (1-\rho^2)\right]^{1/2}}{\sigma_1^2 \sigma_2^2 (1-\rho^2)} \quad (7)$$

$$\alpha^+ = \gamma^- + \frac{\sigma_2^2 - \sigma_1^2}{\sigma_1^2 \sigma_2^2 (1-\rho^2)}, \quad (8)$$

$\sigma_1^2 = \frac{\sigma_n^2}{N}$, $\sigma_2^2 \approx \frac{\sigma_n^2}{N} \frac{\pi}{4}$ and ρ the correlation coefficient, which for large N may be approximated as

$$\rho \approx \frac{\pi}{4\sqrt{16\pi - 4\pi^2}} \approx 0.95 \quad (9)$$

Although eq. (6) is an accurate approximation of the PDF of the actual distribution, the whole expression is not useful to work with. In addition, structural properties

---

[2] Note that to obtain the final PDF of V it is necessary to calculate the PDF of the sum of Rayleigh variables. This is a classical-hard-to-find problem in communications. Some approximations are usually employed. Here, the simplified one in [123, 8] is considered.

cannot be easily extracted from the expression. Likewise, the SLV of the Rician data may be approximated by the distribution of the difference of two dependent Noncentral Chi-square distributions [8]. To avoid the use of such involved PDFs, an alternative simplified model will be considered.

In the case of N IID normalized Gaussian variables, Cochran's theorem shows that the SLV follows a Chi-square distribution with $N-1$ degrees of freedom [122], which becomes a Gamma distribution for non-normalized variables. The resultant distribution does not depend on the mean of the original Gaussian variables. For other variables, however, the exact distribution of the SLV cannot be achieved, and it has to be approximated. In [151] a Gaussian PDF is used, whereas in [152] a more general approximation is proposed, based on a Chi-square distribution.

In MR data, when the SNR is high enough, it is a common task to approximate the Rician distribution by a Gaussian. As a consequence, the SLV of the Rician area of an MR image can be approximated by a Gamma distribution. In the results section we will check the goodness of these approximation. Taking into account this result together with the approximation in [152], one can also think of the Gamma distribution as a proper candidate to approximate the SLV of other distributions.

We will assume that the SLV of both Rician and Rayleigh data follows a Gamma distribution, with PDF:

$$p_V(x) = x^{k_i-1} \frac{e^{-x/\theta_i}}{\Gamma(k_i)\theta_i^{k_i}}. \tag{10}$$

Parameters k_i and θ_i may be estimated from the mean $E\{V\}$ and variance σ_V^2 of the actual distribution (that will be derived next) as

$$k_i = \frac{E^2\{V\}}{\sigma_V^2} \quad \theta_i = \frac{\sigma_V^2}{E\{V\}}.$$

The mean and the variance of the SLV for each distribution are estimated using the moments derived in Appendix 6.1. For the Rayleigh case, assuming N large they become:

$$E\{V\} = 2\sigma_n^2 \left(1 - \frac{\pi}{4}\right)\left(\frac{N-1}{N}\right)$$

$$\approx 2\sigma_n^2\left(1 - \frac{\pi}{4}\right) \tag{11}$$

$$\sigma_V^2 = \frac{\sigma_n^4}{N}\left(4 + 2\pi - \pi^2 + O\left(\frac{1}{N}\right)\right)$$

$$\approx \frac{\sigma_n^4}{N}(4 + 2\pi - \pi^2). \tag{12}$$

Another interesting parameter in MR-related distribution is the mode. It has been previously used as a way to estimate noise parameters and to identify different regions [8]. We will use it later to identify the noise and background contributions. The mode of a Gamma distribution is calculated as

$$\text{mode}\{V\} = E\{V\} - \frac{\sigma_V^2}{E\{V\}}$$

$$= \sigma_n^2 \left(2 - \frac{\pi}{2} - \frac{1}{N} \left(\frac{5\pi^2 - 16\pi}{2\pi - 8} \right) + O(1/N^2) \right)$$

$$\approx \sigma_n^2 \left(2 - \frac{\pi}{2} \right)$$

which fits the value derived in [8].

For the Rician data, in order to approximate the mean and the variance, a series expansion of the Laguerre polynomials is considered, assuming N large and $A^2 >> \sigma_n^2$. For $x = \frac{A^2}{2\sigma_n^2}$ the mean and variance of the SLV will be

$$E\{V\} = \frac{(N-1)\sigma_n^2}{N} \left(1 - \frac{1}{4x} - \frac{1}{8x^2} + O\left(x^{-3} \right) \right)$$

$$\approx \frac{(N-1)\sigma_n^2}{N} \tag{13}$$

$$\sigma_V^2 = \frac{(N-1)\sigma_n^4}{N^2} \left(2 - \frac{1}{x} + \frac{3(2 - 5N + 3N^2)}{8N(N-1)x^2} + O\left(x^{-3} \right) \right)$$

$$\approx \frac{2(N-1)\sigma_n^4}{N^2} \tag{14}$$

Note that the simplifications made only hold for high SNR. In Fig. 1 the mean and the standard deviation are depicted versus these simplified values for different SNR (with SNR= A/σ_n). The results show that for SNR> 5 the approximations are nearly the actual values. (Note that the closer to one the better the approximation).
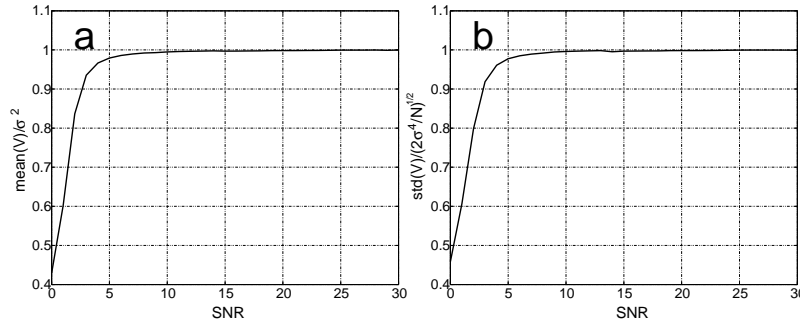
Fig. 1 (a) Actual mean of the SLV of Rician data divided by σ_n^2 (the proposed simplification). (b) Actual standard deviation of the SLV of Rician data divided by $\sqrt{2}\sigma_n^2/\sqrt{N}$ (the proposed simplification).

The mode of the distribution is calculated as

$$\text{mode}\{V\} = \sigma_n^2 \left(1 - \frac{3}{N} + O\left((\sigma/A)^2 \right) \right) \approx \sigma^2$$

Thus, assuming a Rician image with a Rayleigh background and the high SNR in the signal areas, the distribution of the SLV for the whole image can be approximated by a bimodal Gamma distribution, defined as the weighted sum of two Gamma distributions, i.e.

$$p_V(x) = W_B \cdot p_{V_B}(x) + W_I \cdot p_{V_I}(x)$$
$$= W_B \cdot x^{k_B-1} \frac{e^{-x/\theta_B}}{\Gamma(k_B)\theta_B^{k_B}} + W_I \cdot x^{k_I-1} \frac{e^{-x/\theta_I}}{\Gamma(k_I)\theta_I^{k_I}} \tag{15}$$

with $W_B + W_I = 1$. The first term is the Gamma approximation for the Rayleigh area and the second is the Gamma approximation for the Rician area.

The same reasoning may be applied over the central/non-central Chi model. Note that under the high SNR assumption, a non–central Chi may be approximated by a Gaussian, and therefore its SLV may also be approximated by a Gamma distribution.
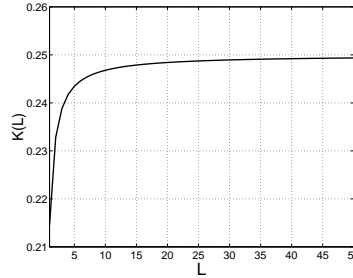
Let us calculate the mean, the variance and the mode of the SLV distribution of central Chi data:

$$E\{V\} = \frac{2\sigma_n^2(N-1)}{N} \left(L - \left(\frac{\Gamma(L+1/2)}{\Gamma(L)} \right)^2 \right) \tag{16}$$

$$\sigma_V^2 = \frac{4\sigma_n^4}{N} \left(L + 8L\frac{\Gamma^2(L+1/2)}{\Gamma^2(L)} - 4\frac{\Gamma^4(L+1/2)}{\Gamma^4(L)} - 4\frac{\Gamma(L+1/2)\Gamma(L+3/2)}{\Gamma^2(L)} \right) + O\left(\frac{1}{N^2} \right) \tag{17}$$

$$\text{mode}\{V\} = 2\sigma_n^2 \left(L - \frac{\Gamma^2(L+1/2)}{\Gamma^2(L)} \right) + O\left(\frac{1}{N} \right) \approx 2K(L)\sigma_n^2 \tag{18}$$

with $K(L) = \left(L - \frac{\Gamma^2(L+1/2)}{\Gamma^2(L)} \right)$. The evolution of this parameter with L is depicted in Fig. 2.

Fig. 2 Evolution of $K(L)$ with L.

For the non–central Chi (assuming high SNR):

$$E\{V\} = \frac{\sigma_n^2(N-1)}{N}\left(1 + \frac{1-2L}{x} + \frac{3-8L+4L^2}{8x^2} + O\left(x^{-3}\right)\right)$$

$$\approx \frac{\sigma_n^2(N-1)}{N} \tag{19}$$

$$\sigma_V^2 = \frac{\sigma_n^4(N-1)}{N^2}\left(2 + \frac{1-2L}{x} + O\left(x^{-2}\right)\right)$$

$$\approx \frac{2\sigma_n^4(N-1)}{N^2} \tag{20}$$

$$\text{mode}\{V\} = \sigma_n^2\left(1 - \frac{3}{N} + O\left((\sigma_n/A)^2\right)\right)$$

$$\approx \sigma_n^2 \tag{21}$$

Note that a series expansion of the Hypergeometric series $_1F_1$ is needed to achieve the result. The distribution of the SLV for the whole image can also be approximated by a bimodal Gamma distribution

$$p_V(x) = W_C \cdot x^{k_C-1}\frac{e^{-x/\theta_C}}{\Gamma(k_C)\theta_C^{k_C}} + W_M \cdot x^{k_M-1}\frac{e^{-x/\theta_M}}{\Gamma(k_M)\theta_M^{k_M}} \tag{22}$$

with $W_C + W_M = 1$. The first term is the Gamma approximation for the central Chi area and the second is the Gamma approximation for the non–central Chi area.

2.3 Background influence on noise estimation

Most of the noise estimation methods for MR data in literature rely on the existence of a homogeneous (either Rayleigh or central Chi) background to carry out the estimation [47]. For some, the background must be segmented [86, 81, 27], while for other techniques a prior segmentation is not needed [37, 136, 83, 8, 84, 47]. Regardless of the chosen method or the necessity of a prior segmentation, this homogeneous background must be present, and the number of points belonging to this background must be large enough when compared with the number of points in the signal areas. However, what *large enough* is, is not properly defined in literature, and it will be related with the distribution of the parameter selected for estimation.

Some post processing software in the scanner usually adds a mask to data, whichs eliminates part of or the whole background, drastically reducing the number of points available for noise estimation. As an illustration, see the MR slice from a real acquisition in Fig. 3. In the right image all the pixels with intensity value equal to 0 are shown in black. A noise pattern can be found around the skull, which corresponds to the actual background, but there is also a greater area where a zero background has been artificially added. This *artificial* zero background is nowadays usually found in MR data when dealing with parallel reconstruction, but not only; even in single coil acquisition a reduction of background areas can be found. See
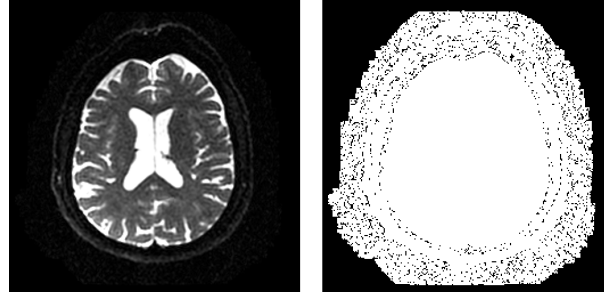
Fig. 3 Left: Slice of an MR volume (from the baseline of a SENSE EPI DT-MRI acquisition). Right: In black all the pixels with intensity value equal to 0. A zero-level background has been artificially added in the scanner.
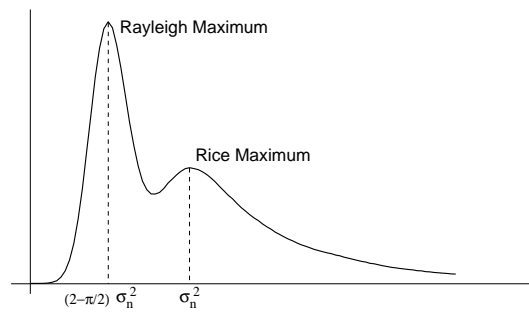
Fig. 4 Position of the maxima of the SLV distribution in MR data, assuming a Rayleigh/Rician model.

for instance the single coil acquisition we will later use for experiments in Fig. 9. The reduction of the background area can make estimation based on Rayleigh distributions fail. Accordingly, for reduced backgrounds, it will be necessary to test if the number of available points is enough for carrying out the estimation with the selected parameter, or if an alternative method must be used instead.

To face this problem we propose to use the properties of the distribution of the SLV of MR data. As described in the previous section, the distribution of the SLV of MR data assuming Rayleigh/Rician data may be modeled as a bimodal Gamma distribution. According to the modes of this Gamma distribution, the maximum of the SLV of the Rayleigh area is located on $\left(2 - \frac{\pi}{2}\right)\sigma_n^2$ and the maximum of the SLV of the Rician data is placed approximately on σ_n^2 (under the high SNR assumption), see Fig. 4 for illustration. Under this assumption, the distribution of the SLV does not depend on the value of the signal, but only on the value of σ_n (and sometimes on the size of the neighborhood used for estimating the SLV, if N is not large). Thus, the shape and position of the distribution of the Rician area will be constant, independently of the different intensity values of the image.

When the weight of Rayleigh-related points outruns the number of Rician-related ones, $W_B > W_I$ in eq. (15), $p(x)$ has a global maximum at $x = \left(2 - \frac{\pi}{2}\right)\sigma_n^2$.

Otherwise the global maximum is at $x = \sigma_n^2$. Thus, a confidence measure can be developed by comparing these local maxima with those of the distribution with the parameter used for estimation.

Method: Let $p_Y(x)$ be the distribution of certain parameter $Y(\mathbf{x})$ used for estimation, whose maximum is placed at x_m:

$$x_m = \arg\max_x p_Y(x) = c_m \sigma_n^2$$

with c_m a constant. Let x_B and x_I be alternatively the Rayleigh and Rician position of the maxima for the SLV distribution:

$$x_B = \arg\max_x p_{V_B}(x) \approx \left(2 - \frac{\pi}{2}\right)\sigma_n^2$$
$$x_I = \arg\max_x p_{V_I}(x) \approx \sigma_n^2$$

The global maximum of the distribution x_G will be either x_B or x_I. However, it is not necessary to identify it as Rician of Rayleigh. If

$$\frac{x_m}{c_m} = \frac{x_G}{2 - \pi/2} \quad \text{or} \quad \frac{x_m}{c_m} = x_G,$$

then the noise estimation is feasible to be done using the parameter $Y(\mathbf{x})$, i.e. there are enough background pixels to perform the estimation. Otherwise, if $\frac{x_m}{c_m} > x_G$, the background assumption fails, and $Y(\mathbf{x})$ is not a proper choice for estimation.

**Example:** As an illustration, consider the second order moment

$$Y(\mathbf{x}) = \langle M(\mathbf{x})^2 \rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} M^2(\mathbf{p}).$$

Noise estimators based on this parameter have been proposed in literature [81, 8, 47]. For the Rayleigh area, the maximum of the distribution of the sample second order moment is placed at $x_m = 2\sigma_n^2$ [8], i.e. $c_m = 2$.

Three different cases are studied: an image with Rayleigh background, an image with a reduced Rayleigh background and an image without the Rayleigh background. The distribution of $Y(\mathbf{x})$, say $p_M(x)$ is depicted together with the distribution of the SLV, say $p_V(x)$ in Fig. 5. Note that:

- For the image with the Rayleigh background, Fig. 5-(a) the maximum of $p_M(x)$ is placed at $2\sigma_n^2$, and the maximum of $p_V(x)$ is the Rayleigh maximum, i.e $x_G = x_B$. So, $\frac{x_m}{c_m} = \frac{2}{4-\pi}x_G$, and $Y(\mathbf{x})$ can therefore be used for estimation.
- When the number of pixels in the background is drastically reduced, the global maximum of the SLV distribution is the one belonging to the Rician area, $x_G = x_I$, see Fig. 5-(b). However, for $Y(\mathbf{x})$, the Rayleigh area is still *stronger* than the Rician one. As a result, the maximum of the distribution of $p_M(x)$ is placed again at $2\sigma_n^2$ and $\frac{x_m}{c_m} = x_G$. $Y(\mathbf{x})$ can still be used for noise estimation.
- When the Rayleigh background is totally removed, the distribution of the SLV reduces to a monomodal Gamma, with the maximum $x_G = x_I$. As there are not
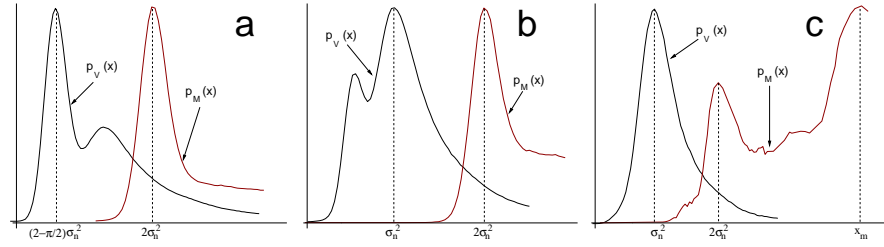
Fig. 5 (Normalized) Distribution of the SLV $p_V(x)$ (black) and the sample second order moment $p_M(x)$ (red); (a) image with a large Rayleigh background; (b) image with a reduced Rayleigh background; (c) image without Rayleigh background.
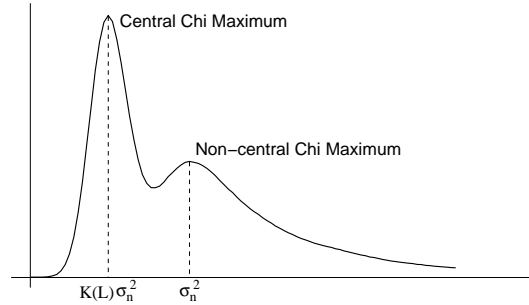
Fig. 6 Position of the maxima of the SLV distribution in MR data, assuming a central/non-central Chi model.

enough point for estimation, the maximum of $p_M(x)$ is no longer related to the Rayleigh area, and it is displaced to higher values, see Fig. 5-(c). In this case $\frac{x_m}{c_m} > x_G$, and consequently $Y(\mathbf{x})$ cannot be used for noise estimation.

The same reasoning can also be applied to the central/non-central Chi model. This time, the maxima of the distributions are located at $2K(L)\sigma_n^2$ for the central Chi data and approximately on σ_n^2 for the non–central Chi (under the high SNR assumption, see Fig. 6 for illustration), i.e:

$$x_C = \arg\max_x p_{V_C}(x) \approx 2K(L)\sigma_n^2$$
$$x_M = \arg\max_x p_{V_M}(x) \approx \sigma_n^2.$$

As a final remark, we must point out that, although the SLV distribution has been used here as a way to assess the viability of some noise estimation technique, it could be used for noise estimation itself. Since a closed expression has been given, i.e. a bimodal Gamma distribution, one can use least squares or a maximum likelihood method to find σ_n, as done in [37, 84, 47] for other parameters.

3 Materials and methods

For the sake of validation and illustration, the following experiments are carried out:

1. Synthetic data is used to test the Gamma distribution assumption for the SLV of the different statistical models. Rayleigh, Rician, central Chi and non–central Chi random data are generated, and the SLV of each set is calculated. The actual distribution is estimated from the histogram and it is compared to the Gamma approximation proposed for all cases. To that end, images of 750×750 points are generated, and the SLV is estimated using 7×7 windows (i.e. $N = 49$ points). Different values of σ_n are considered. For the central/non–central Chi experiment, $L = 8$ is considered in all the cases. For the Rician case, assuming high SNR, the distribution must be independent of the SNR. To better show this property, the Rician experiment is repeated for different values of A ranging in $[100 - 250]$, as shown in Fig. 7.
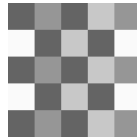
Fig. 7 Image for the Rician experiment.

2. A 256×256 synthetic Rayleigh–distributed image is generated, with a centered 170×170 square Rician inside. Different values of σ_n are considered, and $A = 200$. The distribution of the SLV (using 7×7 windows for estimation) is depicted together with the bimodal Gamma proposed in eq. (15). The values of weights and σ_n are by a least square fitting of the real data.

3. An experiment to test the Gamma approximation over synthetic MR data is carried out. A $181 \times 217 \times 181$ MR synthetic volume, originally noise free, from the Brainweb database [53], Fig. 8-(a), has been artificially corrupted with Rician noise with $\sigma_n = 15$, Fig, 8-(b). Using the mask in Fig. 8-(c) an artificial zero background has been added to the noisy image, Fig. 8-(d). The SLV of the noisy data is calculated using a $3 \times 3 \times 3$ square window and its distribution is depicted for two cases: using the whole Rayleigh background in Fig. 8-(b) and using the reduced Rayleigh background in Fig. 8-(d). The artificial zero background will be removed for variance estimation.

4. To illustrate the proposed method to assess a noise estimation procedure, a new experiment is done. The distribution of the SLV of the noisy synthetic data from BrainWeb of the previous experiment is compared to two other local parameters used for estimation: namely the local mean and the second order moment. Using these moments of the Rayleigh distribution some noise estimators has been defined in literature [81, 86, 8, 47], as, for example, the following:
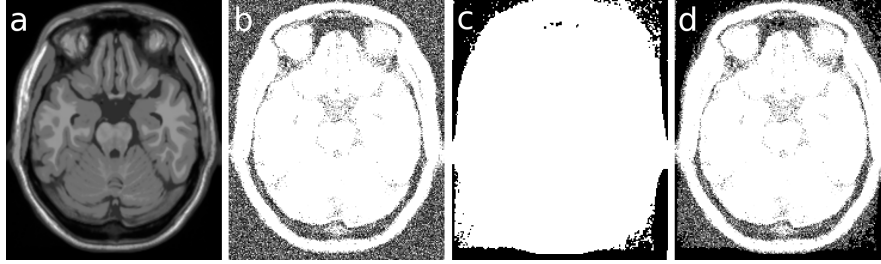
Fig. 8 Slices of a 3D volume from Brainweb. (a) Original image. (b) Image with Rician Noise, $\sigma_n = 15$. (c) Noise mask. Pixels in black are artificially set to 0 in the final image. (d) Image with Rician noise using the mask. Color map in Fig. (b) and (d) has been changed to highlight the noisy areas.

$$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\{\langle M^2(\mathbf{x})\rangle_{\mathbf{x}}\}$$

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\text{mode}\{\langle M(\mathbf{x})\rangle_{\mathbf{x}}\}$$

with $\langle M(\mathbf{x})\rangle_{\mathbf{x}}$ the sample local mean of magnitude image $M(\mathbf{x})$. To make all the distributions be a function of σ_n, the square root of the SLV and the second order moment will be considered. So, for the experiment, the following distributions will be considered:

a. $p_V(x)$: Distribution of the (square root of the) SLV of the noisy image $M(\mathbf{x})$. For the reduced background cases, its maximum will be at σ_n.
b. $p_M(x)$: Distribution of the (square root of the) sample local second order moment of the noisy image $M(\mathbf{x})$. The maximum of the second order moment distribution for Rayleigh data is placed at $2\sigma_n$, so $c_M = \sqrt{2}$ and the maximum of $p_M(x)$ will be then be at $\sqrt{(2)}\sigma_n$ (if the Rayleigh region is big enough).
c. $p_A(x)$: Distribution of the sample local mean of the noisy image $M(\mathbf{x})$. The maxima of the sample mean distribution for Rayleigh data is placed at $\sqrt{2/\pi}\sigma_n$.

5. Finally, the Gamma model for the SLV distribution is tested using real MR data. Two different data sets have been considered. First, the baseline from a single coil DTI acquisition, scanned in a 1.5 Tesla GE system, slick thickness: 5 mm, size $256 \times 256 \times 24$ (real data set 1, RD1, Fig. 9); and second one slice of a multiple coil acquisition, from an 8 coil GE Signa 1.5 Tesla EXCITE 11m4 scanner, FSE Pulse Sequence, TR=500 ms, TE= 13.8 ms, matrix size= 256×256, FOV=20 × 20 cm, slick thickness= 5 mm (real data set 2, RD2, Fig. 10); the magnitude image is reconstructed using SoS. For both data sets, the SLV is estimated using 7×7 square windows, and the PDF is estimated from the histogram. A bimodal Gamma distribution is fitted to each data set using least squares.
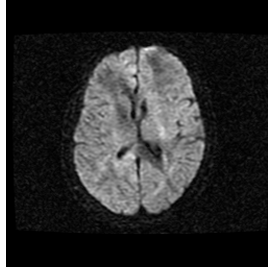
Fig. 9 Real Data set 1 (RD1). Slice from a single coil DTI-MR acquisition.
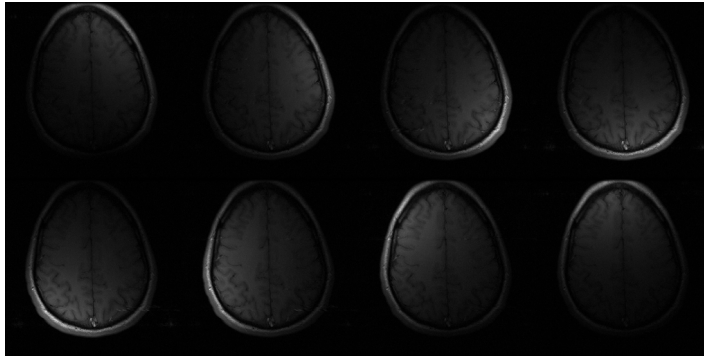
Fig. 10 Real Data set 2 (RD2). Slice from an 8-coil acquisition.

4 Results
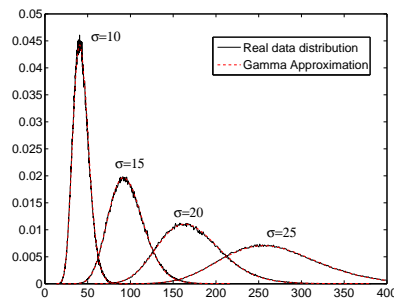
Fig. 11 Distribution of the SLV of Rayleigh data. Real data (black) vs. the proposed Gamma approximation (red) for different σ_n values.

Results for the **first** experiment for Rayleigh data are depicted in Fig. 11. The Gamma approximation totally fits the distribution estimated from the real data for

any σ_n value. We can conclude that the Gamma approximation proposed is a good representation of the SLV distribution for Rayleigh data. To test the effect of the number of points in the approximation, the real SLV data are plotted vs. the proposed Gamma distribution in Fig. 12. Different number of points are considered and $\sigma = 10$ for all 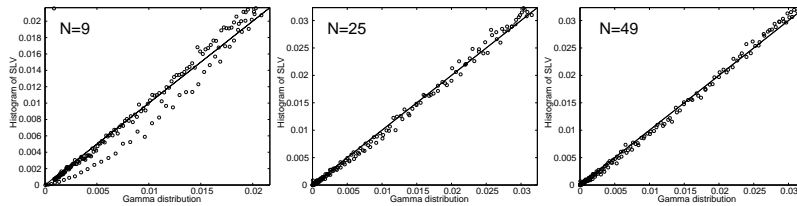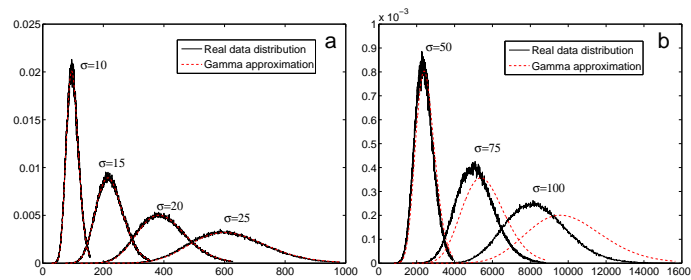cases. Note that the chosen N correspond to 3×3, 5×5 and 7×7 neighborhoods. As N grows, there is a better fit of the data to the proposed model.

Fig. 12 QQ plots of the samples of the SLV of Rayleigh data vs. the proposed Gamma distribution. Different N values are considered and $\sigma = 10$.

Fig. 13 Distribution of the SLV of Rician data for signal value $A = 200$. Real data (black) vs. the proposed Gamma approximation (red) for different σ_n values. Left: High SNR. Right: Low SNR
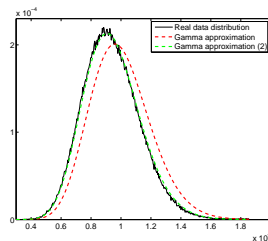
Fig. 14 Distribution of the SLV of Rician data for $SNR = 3$. Real data (black) vs. the proposed Gamma approximation with simplification (red) and with a better approximation of mean and variance (green).

In Fig. 13 the histogram of the SLV of Rician data for different σ_n values is depicted together with the Gamma approximation proposed. As in the Rayleigh case, the Gamma distribution fits the data as long as the high SNR assumption holds. However, for low SNR values (Fig. 13-b) the simplifications made before no longer hold. However, the error commited is not due to the Gamma model proposed, but to the simplification for large SNR done when the mean and the variance are estimated. If necessary, taking more terms in eqs. (13) and (14) will give a better approximation. For instance, taking the series to a higher order, the following mean and variance can be considered:

$$E\{V\} \approx \frac{(N-1)\sigma_n^2}{N}\left(1 - \frac{1}{2\text{SNR}^2}\right)$$
$$\sigma_V^2 \approx \frac{2(N-1)\sigma_n^4}{N^2}\left(1 - \frac{1}{\text{SNR}^2}\right)$$

with SNR=A/σ_n. Note that this time the SNR value must be known. As an illustration, the experiment is repeated, and real data histogram vs. Gamma approximation with these new parameters are shown in Fig. 14. Using the new approximation for the mean and the variance, the Gamma distribution fits again real data.

It is important to recall that, under the high SNR assumption, the resultant distribution of the SLV does not depend on the value of the signal, but only on the value of σ_n and N. Accordingly, regardless of the different intensity values on the image, the shape and position of the distribution will be constant. To test this assumption, a square image with different A values ranging in the interval $A \in [100, 250]$ (see Fig. 7) is now used, see Fig. 15. The Gamma distribution totally fits real data, and the different A values do not influence in the final distribution.

We can conclude that, for the high SNR assumption, the Gamma model is also a good approximation for the SLV of Rician data. However, the model fails when low SNR is considered. The inability of the model to follow the data is not due to the model itself, but to the simplification done when calculating the mean and variance. If more terms are considered in the series expansion, the Gamma approximation will also be a good model for low SNR case. However, the mode will no longer be placed at a value independent of the signal.
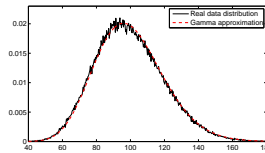
Fig. 15 Distribution of the SLV for a Rician image with multiple levels of signals; $A \in [100, 250]$, $\sigma_n = 10$. Original data (black) vs. theoretical approach.
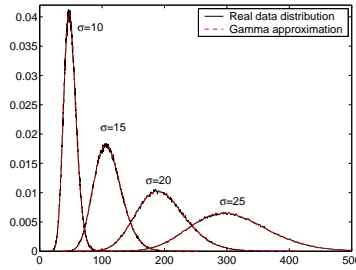
Fig. 16 Distribution of the SLV of central Chi data. Real data (black) vs. the proposed Gamma approximation (red) for different σ_n values and $L = 8$.
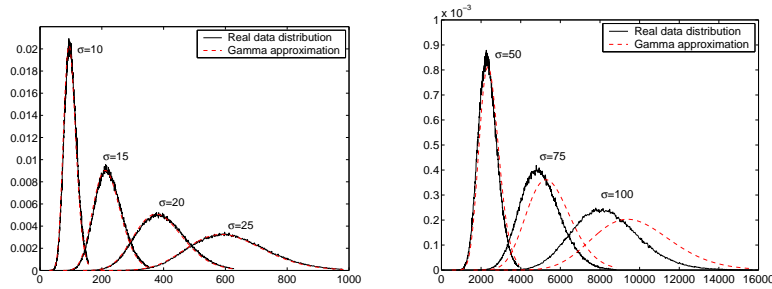
Fig. 17 Distribution of the SLV of non–central Chi data for signal value $A = 200$ and $L = 8$. Real data (black) vs. the proposed Gamma approximation (red) for different σ_n values. Left: High SNR. Right: Low SNR

Results presented in Fig. 16 and Fig. 17 for the central and non-central Chi experiments present the same results as the Rayleigh/Rician case and the same comments can be made

For the **second** experiment, the SLV from a synthetic Rayleigh/Rician image is compared to the Gamma approximation proposed. Results are in Fig. 18 for different SNR values. As expected from the previous experiments, as long as the SNR is high on the Rician area, the assumption holds and the model fits the data. Note that we have used the simplified version of eqs. (13) and (14), hence the error in the Rician area for the low SNR values.

For the **third** experiment, the distribution of the SLV of the noisy images are depicted in Fig. 19 for both cases: the whole Rayleigh background, Fig. 8-(b) (black line), and the smaller Rayleigh background in Fig. 8-(d) (red line). The artificial zero background has been removed for variance estimation. Note that for the first case the Rayleigh maximum is the absolute maximum ($W_B > W_I$), while in the second one, the global maximum is the Rician one. According to the model proposed, the maxima should be related as $\frac{x_B}{2-\pi/2} = x_I = \sigma_n$. From the Figure, it can be effectively seen that $x_I \approx \sigma_n^2 = (15)^2 = 225$ and $x_B = \approx (2-\pi/2)\sigma_n^2 = 98.57$.

The **fourth** experiment is an illustration of the method to test the validity of the background for noise estimation. Fig. 20-(a) is the case showed in Fig.8-(d).
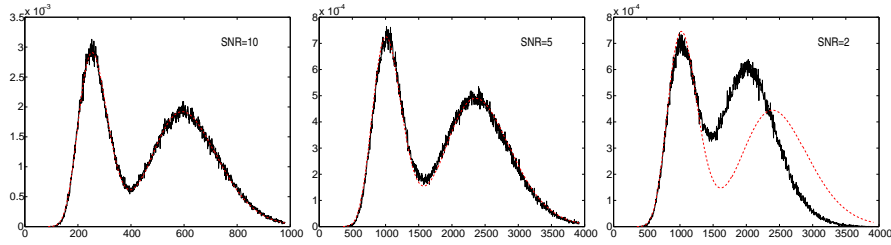
Fig. 18 Histogram of the SLV distribution (black) vs. theoretical distribution (red), using least squares to adjust them.
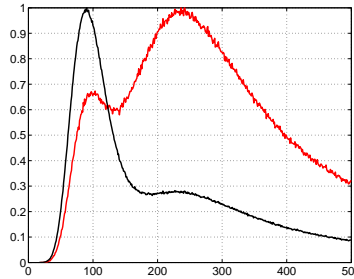
Fig. 19 (Normalized by maximum) distribution of the (square rooted) SLV in the MR noisy volume for $\sigma_n = 15$: for Fig. 8-(b) (black line) and Fig. 8-(d) (red line).

In Fig. 20-(b) the noisy background around the skull has been reduced even more and in Fig. 20-(c) it has been completely removed. For all the cases, the maximum of $p_V(x)$ is placed at a value slightly greater that $\sigma_n = 15$ (the Rician maximum), due to numerical issues. The Rayleigh-related maximum is progressively decreasing as the background is removed. For the first two cases, the maximum of $p_M(x)$ is placed at $x_M = \sqrt{2}\sigma_n = 21.21$, and the maximum of $p_A(x)$ is placed at $x_A = \sqrt{2 - \pi/2}\sigma_n = 18.79$. So, for the first two backgrounds, these parameters may be used for noise estimation, since $x_G = x_I$, $\frac{x_M}{c_M} = x_G$ and $\frac{x_A}{c_A} = x_G$. However, for the third case, since the background has been completely removed, the maxima of $p_M(x)$ and $p_A(x)$ are displaced to higher values, and therefore, for both parameters $\frac{x_M}{c_M} > x_G$ and $\frac{x_A}{c_A} > x_G$, and consequently we can state that there is not enough background for an estimation.

In the **last** experiment, we simply test the Gamma assumption over real data sets. Results for RD1 are depicted on Fig. 21, and for RD2 on Fig. 22. Since both data sets have an homogeneous Rayleigh background, the distributions mainly show the Rayleigh contribution of the data, i.e just one mode. In both cases, the Gamma distributions totally fit the real data distribution. Therefore, we can conclude that the Gamma approximation proposed is a proper representation of the SLV of MR data.
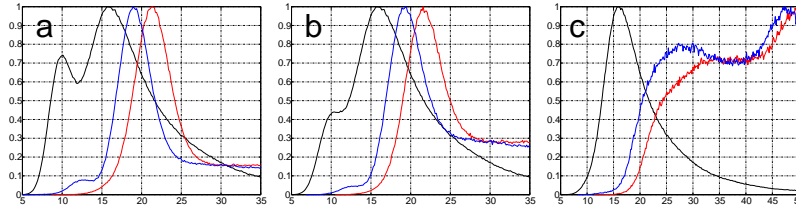
Fig. 20 (Normalized) Distribution $p_V(x)$ (black) vs. $p_M(x)$ (red) and $p_A(x)$ (blue) for the noisy volume in Fig. 8. The maxima of the distributions are used for noise estimation. Three different backgrounds are considered. The size of the background is decreasing from (a) to (c).
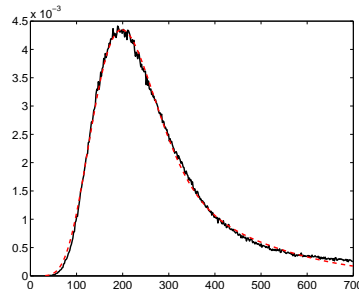
Fig. 21 Single coil real data: Histogram of the SLV distribution of real Data set 1 (black) vs. bimodal Gamma distribution (red), using least squares to adjust them.
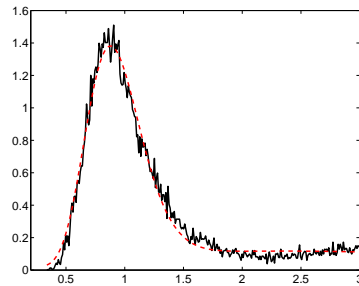
Fig. 22 Multiple coil real data: Histogram of the SLV distribution of real Data set 2 (black) vs. bimodal Gamma distribution (red), using least squares to adjust them.

5 Discussion and conclusions

Two main contributions are presented in the current paper: **first**, a model for the SLV in MR data is derived and **second**, as an example of use, the proposed distribution can be used as a way to test the viability of one particular parameter to be used as a noise estimator.

The SLV is used as a tool to study the behavior of the signal/background areas in different MR models. The SLV of Gaussian data is known to follow a Gamma distribution; this PDF has shown to be a very good model for the distribution of the SLV

of Rayleigh, Rician, central Chi and non–central Chi variables, the usual models for MR, either for one or multiple coil acquired data. The use of this distribution to model the background and signal areas of the image will result in a bimodal Gamma distribution whose maxima are related to the properties of these two areas. Experiments show that the proposed distribution totally fits real data. For low SNR, the model fails in the signal areas due to the simplification of the estimated parameters, but not due to the model itself. If necessary, for low SNR more terms may be added to the calculation of mean and variance, improving the data fitting of the model. The price to pay is that the SNR value must be known. Note that this is not useful in practical situations, where a MR image may have different SNR values in different regions.

In the current paper the SLV has been used exclusively to analyze MR-based models of noise, i.e. the Rayleigh/Rician model and the central/non–central Chi. However, note that the Gamma characterization of the SLV may be extrapolated to other noise models and other probability distributions, such as the statistical characterization of BOLD signals in fMRI [154, 155, 156].

As an illustration of the proposed model, a method to test background-based noise estimation techniques has been developed. If the distribution of the parameter used for estimation has its maximum at a value related to the noise power, a comparison with the maxima of the bimodal Gamma distribution will tell if the number of background pixels is enough to perform such estimation. This test is motivated by the reduced background we can usually find as a result of modern scanning techniques. It is not unusual to find that the scanner has deleted part of the actual background of the image. As a result, the number of pixels available to perform the estimation may be drastically reduced. The proposed method presents a way to test if the number of remaining background pixels is large enough to perform the estimation.

Finally, we want to recall that the SLV distribution model here presented may be used for further purposes, not only as a way to check a noise estimation technique. Following, we propose some possible applications: (1) Alternative noise estimation algorithms may be developed by fitting a bimodal Gamma PDF to the real data SLV histogram. Using least squares or a maximum likelihood algorithm, as done in [37, 84, 47], σ_n-related parameters may be estimated from real data, and from them the variance of noise can be derived. (2) Another application would deal with inhomogeneities in the image and non-stationarity of signals. They both have an impact over the statistical model and consequently they affect to the shape and position of the SLV histogram. A method to detect and characterize such homogeneities may be based on how the real SLV distribution differs from the theoretical one. (3) As the position of the maxima is related to the variance of noise, consistency of the data after some noise reducing procedure may be checked. (4) The position of the signal maximum for the high SNR assumption will be always placed at σ_n^2, regardless of being Rician or non–central Chi data, while the position of the background maximum will be related to the number of coils. The underlying statistical model and the number of coils in the acquisition may be inferred if necessary from a comparison between these maxima positions.

6 Appendices

6.1 About the Sample Local Variance

Let X_i, $i = \{1, \cdots, N\}$ be a set of random variables independent and identically distributed, following a known distribution with moments μ_k. The (biased) SLV is defined as

$$\widehat{\mathrm{Var}}(X_i) = \frac{1}{N} \sum_{i=1}^{N} (X_i)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} X_i \right)^2. \tag{23}$$

Note that it is necessary to multiply eq. (23) by $\frac{N}{N-1}$ to make it unbiased. If we define the variable $V = \widehat{\mathrm{Var}}(X_i)$, the mean and variance of such a distribution are derived as follows:

$$E\{V\} = \left(1 - \frac{1}{N} \right) (\mu_2 - \mu_1^2) \tag{24}$$

$$E\{V^2\} = \frac{1}{N^3}[(N^2 - 2N + 1)\mu_4 + (N^3 - 3N^2 + 5N - 3)\mu_2^2 + (-2N^3 + 12N^2 - 22N + 12)\mu_2\mu_1^2$$
$$+ (N^3 - 6N^2 + 11N - 6)\mu_1^4 + (-4N^2 + 6N - 4)\mu_3\mu_1] \tag{25}$$

$$\mathrm{Var}(V) = \frac{1}{N^3}\left[(-4N^2 + 8N - 4)\mu_3\mu_1 + (8N^2 - 20N + 12)\mu_2\mu_1^2 + (-N^2 + 4N - 3)\mu_2^2 \right.$$
$$\left. + (N^2 - 2N + 1)\mu_4 + (-4N^2 + 10N - 6)\mu_1^4 \right] \tag{26}$$

Multiply the mean by $\frac{N}{N-1}$ and the variance by $\left(\frac{N}{N-1} \right)^2$ for the unbiased case.

6.2 About the moments of the distributions

For the calculation of the mean and variance of the sample local mean, the first four moments of the related distributions are needed. For the Rayleigh distribution:

$$\mu_1 = \sqrt{\frac{\pi}{2}}\sigma \quad \mu_2 = 2\sigma^2$$

$$\mu_3 = 3\sqrt{\frac{\pi}{2}}\sigma^3 \quad \mu_4 = 8\sigma^4$$

For the Rician distribution

$$\mu_1 = \sqrt{\frac{\pi}{2}} L_{1/2} \left(-\frac{A^2}{2\sigma^2} \right) \sigma \quad \mu_2 = A^2 + 2\sigma^2$$

$$\mu_3 = 3\sqrt{\frac{\pi}{2}} L_{3/2} \left(-\frac{A^2}{2\sigma^2} \right) \sigma^3 \quad \mu_4 = A^4 + 8\sigma^2 A^2 + 8\sigma^4$$

where $L_\nu(x)$ is the Laguerre polynomial of order μ-th. A series expansion of these Laguerre polynomials may be considered considered, assuming N large and $A^2 >> \sigma^2$, which actually is the high SNR assumption. For $x = \frac{A^2}{2\sigma^2}$ the polynomials for $\nu = 1/2$ and $\nu = 3/2$ are

$$L_{1/2}(-x) = \frac{2\sqrt{x}}{\sqrt{\pi}} + \frac{1}{2\sqrt{\pi}\sqrt{x}} + \frac{1}{16\sqrt{\pi}x^{3/2}} + O\left(x^{-5/2}\right)$$

$$L_{3/2}(-x) = \frac{4x^{3/2}}{3\sqrt{\pi}} + \frac{3\sqrt{x}}{\sqrt{\pi}} + \frac{3}{8\sqrt{\pi}\sqrt{x}} + \frac{1}{32\sqrt{\pi}x^{3/2}} + O\left(x^{-5/2}\right).$$

The moments for the central Chi distribution are

$$\mu_1 = \sqrt{2}\frac{\Gamma(L + 1/2)}{\Gamma(L)}\sigma \quad \mu_2 = 2L\sigma^2$$

$$\mu_3 = 3\sqrt{2}\frac{\Gamma(L + 3/2)}{\Gamma(L)}\sigma^3 \quad \mu_4 = 4L(L + 1)\sigma^4$$

and for the non–central Chi:

$$\mu_1 = \sqrt{2}\frac{\Gamma(L + 1/2)}{\Gamma(L)}{}_1F_1\left(-\frac{1}{2}, L, -\frac{A_L^2}{2\sigma^2}\right)\sigma \quad \mu_2 = A_L^2 + 2L\sigma^2$$

$$\mu_3 = 3\sqrt{2}\frac{\Gamma(L + 3/2)}{\Gamma(L)}{}_1F_1\left(-\frac{3}{2}, L, -\frac{A_L^2}{2\sigma^2}\right)\sigma^3 \quad \mu_4 = A_L^4 + 4(L + 1)A_L^2\sigma^2 + 4L(L + 1)\sigma^4$$

with ${}_1F_1$ the Kummer confluent hypergeometric function.

Acknowledgments

Part III
Noise estimators for pMRI

Noise Estimation in Parallel MRI: GRAPPA and SENSE

Santiago Aja-Fernández, Gonzalo Vegas-Sánchez-Ferrero, Antonio Tristán-Vega*

Abstract Parallel imaging methods allow to increase the acquisition rate via sub-sampled acquisitions of the $\mathbf{k}-$space. SENSE and GRAPPA are the most popular reconstruction methods proposed in order to suppress the artifacts created by this subsampling. The reconstruction process carried out by both methods yields to a variance of noise value which is dependent on the position within the final image. Hence, the traditional noise estimation methods —based on a single noise level for the whole image— fail. In this paper we propose a novel methodology to estimate the spatial dependent pattern of the variance of noise in SENSE and GRAPPA reconstructed images. In both cases, some additional information must be known beforehand: the sensitivity maps of each receiver coil in the SENSE case and the reconstruction coefficients for GRAPPA.

1 Introduction

Magnetic Resonance Imaging (MRI) is known to be affected by several sources of quality deterioration, due to limitations in the hardware, scanning times, movement of patients, or even the motion of molecules in the scanning subject. Among them, noise is one source of degradation that affects acquisitions. The presence of noise over the acquired MR signal is a problem that affects not only the visual quality of the images, but also may interfere with further processing techniques such as registration or tensor estimation in Diffusion Tensor MRI [157].

Noise has usually been statistically modeled attending to the scanner coil architecture. For a single–coil acquisition, the complex spatial MR data is typically modeled as a complex Gaussian process, where the real and imaginary parts of the

---

original signal are corrupted with uncorrelated Gaussian noise with zero mean and equal variance σ_n^2. Thus, the magnitude signal is the Rician distributed envelope of the complex signal [5]. This Rician distribution whose variance is the same for the whole image is also known as *homogeneous* Rician distribution or, more accurately, *stationary* Rician distribution, and it has been the most used model in literature for multiple applications [4, 8, 37, 23, 25, 58].

When a multiple–coil MR acquisition system is considered, the Gaussian process is repeated for each receiving coil. As a consequence, noise in each coil in the **k**–space can be also modeled as a complex stationary Additive White Gaussian Noise process, with zero mean and equal variance. In that case, the noise in the complex signal in the **x**–space for each coil will also be Gaussian. If the **k**−space is fully sampled, the composite magnitude signal (CMS, i.e. the final real signal after reconstruction) is obtained using methods such as the sum-of-squares (SoS) [27]. Assuming the noise components to be identically and independently distributed, the CMS will follow a non-central chi (nc-χ) distribution [27]. If the correlation between coils is taken into account, the data does not strictly follow a nc-χ but, for practical purposes, it can be modeled as such, but taking into account effective parameters [49, 50].

However, in multiple–coil systems, fully sampling the **k**−space acquisition is not the common trend in acquisition. Nowadays, due to time restrictions, most acquisitions are usually accelerated by using parallel MRI (pMRI) reconstruction techniques, which allow to increase the acquisition rate via subsampled acquisitions of the **k**−space. This acceleration goes together with an artifact known as *aliasing*. Many reconstruction methods have been proposed in order to suppress the aliasing created by this subsampling, being SENSE (Sensitivity Encoding for Fast MRI) [32] and GRAPPA (Generalized Autocalibrating Partially Parallel Acquisition) [33] dominant among them. From a statistical point of view, both reconstruction methods will affect the stationarity of the noise in the reconstructed data, i.e. the spatial distribution of the noise across the image. As a result, if SENSE is used, the magnitude signal may be considered Rician distributed [46, 43] but the value of the statistical parameters and, in particular, the variance of noise σ_n^2, will vary for different image locations, i.e. it becomes **x**-dependent. Similarly, if GRAPPA is used, the CMS may be approximated by a non-stationary nc-χ distribution [42, 43] with effective parameters.

Noise estimators proposed in literature are based on the assumption of a single σ_n^2 value for all the pixels in the image, either assuming a Rician model[7, 81, 37, 8, 47, 84] or a nc-χ [27, 47, 48, 49]. Accordingly, those methods do not apply when dealing with pMRI and non-stationary noise. Noise estimators must therefore be reformulated in order to cope with these new image modalities.

In this paper we propose different methodologies to estimate the spatially distributed variance of noise σ_n^2 from the magnitude signal when SENSE or GRAPPA are used as pMRI technique.

2 Noise statistical models in pMRI

As previously stated, most noise estimation methods in literature rely on the assumption of a single value of σ_n^2 for every pixel within the image. However, this is no longer the case when pMRI protocols are considered.

In multiple coil systems, the acquisition rate may be increased by subsampling the $\mathbf{k}-$space data [30, 31], while reducing phase distortions when strong magnetic field gradients are present. The immediate effect of the $\mathbf{k}-$space subsampling is the appearance of aliased replicas in the image domain retrieved at each coil. In order to suppress or correct this aliasing, pMRI combines the redundant information from several coils to reconstruct a single non-aliased image domain.

The commonly used (stationary) Rician and nc-χ models do not necessarily hold in this case. Depending on the way the information from each coil is combined, the statistics of the image will follow different distributions. It is therefore necessary to study the behavior of the data for a particular reconstruction method. We will focus on two of the most popular methods, SENSE [32] and GRAPPA [33], in their most basic formulation.

In the following sections we will assume an L-coil configuration, being L the number of coils in the system. $s_l^{\mathcal{S}}(\mathbf{k})$ is the subsampled signal at the l-th coil of the $\mathbf{k}-$space ($l = 1, \cdots, L$), $S_l^{\mathcal{S}}(\mathbf{x})$ is the subsampled signal in the image domain, i.e., the $\mathbf{x}$-space, and r is the subsampling rate. The $\mathbf{k}$-space data at each coil can be accurately described by an Additive White Gaussian Noise (AWGN) process, with zero mean and variance σ_K^2:

$$s_l^{\mathcal{S}}(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; \sigma_{K_l}^2), \quad l = 1, \cdots, L \qquad (1)$$
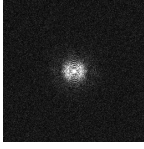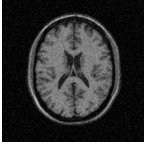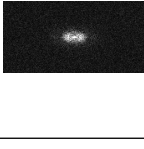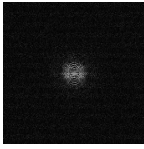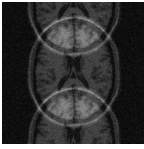
with $a_l(\mathbf{k})$ the noise-free signal and $n_l(\mathbf{k}; \sigma_{K_l}^2) = n_{l_r}(\mathbf{k}; \sigma_{K_l}^2) + jn_{l_i}(\mathbf{k}; \sigma_{K_l}^2)$ the AWGN process, which is initially assumed stationary so that $\sigma_{K_l}^2$ does not depend on $\mathbf{k}$. The complex $\mathbf{x}$-space is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l^{\mathcal{S}}(\mathbf{k})$ for each slice or volume, so the noise in the complex $\mathbf{x}$-space is still Gaussian [43]:

$$S_l^{\mathcal{S}}(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; \sigma_l^2), \quad l = 1, \cdots, L$$

where $N_l(\mathbf{x}; \sigma_l^2) = N_{l_r}(\mathbf{x}; \sigma_l^2) + jN_{l_i}(\mathbf{x}; \sigma_l^2)$ is also a complex AWGN process (note we are assuming that there are not any spatial correlations) with zero mean and covariance matrix:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1L} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L1} & \sigma_{L2} & \cdots & \sigma_L^2 \end{pmatrix}. \qquad (2)$$

The relation between the noise variances in the $\mathbf{k}$ and $\mathbf{x}$-domains is given by the number of points used for the iDFT:

Noise relations			
k−space	Parameters	**x**-space	Relation
	Fully sampled, $\sigma^2_{K_l}$ **k**-size: $\lvert\Omega\rvert$		$\sigma^2_l = \frac{1}{\lvert\Omega\rvert}\sigma^2_{K_l}$, **x**-size: $\lvert\Omega\rvert$
	Subsampled r, $\sigma^2_{K_l}$ **k**-size: $\lvert\Omega\rvert/r$		$\sigma^2_l = \frac{r}{\lvert\Omega\rvert}\sigma^2_{K_l}$, **x**-size: $\lvert\Omega\rvert/r$ (SENSE)
	Subsampled r, $\sigma^2_{K_l}$ **k**-size: $\lvert\Omega\rvert$ (zero padded)		$\sigma^2_l = \frac{1}{\lvert\Omega\rvert\cdot r}\sigma^2_{K_l}$, **x**-size: $\lvert\Omega\rvert$ (GRAPPA)

Table 1 Relations between the variance of noise in complex MR data for each coil in the **k**−space and the image domain.

$$\sigma^2_l = \frac{r}{\lvert\Omega\rvert}\sigma^2_{K_l}.$$

with $\lvert\Omega\rvert$ the final number of pixels in the field of view (FOV). Note that the final noise power is greater than in the fully sampled case due to the reduced **k**−space averaging, as it will be the case with SENSE (see below). On the contrary, the iDFT may be computed after zero-padding the missing (not sampled) **k**−space lines, and then we have [42]:

$$\sigma^2_l = \frac{1}{\lvert\Omega\rvert\cdot r}\sigma^2_{K_l}.$$

In the latter case the noise power is reduced with respect to the fully sampled case, since we average exactly the same number of samples but only 1 of each r of them contributes a noise sample (this will also be the case with GRAPPA). Finally, note that although the level of noise is smaller in GRAPPA due to the zero padding, the SNR does not increase, since the zero padding produces also a reduction of the level of the signal.

Relations between the variance of noise in complex **x**-space and **k**−space for each coil is summarized in Table 1.

2.1 Statistical Noise Model in SENSE Reconstructed Images

Prior to the definition of the estimators, the statistical noise model in SENSE must be properly defined. Many studies have been made about this topic from a SNR or a g-factor (noise amplification) point of view [32, 43, 52]. Since this paper is focused on the σ_n^2 value estimation rather than a SNR level, an equivalent reformulation must be done, more coherent with the signal and noise analysis usually assumed for noise estimation.

In multiple coil scanners, the signal acquired in each coil, $l = 1, 2, \cdots, L$, can be modeled in the $\mathbf{k}-$space by the following equation [30, 158]:

$$s_l(\mathbf{k}) = \int_V C_l(\mathbf{x}) S_0(\mathbf{x}) e^{j2\pi k \cdot \mathbf{x}} d\mathbf{x},$$

where $S_0(\mathbf{x})$ is the excited spin density function throughout the volume V (it is sometime denoted by $\rho(\mathbf{x})$, and it can be seen as an *original image* weighted by the spatial sensitivity of coil l-th, $C_l(\mathbf{x})$. In the $\mathbf{x}$-space this is equivalent to [30, 34]:

$$S_l(\mathbf{x}) = C_l(\mathbf{x}) S_0(\mathbf{x}), \quad l = 1, \cdots, L. \tag{3}$$

An accelerated pMRI acquisition with a factor r will reduce the matrix size of the image at every coil. The signal in one pixel at location (x, y) of l-th coil can be now written as [32, 34]:

$$S_l(x, y) = C_l(x, y_1) S_0(x, y_1) + \cdots + C_l(x, y_r) S_0(x, y_r). \tag{4}$$

Let us call $S_l^{\mathcal{S}}(x, y)$ to the subsampled signal at coil l-th and $S^{\mathcal{R}}(x, y)$ to the final reconstructed image. Note that the latter can be seen as an estimator of the original image $S^{\mathcal{R}}(x, y) = \widehat{S_0}(x, y)$ that can be obtained from eq. (4)

$$
\begin{aligned}
S_l^{\mathcal{S}}(x, y) &= C_l(x, y_1) \widehat{S_0}(x, y_1) + \cdots + C_l(x, y_r) \widehat{S_0}(x, y_r) \\
&= C_l(x, y_1) S^{\mathcal{R}}(x, y_1) + \cdots + C_l(x, y_r) S^{\mathcal{R}}(x, y_r) \quad l = 1, \cdots, L
\end{aligned}
$$

$S^{\mathcal{R}}(x, y)$ can be easily derived from this relation. For instance, for $r = 2$ for pixel (x, y), $S^{\mathcal{R}}(x, y)$ becomes [32, 30, 34]

$$\begin{bmatrix} S_1^{\mathcal{R}} \\ S_2^{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 \ \mathbf{W}_2 \end{bmatrix} \times \begin{bmatrix} S_1^{\mathcal{S}} \cdots S_L^{\mathcal{S}} \end{bmatrix}, \tag{5}$$

where $S_i^{\mathcal{R}}$ stands for each of the r reconstructed pixels. In matrix form for each pixel and an arbitrary r

$$S_i^{\mathcal{R}} = \mathbf{W}_i \times \mathbf{S}^{\mathcal{S}} \quad i = 1, \cdots, r, \tag{6}$$

with $\mathbf{W} = [\mathbf{W}_1, \cdots \mathbf{W}_r]$ a reconstruction matrix created from the sensitivity maps at each coil. These maps, $\mathbf{C} = [\mathbf{C}_1, \cdots, \mathbf{C}_L]$ are estimated through calibration

right before each acquisition session. Once they are known, the matrix $\mathbf{W}$ reduces to a least-squares solver for the overdetermined problem $\mathbf{C}(x,y) \times \mathbf{S}^{\mathcal{R}}(x,y) \simeq \mathbf{S}^{\mathcal{S}}(x,y)$ [32, 34]:

$$\mathbf{W}(x,y) = (\mathbf{C}^*(x,y)\mathbf{C}(x,y))^{-1}\mathbf{C}^*(x,y). \qquad (7)$$

The correlation between coils may be incorporated in the reconstruction as a pre-whitening matrix for the measurements, and $\mathbf{W}(x,y)$ becomes then a weighted least squares solver with correlation matrix $\mathbf{\Sigma}$:

$$\mathbf{W}(x,y) = (\mathbf{C}^*(x,y)\mathbf{\Sigma}^{-1}\mathbf{C}(x,y))^{-1}\mathbf{C}^*(x,y)\mathbf{\Sigma}^{-1}.$$

The SNR of the fully sampled image and the image reconstructed with SENSE are related by the so-called g-factor, g [52, 34]:

$$\mathrm{SNR}_{\mathrm{SENSE}} = \frac{\mathrm{SNR}_{\mathrm{full}}}{\sqrt{r} \cdot g} \qquad (8)$$

However, in our problem we are more interested on the actual noise model underlying the SENSE reconstruction and on the final variance of noise. The final signal $S_i^{\mathcal{R}}$ is obtained as a linear combination of $S_l^{\mathcal{S}}$, where the noise is Gaussian distributed. Thus, the resulting signal is also Gaussian, with variance:

$$\sigma_i^2 = \mathbf{W}_i^* \mathbf{\Sigma} \mathbf{W}_i. \qquad (9)$$

Since $\mathbf{W}_i$ is position-dependent, i.e. $\mathbf{W}_i = \mathbf{W}_i(x,y)$, so will be the variance of noise, $\sigma_i^2(x,y)$. For further reference, when the whole image is taken into account, let us denote the variance of noise for each pixel in the reconstructed data by $\sigma_{\mathcal{R}}^2(\mathbf{x})$.

All in all, noise in the final reconstructed signal $S^{\mathcal{R}}(x,y)$ will follow a complex Gaussian distribution. If the magnitude is considered, i.e. $M(x,y) = |S^{\mathcal{R}}(x,y)|$, the final magnitude image will follow a Rician distribution [43], just like single-coil systems.

To sum up: (1) Subsampled multi coil MR data reconstructed with Cartesian SENSE follow a Rician distribution at each point of the image; (2) The resulting distribution is non-stationary. This means that the variance of noise will vary from point to point across the image; (3) The final value of the variance of noise at each point will only depend on the covariance matrix of the original data (prior to reconstruction) and on the sensitivity map, and not on the data themselves.

2.2 Noise statistical model in GRAPPA

The GeneRalized Autocalibrated Partially Parallel Acquisitions (GRAPPA) [33] reconstruction strategy estimates the full $\mathbf{k}-$space in each coil from a sub–sampled $\mathbf{k}-$space acquisition. The reconstructed lines are estimated through a linear com-

bination of the existing samples. Weighted data in a neighborhood $\eta(\mathbf{k})$ around the estimated pixel from several coils is used for such an estimation. While the sampled data $s_l^{\mathcal{S}}(\mathbf{k})$ remain the same, the reconstructed lines $s_l^{\mathcal{R}}(\mathbf{k})$ are estimated through a linear combination of the existing samples. Weighted data in a neighborhood $\eta(\mathbf{k})$ around the estimated pixel from several coils is used for such an estimation:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} \sum_{\mathbf{c} \in \eta(\mathbf{k})} s_m^{\mathcal{S}}(\mathbf{k} - \mathbf{c})\omega_m(l, \mathbf{c}), \tag{10}$$

with $s_l(\mathbf{k})$ the complex signal from coil l at point $\mathbf{k}$ and $\omega_m(l, \mathbf{k})$ the complex reconstruction coefficients for coil l. These coefficients are determined from the low-frequency coordinates of $\mathbf{k}$–space, termed the Auto Calibration Signal (ACS) lines, which are sampled at the Nyquist rate (i.e. unaccelerated). Breuer *et al.* [51] pointed out that eq. (10) can be rewritten using the *convolution* operator:

$$s_l^{\mathcal{R}}(\mathbf{k}) = \sum_{m=1}^{L} s_m^{\mathcal{S}}(\mathbf{k}) \circledast \mathsf{w}_m(l, \mathbf{k}), \tag{11}$$

where $\mathsf{w}_m(l, \mathbf{k})$ is a convolution kernel that can be easily derived from the GRAPPA weight set $\omega_m(l, \mathbf{k})$. Since a (circular) convolution in the $\mathbf{k}$–space is equivalent to a product into the $\mathbf{x}$-space, we can write:

$$S_l^{\mathcal{R}}(\mathbf{x}) = |\Omega| \sum_{m=1}^{L} S_m^{\mathcal{S}}(\mathbf{x}) \times W_m(l, \mathbf{x}),$$

with $W_m(l, \mathbf{x})$ the GRAPPA reconstruction coefficients in the $\mathbf{x}$-space and $|\Omega|$ the size of the image in each coil. The CMS can be obtained using the SoS of the signal in each coil:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l^{\mathcal{R}}(\mathbf{x})|^2}. \tag{12}$$

In [42] authors pointed out that the resultant distribution of the CMS in eq. (3) is not strictly a nc-χ, but its behavior will be very similar and could be modeled as such with a small approximation error. However, the reconstruction method will highly increase the correlations between the reconstructed signals in each coil, what translates into a decrease of the number of Degrees of Freedom of the distribution. As a consequence, the final distribution will show a (reduced) *effective number of coils* L_{eff} and an (increased) *effective variance of noise* σ_{eff}^2:

$$L_{\text{eff}}(\mathbf{x}) = \frac{|\mathbf{A}|^2 \operatorname{tr}\left(\mathbf{C}_X^2\right) + \left(\operatorname{tr}\left(\mathbf{C}_X^2\right)\right)^2}{\mathbf{A}^* \mathbf{C}_X^2 \mathbf{A} + ||\mathbf{C}_X^2||_F^2}, \tag{13}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\operatorname{tr}\left(\mathbf{C}_X^2\right)}{L_{\text{eff}}}, \tag{14}$$

where $\mathbf{C}_X^2(\mathbf{x}) = \mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^*$ is the covariance matrix of the *interpolated* data at each spatial location, $\mathbf{A}(\mathbf{x}) = [A_1, \cdots, A_L]^T$ is the noise-free reconstructed signal, $||.||_F$ is the Frobenious norm, $\boldsymbol{\Sigma}$ is the covariance matrix of the *original* data and $\mathbf{W}(\mathbf{x})$ is the GRAPPA interpolation matrix for each $(\mathbf{x})$:

$$\mathbf{W}(\mathbf{x}) = \begin{pmatrix} W_1(1,\mathbf{x}) & \cdots & W_1(L,\mathbf{x}) \\ \vdots & \ddots & \vdots \\ W_L(1,\mathbf{x}) & \cdots & W_L(L,\mathbf{x}) \end{pmatrix}$$

Although the nc-χ model is feasible for GRAPPA, the resulting distribution is non-stationary since the effective parameters are spatially dependent.

2.3 Practical simplifications over the GRAPPA model

For practical purposes, in order to make the noise estimation feasible, some simplifications can be made over eq. (13) and eq. (14). We will simplify the problem by assuming that the variance of noise is the same for every coil, $\sigma_l^2 = \sigma_n^2$, and that the signal is also the same $A_i = A_j$ for all i, j. The covariance matrix can therefore be written as:

$$\boldsymbol{\Sigma} = \sigma_n^2 \cdot \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1L} \\ \rho_{21} & 1 & \cdots & \rho_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{L1} & \rho_{L2} & \cdots & 1 \end{pmatrix}. \tag{15}$$

Accordingly, matrix $\mathbf{C}_X^2$ becomes

$$\mathbf{C}_X^2(\mathbf{x}) = \sigma_n^2 \cdot \mathbf{W} \times \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1L} \\ \rho_{21} & 1 & \cdots & \rho_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{L1} & \rho_{L2} & \cdots & 1 \end{pmatrix} \times \mathbf{W}^* = \sigma_n^2 \cdot \boldsymbol{\Theta}(\mathbf{x}). \tag{16}$$

The effective values may be now simplified to:

$$L_{\text{eff}}(\mathbf{x}) = \frac{\text{SNR}^2\, L\, \text{tr}\,(\boldsymbol{\Theta}) + (\text{tr}\,(\boldsymbol{\Theta}))^2}{\text{SNR}^2 ||\boldsymbol{\Theta}||_1 + ||\boldsymbol{\Theta}||_F^2}, \tag{17}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \sigma_n^2 \frac{\text{SNR}^2 ||\boldsymbol{\Theta}||_1 + ||\boldsymbol{\Theta}||_F^2}{\text{SNR}^2\, L + \text{tr}\,(\boldsymbol{\Theta})}, \tag{18}$$

with $\text{SNR}^2(\mathbf{x}) = \frac{A_T^2(\mathbf{x})}{L\sigma_n^2}$. For these equations, two extreme cases can be considered:

1. In the background, where no signal is present and hence SNR=0, the effective values are:

$$L_{\text{eff},B} = \frac{(\text{tr}(\boldsymbol{\Theta}))^2}{||\boldsymbol{\Theta}||_F^2} \tag{19}$$

$$\sigma_{\text{eff},S}^2 = \sigma_n^2 \frac{||\boldsymbol{\Theta}||_F^2}{\text{tr}(\boldsymbol{\Theta})}. \tag{20}$$

2. For high SNR areas, say SNR $\rightarrow \infty$:

$$L_{\text{eff},S} = L \cdot \frac{\text{tr}(\boldsymbol{\Theta})}{||\boldsymbol{\Theta}||_1} \tag{21}$$

$$\sigma_{\text{eff},B}^2 = \sigma_n^2 \frac{||\boldsymbol{\Theta}||_1}{L}. \tag{22}$$

These two cases give respectively the lower and upper bounds of $\sigma_{\text{eff}}^2(\mathbf{x})$ within the image (vice-versa for L_{eff}). Using the simplified version of the effective variance of noise in eq. (22) we can write:

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \phi_n(\mathbf{x}) \cdot \sigma_{\text{eff},B}^2 + (1 - \phi_n(\mathbf{x})) \cdot \sigma_{\text{eff},S}^2 \tag{23}$$

with

$$\phi_n(\mathbf{x}) = \frac{\text{tr}(\boldsymbol{\Theta}(\mathbf{x}))}{L \, \text{SNR}^2(\mathbf{x}) + \text{tr}(\boldsymbol{\Theta}(\mathbf{x}))}. \tag{24}$$

Note that $\phi_n(\mathbf{x})$ becomes 1 in the background (when SNR$\rightarrow 0$) and becomes 0 in high SNR areas (when SNR$\rightarrow \infty$).

The simplified model here presented is far from the *standard* stationary nc-χ generally used, and clearly very far from the stationary Rician model. If we consider results in eq. (20) and (22) we can see that the variance of noise in the background and in the signal areas will be different. If the estimation of noise is done using only the background (as it has been traditionally done) and no corrections are done, there will be a bias when used over the signal areas.

3 Noise estimation

3.1 Noise Estimation in SENSE

In the background of a SENSE MR image, where the SNR is zero, the Rician PDF simplifies to a (non-stationary) Rayleigh distribution, whose second order moment is defined as

$$\text{E}\{M^2(\mathbf{x})\} = 2 \cdot \sigma_{\mathcal{R}}^2(\mathbf{x}). \tag{25}$$

Since $\sigma_{\mathcal{R}}^2(\mathbf{x})$ is $\mathbf{x}$-dependent, $\text{E}\{M^2(\mathbf{x})\}$ will also show a different value for each $\mathbf{x}$ position. Let us assume that each coil in the $\mathbf{x}$-space is initially corrupted with uncorrelated Gaussian noise with the same variance σ_n^2 and there is a correlation between coils ρ so that matrix $\boldsymbol{\Sigma}$ becomes

$$\mathbf{\Sigma} = \sigma_n^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} = \sigma_n^2 \left(\mathbf{I} + \rho[\mathbf{1} - \mathbf{I}] \right).$$

with $\mathbf{I}$ the $L \times L$ identity matrix and $\mathbf{1}$ a $L \times L$ matrix of 1's. For each $\mathbf{x}$ value, we define the global map

$$\mathcal{G}_{W_i} = \mathbf{W}_i^* \left(\mathbf{I} + \rho[\mathbf{1} - \mathbf{I}] \right) \mathbf{W}_i, \quad i = 1, \cdots, r$$

Global map $\mathcal{G}_W(\mathbf{x})$ can be easily infered from the $\mathcal{G}_{W_i}$ values. Note that $\mathcal{G}_W(\mathbf{x})$ is strongly related to the g-factor [52]. Eq. (25) then becomes

$$E\{M^2(\mathbf{x})\} = 2\,\sigma_n^2\,\mathcal{G}_W(\mathbf{x}) \tag{26}$$

and

$$\sigma_n^2 = \frac{E\{M^2(\mathbf{x})\}}{2\,\mathcal{G}_W(\mathbf{x})} \tag{27}$$

By using this regularization, we can assure a single σ_n^2 value for all the points in the image. Following the noise estimation philosophy in [8, 47], we can now define a noise estimator based on the local sample estimation of the second order moment:

$$\langle M^2(\mathbf{x}) \rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} M^2(\mathbf{p}),$$

with $\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$. $\langle M^2(\mathbf{x}) \rangle_{\mathbf{x}}$ is known to follow a Gamma distribution [47] whose mode is $2\sigma_n^2(|\eta(\mathbf{x})| - 1)/|\eta(\mathbf{x})|$. Then

$$\text{mode}\left\{ \frac{\langle M_L^2 \rangle_{\mathbf{x}}}{\mathcal{G}_W(\mathbf{x})} \right\} = 2\sigma_n^2 \frac{|\eta(\mathbf{x})| - 1}{|\eta(\mathbf{x})|} \approx 2\sigma_n^2$$

when $|\eta(\mathbf{x})| >> 1$. The estimator is then defined as

$$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\left\{ \frac{\langle M_L^2(\mathbf{x}) \rangle_{\mathbf{x}}}{\mathcal{G}_W(\mathbf{x})} \right\} \tag{28}$$

and consequently the noise in each pixel is estimated as

$$\widehat{\sigma_{\mathcal{R}}^2}(\mathbf{x}) = \frac{1}{2}\text{mode}\left\{ \frac{\langle M_L^2(\mathbf{x}) \rangle_{\mathbf{x}}}{\mathcal{G}_W(\mathbf{x})} \right\} \mathcal{G}_W(\mathbf{x}) \tag{29}$$

This estimator is only valid over the background pixels. However, as shown in [8, 47], no segmentation of these pixels is needed: the use of the mode allows us to work with the whole image. On the other hand, to carry out the estimation, the sensitivity map of each coil and the correlation between coils must be known be-

forehand. These parameters are needed for the SENSE encoding, and thus, they can be easily obtained.

3.2 Noise estimation in GRAPPA

The background area of a GRAPPA reconstructed image may be approximated by a c-χ distribution, whose second order moment is defined as

$$\text{E}\{M_L^2\} = 2\sigma_n^2 L. \tag{30}$$

Effective parameters $L_{\text{eff}}(\mathbf{x})$ and $\sigma_{\text{eff}}^2(\mathbf{x})$ must me taken into account. Since both are $\mathbf{x}$-dependent, $\text{E}\{M_L^2\}$ will also show a different value for each $\mathbf{x}$ position:

$$\begin{aligned} \text{E}\{M_L^2\}(\mathbf{x}) &= 2\,\sigma_{\text{eff}}^2(\mathbf{x})\,L_{\text{eff}}(\mathbf{x}) \\ &= 2\,\text{tr}\left(\mathbf{C}_X^2(\mathbf{x})\right) \end{aligned}$$

and assuming the simplifications proposed in section 2.3:

$$\text{E}\{M_L^2\}(\mathbf{x}) = 2\,\sigma_n^2\,\text{tr}\left(\boldsymbol{\Theta}(\mathbf{x})\right).$$

In order to estimate a possible value of σ_n^2 matrices $\mathbf{W}(\mathbf{x})$ (the GRAPPA weights) must be known before hand. In addition, some assumption must be also made over covariance matrix $\boldsymbol{\Sigma}$. One possible assumption is the same correlation between all coils, as done in SENSE:

$$\boldsymbol{\Sigma} = \sigma_n^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} = \sigma_n^2\left(\mathbf{I} + \rho[\mathbf{1} - \mathbf{I}]\right).$$

or, in a much simplified case, no correlations between coils, $\boldsymbol{\Sigma} = \sigma_n^2\,\mathbf{I}$. In any case, from eq. (30) we can always derive

$$\sigma_n^2 = \frac{\text{E}\{M_L^2\}(\mathbf{x})}{2\,\text{tr}\left(\boldsymbol{\Theta}(\mathbf{x})\right)} \tag{31}$$

Following the same noise estimation philosophy proposed for SENSE, we can define a noise estimator based on the local sample estimation of the second order moment:

$$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\left\{\frac{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}}{\text{tr}\left(\boldsymbol{\Theta}(\mathbf{x})\right)}\right\} \tag{32}$$

This estimator is only valid over the background pixels. However, as showed in [8, 47], no segmentation of these pixels is needed. On the other hand, to carry out the estimation, the GRAPPA interpolation weights must be known beforehand.

3.3 Estimation of effective values in GRAPPA

Although many methods and applications based on the nc-χ use only the σ_n^2 value, there are other situations in which the effective value of noise is needed. Note that this effective value will now be **x**-dependent.

Assuming that we know the GRAPPA weights beforehand, we can use the estimation $\widehat{\sigma_n^2}$ in eq. (32) to estimate $\widehat{\sigma_{n,B}^2}$ and $\widehat{\sigma_{n,S}^2}$, using eq. (20) and eq. (22) respectively. These two values give the lower and upper bounds of the actual $\sigma_{\text{eff}}^2(\mathbf{x})$ across the image. Using the simplified version of the effective variance of noise in eq. (23):

$$\widehat{\sigma^2}_{\text{eff}}(\mathbf{x}) = \widehat{\phi_n}(\mathbf{x}) \cdot \widehat{\sigma^2}_{\text{eff},B} + (1 - \widehat{\phi_n}(\mathbf{x})) \cdot \widehat{\sigma^2}_{\text{eff},S} \tag{33}$$

A rough estimation of $\phi_n(\mathbf{x})$ can be done using the sample second order moment (although more complex estimation could also be considered). Since

$$\mathrm{E}\{M_L^2\}(\mathbf{x}) = A_T^2 + 2\,\sigma_n^2\,\mathrm{tr}\left(\mathbf{\Theta}(\mathbf{x})\right).$$

we can write

$$\phi_n = \frac{\mathrm{tr}\left(\mathbf{\Theta}\right)}{\frac{A_T^2}{\sigma_n^2} + \mathrm{tr}\left(\mathbf{\Theta}\right)}$$

$$= \frac{\mathrm{tr}\left(\mathbf{\Theta}\right)\,\sigma_n^2}{A_T^2 + \mathrm{tr}\left(\mathbf{\Theta}\right)\,\sigma_n^2}$$

Therefore, a simple estimation would be

$$\widehat{\phi_n}(\mathbf{x}) = \frac{\mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}}{\langle M^2 \rangle - \mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}}. \tag{34}$$

Finally, the estimated effective noise variance becomes:

$$\widehat{\sigma^2}_{\text{eff}}(\mathbf{x}) = \widehat{\sigma_n^2}\left[\frac{\mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}}{\langle M^2 \rangle - \mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}} \cdot \frac{||\mathbf{\Theta}||_1}{L} + \left(1 - \frac{\mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}}{\langle M^2 \rangle - \mathrm{tr}\left(\mathbf{\Theta}\right)\,\widehat{\sigma_n^2}}\right) \cdot \frac{||\mathbf{\Theta}||_F^2}{\mathrm{tr}\left(\mathbf{\Theta}\right)}\right].$$
$$\tag{35}$$
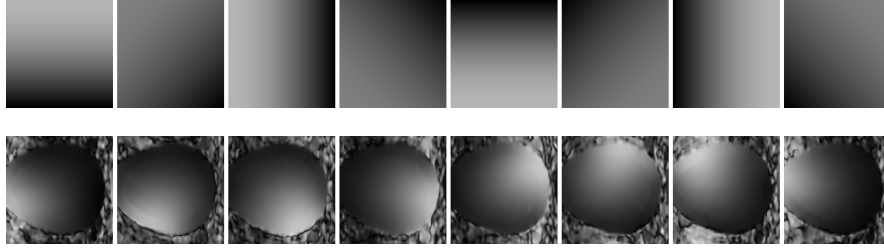
Fig. 1 Sensitivity Maps used for the experiments. Top: synthetic sensitivity map. Bottom: Map estimated from real acquisition.
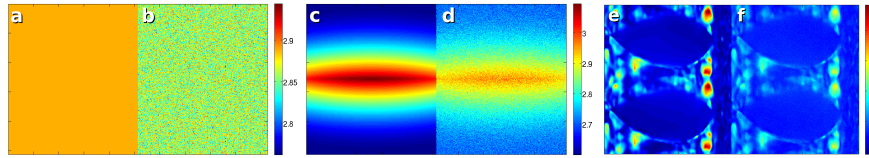
Fig. 2 Maps of $\sigma_{\mathcal{R}}^2(\mathbf{x})$ in the final image: (a-c-e): Theoretical values. (b-d-f): Estimated from samples. (a-b) Synthetic Sensitivity Map with no correlation. (c-d) Synthetic Sensitivity Map with correlation between coils. (e-f) Real sensitivity map with correlation between coils (log scale).

4 Experiments and Results

For the sake of validation of the noise estimators proposed, some experiments are carried out. We will focus first in SENSE and later in GRAPPA.

4.1 Noise estimation in SENSE

We will **first** test the variation of parameter $\sigma_{\mathcal{R}}^2(\mathbf{x})$ across the image in SENSE. To that end, we work with two sensitivity maps belonging to 8-coil systems as shown in Fig. 1: one synthetic sensitivity map (top) and a real map (bottom), estimated from a T1 acquisition done in a GE Signa 1.5T EXCITE, FSE pulse sequence, 8 coils, TR=500msec, TE=13.8msec, 256×256 and FOV: 20cm×20cm. For the sake of simplicity we assume a normalized variance at each coil $\sigma_l^2 = 1$ since it will not affect the experiment. We will simulate two different configurations, first, assuming that there is no initial correlation between coils, and second, assuming a correlation coefficient of $\rho = 0.1$. From the data, and using the theoretical expression in eq. (9) we calculate the variance of noise for each pixel in the final image. In order to test the theoretical distributions, 5000 samples of 8 complex 256×256 Gaussian images with zero mean and covariance matrix $\mathbf{\Sigma}$ are generated. The **k**-space of the data is subsampled by a 2x factor and reconstructed using SENSE and the synthetic sensitivity field. We estimate the variance of noise in each point using the second

order moment of the Rayleigh distribution [47]:

$$\sigma_{\mathcal{R}}^2(\mathbf{x}) = \frac{1}{2}E\{M^2(\mathbf{x})\}.$$

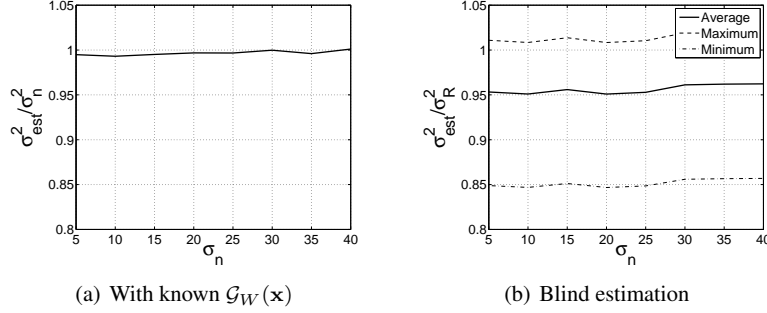We estimate the $E\{M^2(\mathbf{x})\}$ along the 5000 samples.

Visual results are depicted in Fig 2. For the synthetic maps, when no correlations are considered, the final variance of noise will not depend on the position $\mathbf{x}$. Therefore, in this particular case $\sigma_{\mathcal{R}}^2(\mathbf{x}) = \sigma_{\mathcal{R}}^2$. The estimated values in Fig 2-(b) show a noise pattern that slightly varies around the real value (note the small range of variation). In this very particular case, the noise can be considered to be spatially stationary, and the final image (leaving the correlation between pixels aside) is equivalent to one obtained from a single-coil scanner.

When correlations are taken into account, even using the same synthetic sensitivity map, results differ. In Fig. 2-(c), the theoretical value shows that the standard deviation of noise of the reconstructed data is not the same for every pixel, i.e., the noise is no longer spatial-stationary. The center of the image shows a larger value that decreases going north and south. So, in this more realistic case, the $\sigma_{\mathcal{R}}^2(\mathbf{x})$ will depend on $\mathbf{x}$, which can have serious implications for future processing, such as model based filtering techniques. The estimated value in Fig. 2-(d) shows exactly the same non-homogeneous pattern across the image. In the last experiment, Fig. 2-(e) and Fig. 2-(f), a real sensitivity map is used, and correlation between coils is also assumed. Again, the noise is non-stationary. To increase the dynamic range of the images, the logarithm has been used to show the data.

Secondly, we will validate the noise estimation capability of the proposed method by carrying out an experiment with a 2D synthetic slice from a BrainWeb MR volume [53], with intensity values in $[0 - 255]$. The average intensity value for the White Matter is 158, for the Gray Matter is 105, for the cerebrospinal fluid 36 and 0 for the background. An 8-coil system is simulated using the artificial sensitivity in Fig. 1. Image in each coil is corrupted with additive circular complex Gaussian noise with std σ_n ranging in $[5 - 40]$ and $\rho = 0.1$ between all coils. The $\mathbf{k}$-space is uniformly subsampled by a factor of 2 and reconstructed using SENSE. Note that the variance of noise of the subsampled images in each coil is amplified by a factor r [32]: $(\sigma_n^2)_{\text{sub}} = r \times \sigma_n^2$.

Results for the experiment are shown in Fig 3-(a): the average of the 100 experiments divided by the actual value of σ_n^2 is depicted. Accordingly, the closer to 1, the better the estimation. From the figure it can be seen that the estimation is very accurate for all the considered values of σ_n. The estimation is similar to the one carried out for single coil data in [8]. However, the goodness of the estimation lies in the fact that the sensitivity maps are available. We repeat the estimation assuming that the maps are not available, and considering a single $\sigma_{\mathcal{R}}^2$ value for the whole image:

$$\widehat{\sigma_{\mathcal{R}}^2} = \frac{1}{2}\text{mode}\left\{\langle M_L^2(\mathbf{x})\rangle_{\mathbf{x}}\right\} \tag{36}$$

(a) With known $\mathcal{G}_W(\mathbf{x})$ (b) Blind estimation

Fig. 3 Estimation of the variance of noise from SENSE. The average of 100 experiments is considered.
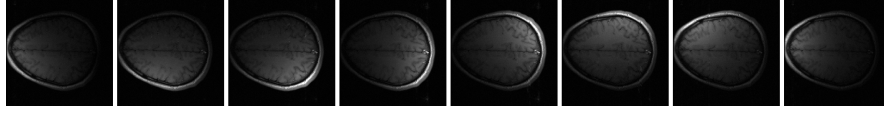
Fig. 4 Slice from a brain T1 acquisition done in a GE Signa 1.5T EXCITE with 8 coils.

We define the ratio $\widehat{\sigma_{\mathcal{R}}^2}/\sigma_{\mathcal{R}}^2(\mathbf{x})$ and we calculate the average, the minimum and maximum values across the image, and the average along 100 samples. Results are depicted in Fig 3-(b). The estimated value presents a constant bias of around 5% for all values. The estimated value will be in a range from 85% to 100% of the original value. Hence, if $\mathcal{G}_W(\mathbf{x})$ is unknown, estimating an individual value of σ_n^2 will only be acceptable for certain applications, whenever they are robust enough to cope with a bit deal of bias and a higher deal of uncertainty in this parameter.

Finally, an experiment is carried out with data from a real acquisition, see Fig. 4, with sensitivity map in Fig. 1-bottom. First, as a golden standard, parameter σ_n is estimated from the Gaussian complex data:

Real component	$\widehat{\sigma_n} = 4.1709$
Imag. component	$\widehat{\sigma_n} = 4.0845$

Then a subsampled acquisition is simulated and reconstructed with SENSE. σ_n is first estimated using eq. (32) and then, assuming the map $\mathcal{G}_W(\mathbf{x})$ is unknown, using eq. (36). Results are as follows:

Magnitude ($\mathcal{G}_W(\mathbf{x})$ known)	$\widehat{\sigma_n}/\sqrt{r} = 4.1728$
Magnitude ($\mathcal{G}_W(\mathbf{x})$ unknown)	$\widehat{\sigma_n}/\sqrt{r} = 4.8404$

Note that the value estimated using the proposed method is totally consistent with the estimation done over the original complex Gaussian data. The blind estimation method, on the other hand, overestimates the noise level. This is caused because in

eq. (29) the map given by $\mathcal{G}_W(\mathbf{x})$ is basically a normalization. The lack of knowledge of this parameter displaces the mode of the distribution from its actual value, hence the mismatch. However, for some applications in which a great accuracy is not needed, there could still be a valid value that gives a rough approximation to the variance of noise.

4.2 Noise estimation in GRAPPA

For the sake of validation, several experiments are considered. **First**, synthetic experiments were carried out using the same 2D synthetic slice from a BrainWeb MR volume used for SENSE. Image in each coil is again corrupted with Gaussian noise with std σ_n ranging in $[5-40]$ and $\rho = 0$. The **k**-space is uniformly subsampled by a factor of 2, keeping 32 ACS lines. The CMS is reconstructed using GRAPPA and SoS. The sample local moments have been calculated using 7×7 neighborhoods. Two different cases are considered in the simulation, 4 and 8 coils.

| (a) Mean | (b) Standard deviation |
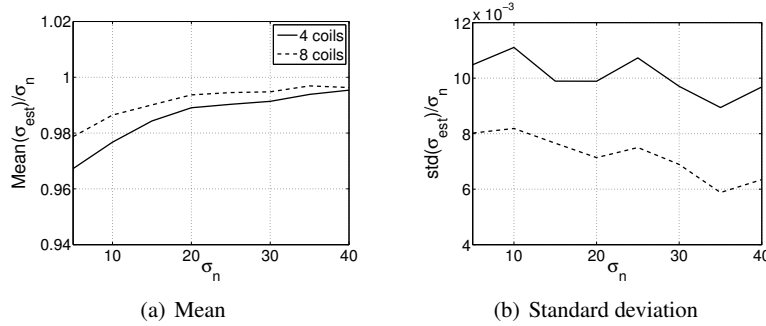
Fig. 5 Results of σ_n estimation using the proposed method; 100 experiments are considered for each sigma value. (a) Mean of the estimated value divided by the actual value. (b) Standard deviation of the estimated values.

Results for the experiment are shown in Fig. 5: in Fig 5-(a), the mean of the 100 experiments divided by the actual value of σ_n is depicted. Accordingly, the closer to 1, the better the estimation; in Fig 5-(b), the standard deviation of the experiments divided by the actual value is shown; the lower the value, the better the estimation.

From the results it can be seen that the estimation is very accurate, although a small bias appears for low values of σ_n. This bias is surely motivated by a mismatch between the GRAPPA reconstructed image and the nc-χ model: according to [42] the error of approximating the CMS by a nc-χ is larger for very low σ_n values. All in all, the proposed method shows a very good average behavior –the values are in a small range between 0.97 and 1–, with a small biased mean and a very low variance, which assures a consistent estimation.
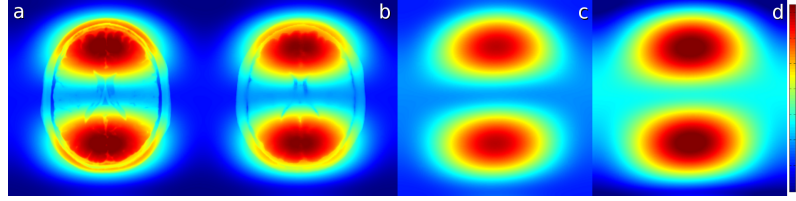
Fig. 6 Effective standard deviation of noise: (a) Original $\sigma_{\text{eff}}(\mathbf{x})$, derivated from the GRAPPA weights and eq. (14); (b) Estimated $\widehat{\sigma}_{\text{eff}}(\mathbf{x})$ from eq. (35); (c) Estimation of effective std of noise for SNR=0, $\widehat{\sigma}_{\text{eff,B}}(\mathbf{x})$; (d) Estimation of effective std of noise for high SNR, $\widehat{\sigma}_{\text{eff,S}}(\mathbf{x})$.
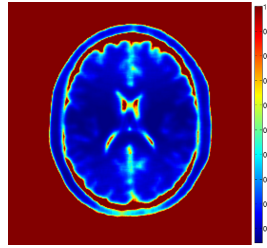
Fig. 7 Estimation of correction factor $\widehat{\phi_n}(\mathbf{x})$ from eq (34).
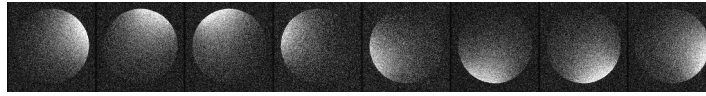
Fig. 8 Slice of an 8-coil 2D acquisition of the phantom used for the experiments.

For the sake of illustration, the map of the effective values of noise is also calculated for one single experiment with $\sigma_n = 10$. For that experiment, the theoretical value of $\sigma_{\text{eff}}^2(\mathbf{x})$ is calculated using eq. (14). From the expression in eq. (35), using the estimated noise $\widehat{\sigma^2}_n$ and the GRAPPA weights coded in $\boldsymbol{\Theta}$, the variance of noise for the two extreme cases (SNR=0 and high SNR) are estimated. Using the correction factor $\widehat{\phi_n}(\mathbf{x})$, a global value for $\widehat{\sigma^2}_{\text{eff}}(\mathbf{x})$ is obtained.

Results are depicted in Fig. 6-(a) $(\sigma_{\text{eff}}(\mathbf{x}))$; Fig. 6-(b) $(\widehat{\sigma}_{\text{eff}}(\mathbf{x}))$; Fig. 6-(c) $\widehat{\sigma}_{\text{eff,B}}(\mathbf{x})$; Fig. 6-(d) $\widehat{\sigma}_{\text{eff,S}}(\mathbf{x})$. The correction factor $\widehat{\phi_n}(\mathbf{x})$ is depicted in Fig. 7.

From the illustrations it is easy to see that the variance of noise has a high variation itself across the image. $\sigma_{\text{eff}}(\mathbf{x})$ ranges from 10 to 45. Even inside the same tissue, there is a huge variation (from 25 to 45). There is, also, a high mismatch between the head and the background areas. Some interesting conclusions can be raised from this: (1) The assumption of a single σ_n^2 value for the whole volume does not hold in GRAPPA. Assuming this single value will clearly bias any further processing; (2) In this example, the noise values in the background are much smaller than those inside the tissue. If the background is used to estimate the noise, and no correction is applied, there can be a huge mismatch between the real noise and the estimated value.

Fig. 9 Values of map $\mathrm{tr}(\mathbf{\Theta})(\mathbf{x})$ from the GRAPPA reconstruction coefficients of the second experiment.

For the **second** experiment, real acquisitions are considered. 100 repetitions of the same slice of a phantom, scanned in an 8-channel head coil on a GE Signa 1.5T EXCITE 12m4 scanner with FGRE Pulse Sequence to generate low SNR, see Fig. 3. Matrix size= 128×128, TR/TE=8.6/3.38 ms, FOV 21x21cm, slice thickness = 1mm. Noise variance σ_n^2 is initially estimated using the variance of the real part of every coil of every sample, where the noise is known to be additive Gaussian [108]. This value σ_0^2 is taken as Golden Standard. Then, all the 100 samples are 2× subsampled. The GRAPPA reconstruction coefficients are derived from one sample, using 32 ACS lines, and used for interpolation in all samples. The CMS is obtained by SoS. Noise is estimated over each CMS using eq. (32). For the sake of illustration, values for $\mathrm{tr}(\mathbf{\Theta})(\mathbf{x})$ derived from the GRAPPA coefficients are depicted in Fig. 9. Estimation results are as follows:

σ_0	mean$\{\widehat{\sigma_n}\}$	mean$\{\widehat{\sigma_n}\}/\sigma_0$	std$\{\widehat{\sigma_n}\}/\sigma_0$
0.0428	0.0424	0.9905	0.0113

Results obtained estimating the noise with the proposed method is totally consistent with the value obtained over the complex Gaussian images without subsamplig. There is a very small bias in the estimation and the method also shows a very small variance, as also seen in the synthetic experiments. The map of $\mathrm{tr}(\mathbf{\Theta})(\mathbf{x})$ depicted in Fig. 9 shows that, in this real case, there is also a great variation of the noise parameter across the image.

Finally, for the sake of comparison with SENSE estimation, a new experiment is carried out with the data from the real acquisition in Fig. 4, as a golden standard for parameter σ_n the estimation from the Gaussian complex already done for SENSE ($\widehat{\sigma_n} = 4.1709$ for the real component.) The complex data is subsampled with $r = 2$. The **k**-space is reconstructed using GRAPPA and 32 ACS lines and the CMS is obtained by SoS. Noise is estimated over the CMS using eq. (32). Two different estimations have been done: (1) using the GRAPPA coeffficients; (2) assiming the coefficients unknown. In the last case, matrix $\mathbf{\Theta}(\mathbf{x})$ is replaced by a 8×8 identity matrix.

Results are as follows:

Magnitude (GRAPPA coefficients known)	$\widehat{\sigma_n} \times \sqrt{r} = 4.1097$
Magnitude (GRAPPA coefficients unknown)	$\widehat{\sigma_n} \times \sqrt{r} = 5,1933$

Again, like in SENSE; the value estimated using the proposed method is consistent with the estimation done over the original complex Gaussian data. The blind estimation method, on the other hand, overestimates the noise level. Note that there is a great lack of knowledge of a normalization level, hence the error. However, note that can still be valid to estimate the order of magnitude of the variance of noise, or in case a rough estimation is needed.

5 Conclusions

The proper modeling of the statistics of thermal noise in MRI is crucial for many image processing and computer aided diagnosis tasks. While the stationary Rician and nc-χ models have been the keystone of statistical signal processing in MR for years, the stationarity assumption cannot be applied when parallel imaging reconstruction is considered: the main assumption of a single value of σ_n^2 to characterize the whole data set is no longer valid. When pMRI techniques are used, due to the reconstruction process, the variance of noise becomes **x**-dependent, with a different value for each pixel.

To overcome the problems of non-stationarity, we have proposed a novel noise estimation technique to be used with SENSE and GRAPPA reconstructed data. The estimation of the spatially variant $\sigma_n^2(\mathbf{x})$ is of paramount importance, since the knowledge of this parameter will allow us to re-use many of the methods proposed in literature for stationary models. In most cases it will suffice with changing an scalar σ_n^2 value by the spatially dependent $\sigma_n^2(\mathbf{x})$.

The estimation methods proposed have shown to be accurate, robust and easy to use. However, it also shows some limitations. First, correlation between coils must be known beforehand, as well as the sensitivity map from each coil (in SENSE) or the reconstruction weights (in GRAPPA). Finally, some post processing software in the scanner may add a mask to data, which eliminates part of the background, drastically reducing the number of points available for noise estimation [107]. The estimation method selected must be properly adjusted to this problem. Note that if the background is totally removed, the estimation should be done using methods that do not rely on the background, but on the signal areas.

Acknowledgements

Blind estimation of spatially variant noise in MRI

Santiago Aja-Fernández, Tomasz Pieciak, Gonzalo Vegas-Sánchez-Ferrero*

Abstract The reliable estimation of noise characteristics in MRI is a task of great importance due to the influence of noise features in extensively used post-processing algorithms. Many methods have been proposed in the literature to retrieve noise features from the magnitude signal. However, most of them assume a stationary noise model, i.e., the features of noise do not vary with the position inside the image. This assumption does not hold when modern scanning techniques are considered, e.g., in the case of parallel reconstruction and intensity correction. Therefore, new noise estimators must be found to cope with non-stationary noise. Some methods have been recently proposed in the literature. However, they require multiple acquisitions or extra information which is usually not available (biophysical models, sensitivity of coils). In this work we overcome this drawback by proposing a new method that can accurately estimate the non-stationary parameters of noise from just a single magnitude image. In the derivation, we considered the noise to follow a non-stationary Rician distribution, since it is the most common model in real acquisitions (e.g., SENSE reconstruction), though it can be easily generalized to other models. The proposed approach makes use of a homomorphic separation of the spatially variant noise in two terms: a stationary noise term and one low frequency signal that correspond to the x-dependent variance of noise. The non-stationary variance of noise is then estimated by a low pass filtering with a Rician bias correction. Results in real and synthetic experiments evidence the better performance and the lowest error variance of the proposed methodology when compared to the state-of-the-art methods.

---

1 Introduction

Noise is known to be one of the most common sources of deterioration of the quality of Magnetic Resonance Imaging (MRI) data. The principal source of noise in most MR scans is the subject or object to be imaged, followed by electronics noise during the acquisition of the signal in the receiver chain. It is produced by the stochastic motion of free electrons in the RF coil, which is a conductor, and by eddy current losses in the patient, which are inductively coupled to the RF coil. The presence of noisy patterns on the acquired MR signal is a problem that affects not only the visual quality of the images, but also may interfere with further processing techniques such as registration, fMRI analysis or tensor estimation in Diffusion Tensor MRI ([4, 5, 6, 157, 25]). The accurate modeling of signal and noise statistics in MR data usually underlies the tools for processing and interpretation within MRI.

The stationary Rician distribution ([5]) has been widely accepted in literature as a suitable model for noise in MR magnitude images. Many authors have precisely introduced the Rician statistics in the estimation of diffusion models ([55]), curve fitting for quantitative perfusion measurements ([56, 58]), hypothesis tests to assess the activation level in functional MRI ([24]), in a preprocessing step to remove the bias of the data for the subsequent processing stages ([48]) and denoising techniques ([9, 12, 13]).

The main assumption for single coil Rician acquisitions is that the noise is stationary, and therefore a single value of σ characterizes the whole data set. However, this premise will mostly fail when considering modern scanners with multiple-coil antennae and scanning software to correct artifacts, and to improve the final appearance of the image. Linear operations carried out over the complex Gaussian data modify the variance of the noise σ^2 differently for each position. As a result, the final magnitude signal will have an $\mathbf{x}-$dependent value of σ, i.e. $\sigma(\mathbf{x})$, generating a *non-stationary* (or *non-homogeneous*), distribution. For instance, this is the case of a Sum-of-Squares ([27]) reconstruction of a multiple-coil acquisition, where the composite magnitude signal can be approximated by a non-stationary non-central χ (nc-χ) distribution ([49]). It is also the case for accelerated acquisitions with parallel MRI (pMRI) reconstruction techniques. If GRAPPA (Generalized Autocalibrating Partially Parallel Acquisition, [33]) is used for reconstruction, the final noise is known to follow a non-stationary nc-χ distribution ([42, 43]), while if SENSE (Sensitivity Encoding for Fast MRI, [32]) is used, the magnitude signal may be considered Rician distributed with a spatially variant value of $\sigma(\mathbf{x})$ ([46, 43, 70]).

Nowadays, SENSE has become practically a *de facto* standard in most acquisitions. However, many processing techniques still assume the stationary Rician distribution as a model for the signal and noise, forgetting about the non-stationarity of the data. This is probably due to the fact that most noise estimators in literature are based on a single σ^2 value for all the pixels in the image, either assuming a Rician model ([7, 81, 37, 8, 47, 84]) or an nc-χ ([27, 47, 48, 49]). There have also been some proposals to carry out a rough estimation of non-stationary noise maps. However, these approaches require extra information beyond the simple magnitude signal: multiple acquisitions or different signals are required ([67, 68, 69]), a bio-

physical model must be defined ([69]), or even acquisition information such the estimated sensitivity of the coils is needed ([70]). This need of extra information has supposed a drawback in the usage of more complex noise models.

In this paper we propose a new technique that allows the estimation of the spatially variant maps of noise $\sigma(\mathbf{x})$ from the magnitude signal when only a single image is available and no additional information is required. The estimator is developed for the non-stationary Rician case, and it is complemented with the estimators for Gaussian and Rayleigh cases. The methodology here presented is totally compatible with other noise models, such as the non-stationary nc-χ distribution, and the extension would be straightforward. The initial assumption needed is that the variability of the map of noise is smaller than the variability of the noise itself, i.e., $\sigma(\mathbf{x})$ is a low frequency signal when compared to the noise, which is a rational assumption in MRI acquisitions. Both sources of variability are separated by using a homomorphic transformation ([159]). This technique allows us to improve the estimation of the map of noise while it avoids the granularity produced by most local methods.

2 Background: non-stationary noise in MRI

When dealing with MRI data, before performing any processing that may involve noise related parameters, it is necessary to identify the specific noise model present in your data. The probability model of noise in the data depends on the coil configuration of the scanner and on the kind of processing the MR data goes through before producing the final magnitude image. The purpose of this section is precisely to provide a general framework of the most usual models of noise in MR data, as well as the most common procedures that originates these distributions.

Most applications dealing with noise in MRI rely on the assumption of a single value of the variance of noise σ^2 for every pixel within the image, i.e., they assume a stationary noise model: the features of noise do not change with position. However, this is not entirely the case in modern acquisition systems, when pMRI protocols and artifact correction techniques are applied. The linear manipulation of the original Gaussian data, the combination of different coil information and adaptive processing change the features of noise differently in every location of the image. However, due to the kind of processing carried out, most of the times the Rician and nc-χ assumptions still hold, although the stationarity does not.

In what follows the procedures that lead to non-stationary noise models in MRI will be put together, and the main proposals in the literature for non-stationary noise estimation will be reviewed.

2.1 Noise model for each coil

The first step in modeling the final magnitude image in MRI is to model the noise distribution in each of the scanner coils and, then, to propagate the model along the processing pipeline.

Let us assume an L-coil antenna configuration, being L the number of coils in the system. We denote $s_l(\mathbf{k})$ the signal in the $\mathbf{k}-$space acquired by the l-th coil, which corresponds to a complex signal $S_l(\mathbf{x})$ in the image domain. Signals $s_l(\mathbf{k})$ are considered to be corrupted by Additive White Gaussian Noise (AWGN), with zero mean and variance $\sigma_{K_l}^2$:

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; \sigma_{K_l}^2), \quad l = 1, \cdots, L \tag{1}$$

with $a_l(\mathbf{k})$ the noise-free signal and $n_l(\mathbf{k}; \sigma_{K_l}^2) = n_{l_r}(\mathbf{k}; \sigma_{K_l}^2) + j \cdot n_{l_i}(\mathbf{k}; \sigma_{K_l}^2)$ the AWGN process, which is initially assumed stationary, so that $\sigma_{K_l}^2$ does not depend on $\mathbf{k}$. The complex $\mathbf{x}-$space is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice or volume, so the noise in the complex $\mathbf{x}-$space is still Gaussian ([43]):

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; \sigma_l^2), \quad l = 1, \cdots, L$$

where $N_l(\mathbf{x}; \sigma_l^2)$ is also a complex AWGN process (note that we are assuming no spatial correlations) with zero mean and covariance matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1L} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L1} & \sigma_{L2} & \cdots & \sigma_L^2 \end{pmatrix}. \tag{2}$$
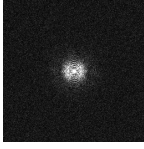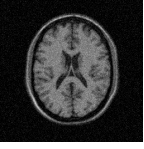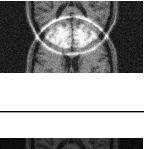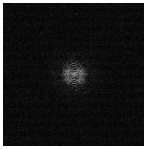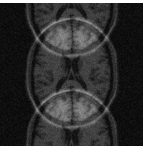
The relation between the noise variances in the $\mathbf{k}-$ and $\mathbf{x}-$domains is given by the number of points used for the iDFT:

$$\sigma_l^2 = \frac{1}{|\Omega|} \sigma_{K_l}^2, \tag{3}$$

with $|\Omega|$ the final number of pixels in the field of view (FOV).

2.2 Noise models in pMRI

In order to accelerate the acquisition rate in multiple coil systems, pMRI techniques are used. This acceleration is achieved by subsampling the $\mathbf{k}-$space data in each coil ([30, 31]), i.e., not all the frequency lines are acquired. The immediate effect of this $\mathbf{k}-$space subsampling is the appearance of aliased replicas in the image domain retrieved at each coil. In order to suppress or correct this aliasing, pMRI

Noise relations			
$\mathbf{k}-$space	Parameters	$\mathbf{x}-$space	Relation
	Fully sampled, $\sigma^2_{K_l}$ $\mathbf{k}-$size: $\|\Omega\|$		$\sigma_l^2 = \frac{1}{\|\Omega\|}\sigma^2_{K_l}$, $\mathbf{x}$-size: $\|\Omega\|$
	Subsampled r, $\sigma^2_{K_l}$ $\mathbf{k}-$size: $\|\Omega\|/r$		$\sigma_l^2 = \frac{r}{\|\Omega\|}\sigma^2_{K_l}$, $\mathbf{x}-$size: $\|\Omega\|/r$
	Subsampled r, $\sigma^2_{K_l}$ $\mathbf{k}-$size: $\|\Omega\|$ (zero padded)		$\sigma_l^2 = \frac{1}{\|\Omega\|\cdot r}\sigma^2_{K_l}$, $\mathbf{x}-$size: $\|\Omega\|$

Table 1 Relations between the variance of noise in complex MR data for each coil in the $\mathbf{k}-$space and the image domain.

combines the redundant information from several coils to reconstruct a single non-aliased image domain.

The commonly used stationary Rician and nc-χ models do not necessarily hold in this case. Depending on the way the information from each coil is combined, the statistics of the image follows different distributions. It is therefore necessary to study the behavior of the data for each particular reconstruction method.

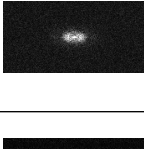Let us call $s_l^{\mathcal{S}}(\mathbf{k})$ to the subsampled signal at the l-th coil in the $\mathbf{k}-$space and $S_l^{\mathcal{S}}(\mathbf{x})$ in the image domain. The relation between the noise variances in the $\mathbf{k}-$ and $\mathbf{x}-$domains in eq. (3) depends on the kind of reconstruction used for the iDFT. If a $S_l^{\mathcal{S}}(\mathbf{x})$ is direct iDFT of $s_l^{\mathcal{S}}(\mathbf{k})$, the relation, assuming an acceleration rate of r, becomes

$$\sigma_l^2 = \frac{r}{\|\Omega\|}\sigma^2_{K_l}. \tag{4}$$

On the contrary, if the iDFT is computed after zero-padding, the missing (not sampled) $\mathbf{k}-$space lines, the relation is:

$$\sigma_l^2 = \frac{1}{\|\Omega\|\cdot r}\sigma^2_{K_l}. \tag{5}$$

Relations between the variance of noise in complex $\mathbf{x}-$space and $\mathbf{k}-$space for each coil are summarized in Table 1.

Many different methods have been defined to reconstruct the final image from subsampled versions of the signals in each coil, being SENSE and GRAPPA dominant in commercial scanners. However, new reconstruction methods and modifications of the existing ones are continuously proposed. From a statistical point of view, reconstruction methods carry out linear operations over the subsampled signals $S_l^{\mathcal{S}}(\mathbf{x})$, in order to obtain a final reconstructed magnitude image, which is the one of the main causes of the non-stationarity of noise. There are mainly two different approaches for signal reconstruction:

1. *Reconstruction of a single complex image:* The reconstruction process combines the data of the different coils with some extra information (such as the sensitivity map of each coil or the covariance matrix) to obtain a single image:

$$S^{\mathcal{R}}(\mathbf{x}) = f\left(\{S_l^{\mathcal{S}}(\mathbf{x}), l = 1, \cdots, L\}; \boldsymbol{\Theta}\right). \tag{6}$$

with $f(.)$ a linear reconstruction function (see some specific functions in [32, 34]) and $\boldsymbol{\Theta}$ any additional information needed. The linear operations over the Gaussian data generate correlated Gaussian data. However, the reconstruction affects the stationarity of the noise in the resulting image. Thus, the final signal can be seen as a reconstructed signal corrupted with Gaussian noise whose variance depends on the position:

$$S^{\mathcal{R}}(\mathbf{x}) = A^{\mathcal{R}}(\mathbf{x}) + N^{\mathcal{R}}(\mathbf{x}; \sigma_{\mathcal{R}}^2(\mathbf{x})) \tag{7}$$

where $N^{\mathcal{R}}(\mathbf{x}; \sigma_{\mathcal{R}}^2(\mathbf{x}))$ is a non-stationary complex AWGN process. The final magnitude image is obtained by using the absolute value:

$$M(\mathbf{x}) = |S^{\mathcal{R}}(\mathbf{x})| \tag{8}$$

and therefore it follows a non-stationary Rician distribution, with the parameter $\sigma_{\mathcal{R}}^2(\mathbf{x})$ being spatially variant. This is the case, for instance, of pMRI data reconstructed with SENSE in its original form.

2. *Reconstruction of multiple complex images:* The reconstruction process combines the data of the different coils to obtain a reconstructed image per coil:

$$S_l^{\mathcal{R}}(\mathbf{x}) = f_l\left(\{S_m^{\mathcal{S}}(\mathbf{x}), m = 1, \cdots, L\}; \boldsymbol{\Theta}\right), \text{ with } l = 1, \cdots, L. \tag{9}$$

with $f_l(.)$ a linear reconstruction function over each coil (see one specific function in [33]). As in the previous case, the linear operations over the Gaussian data generate non-stationary Gaussian data in each coil. The final signal in each coil can be seen as a reconstructed signal corrupted with Gaussian noise whose variance depends on the position:

$$S_l^{\mathcal{R}}(\mathbf{x}) = A_l^{\mathcal{R}}(\mathbf{x}) + N_l^{\mathcal{R}}(\mathbf{x}; \sigma_{\mathcal{R}_l}^2(\mathbf{x})), \text{ with } l = 1, \cdots, L. \tag{10}$$

The probability distribution of the final magnitude image depends on the method used to merge the information of the multiple reconstructed coils into one single

image. To avoid any extra information, one of the most common approaches is the Sum-of-Squares (SoS):

$$M_{\text{SoS}}(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l^{\mathcal{R}}(\mathbf{x})|^2}. \tag{11}$$

The distribution of the magnitude signal $M_{\text{SoS}}(\mathbf{x})$ depends on the relation of noise with the reconstructed signal in each coil, $S_l^{\mathcal{R}}(\mathbf{x})$. For a GRAPPA procedure, for instance, it can be approximated by a non-stationary nc-χ distribution ([42]). However, since it is the sum of multiple signals, the Rician distribution does no longer hold. Due to the bias that the method can introduce over the resulting signal, manufacturers are lately trying to avoid the SoS to obtain $M(\mathbf{x})$. Even for GRAPPA, some new approaches use a reconstruction similar to the first case, where data is Rician distributed.

Other common strategy is based on the spatial matched filter approach, which linearly combines the complex signals of each coil and produces voxelwise complex signals ([160]). This is the same methodology used by SENSE, and it requires extra information of the sensitivity of each coil. One simple way to implement this filtering is the following:

$$S_T^{\mathcal{R}}(\mathbf{x}) = \sum_{l=1}^{L} S_l^{\mathcal{R}}(\mathbf{x}) \cdot C_l(\mathbf{x}), \tag{12}$$

with $C_l(\mathbf{x})$ an estimation of the sensitivity map in each coil. The complex image $S_T^{\mathcal{R}}(\mathbf{x})$ now follows a complex Gaussian distribution, similar to the previous case. The magnitude signal is then obtained:

$$M_T(\mathbf{x}) = |S_T^{\mathcal{R}}(\mathbf{x})|$$

and therefore it follows a non-stationary Rician distribution, with the parameter $\sigma_{\mathcal{R}}^2(\mathbf{x})$ being spatially variant.

2.3 Extrapolation to other data processing

Parallel reconstruction methods are not the only source of non-stationarity of the data. Even when no acceleration is present, certain *correction* and *preprocessing* techniques can make the features of noise spatially variant. A deep study of these techniques is beyond the scope of this paper. However, as an illustration of other cases in which the same methodology can be used, two examples are presented:

- *Definition of composite magnitude image from multiple coils:* even when non acceleration is considered, if the acquisition is done in multiple coils, complex signals from each coil have to be merged into one single final image. If the SoS

of eq. (3) is used, due to the correlation between coils, the final image can be approximated by a non-stationary nc-χ distribution with adaptive parameters ([49]). On the contrary, if the matched filter in eq. (12) is used, the final $M(\mathbf{x})$ follows a Rician distribution, and the variance of noise may become spatially variable due to the estimated sensitivity coils.

- *Spatial linear processing with space dependent weights:* any linear processing that uses spatial-dependent features over the complex Gaussian data also change the spatial features of noise. Assume, for instance, an illumination correction in a single coil signal:

$$S_F(\mathbf{x}) = S(\mathbf{x}) \cdot C^{-1}(\mathbf{x}),$$

where $C(\mathbf{x})$ is an estimation of the image illumination, related to the coil sensitivity. Note that the final image $S_F(\mathbf{x})$ will remain Gaussian, but the variance of noise is affected by the following spatial correction:

$$\sigma_F^2(\mathbf{x}) = \sigma^2 \cdot (C^{-1}(\mathbf{x}))^2.$$

2.4 Non-stationary noise estimation

Although most of the noise estimators in literature cope with the problem of a single value of σ for the whole image ([7, 81, 37, 8, 47, 84]), there have been some attempts to estimate non-homogeneous maps of noise, not only in the MRI context. The most usual approaches are based on wavelet decomposition or on multiple acquisitions

One of the first attempts to estimate spatially variant noise in images is due to [71]. Authors estimate the spatially variant map of noise in images assuming they are corrupted by a non-stationary AWGN process. To separate the signal and the noise, they use a wavelet transform, assuming that the high-high subband is strictly noise. In [98], authors propose the joint estimation of non-homogeneous noise and signal in Rician data, using an expectation-maximization (EM) algorithm to find the maximum likelihood (ML) estimate for the parameters in synthetic aperture radar images. To carry out the algorithm, multiple samples of the receiving signal are necessary.

In MRI, [72] proposed a method to estimate spatially variant noise by suppressing the signal component. To that end, the stationary wavelet transform of the magnitude image at the scale $s = 1$ is calculated and the low-low subband coefficients are removed. The estimation is done assuming that the remaining signal is only noise following a Rayleigh distribution. This assumption, however, is not entirely true, as can be experimentally proved. A similar approach, also based in image removal to get an *only noise* image is carried out by [100]. They model the variance of noise as a minimal distance between local neighborhood (patch) of the current pixel and the remaining patches in the non-local means filtering scheme. The estimate is corrected for low Signal-to-Noise Ratio (SNR) regions applying the technique described in [48]. Thus, a SNR iterative estimation is needed.

The median absolute deviation (MAD) estimator for stationary Rician noise estimation proposed in [78] has been extended by some authors to non-stationary noise using a *local* version of the MAD, see for instance [68] and [74], with very similar approaches, and [67] with a very effective technique that needs information of multiple diffusion weighted images of the same slice to carry out the estimation. All these methods also need an estimation of the SNR.

An alternative technique was proposed in [69], based in a Q_n estimator followed by a regularization procedure using coil sensitivity model. Although this method has been proposed to cope with multiple independent MR scans, in its basic scenario it can be used to estimate the noise map on a single image.

Other significant methods are the following: [99] proposed a local variance as a noise level estimator after edges exclusion by means of local mutual information and k-means segmentation. This method suffers from edges overestimation, thus, a mathematical morphology filter is applied to suppress this undesirable effect; [102] calculate the noise map from the receiver coil noise matrix, which, in fact, is not always available in a clinical routine; [101] proposed a method to assess temporary random noise in dynamic MR image series, e.g., cardiac function imaging or blood flow velocity mapping.

Finally, in [70], authors propose a method to estimate the variable noise maps in MRI assuming SENSE and GRAPPA reconstruction. Although the results are precise, the drawback of this method is that some prior information about the reconstruction process is needed: namely the sensitivity map estimated for each coil (for SENSE) or the GRAPPA reconstruction coefficients. Although these parameters can be obtained from the scanner, they are not always available in clinical acquisitions. This limitation, present in most of the reviewed methods, is the main motivation of this current work.

3 A homomorphic approach to non-stationary noise estimation

In what follows we will assume that, due to the processing done in the scanner, the final image $I(\mathbf{x})$ is corrupted with noise whose variance is $\mathbf{x}-$dependent. For the sake of generality and simplicity, no specific pMRI reconstruction method will be considered. We will also assume that only one single 2D image is available and no extra parameters are known. Note that, unlike our proposal, some of the estimation methods proposed in the literature are based on the availability of multiple repetitions or information about the reconstruction process. The methodology presented is based on the initial assumption that the spatial dependent $\sigma^2(\mathbf{x})$ shows a low variability, i.e., it can be considered a low pass signal. Therefore, it can be separated from the noise pattern that multiplies it. This is a rational assumption in MRI acquisitions.

In the following study, three different cases will be considered: Gaussian, Rayleigh and Rician. Although the latest is the most suitable model for MRI, the first one provides a good alternative for high SNR. It also presents a solid alternative to an

automatic estimator built-in inside the scanning software before the magnitude is calculated. Finally, for the sake of completeness, the Rayleigh case is also considered since it is the lower boundary for the Rician case when the SNR tends to zero.

3.1 The Gaussian case

Let us first assume a simple case in which an image $A(\mathbf{x})$ is corrupted with additive Gaussian noise with zero mean and spatially-dependent variance $\sigma^2(\mathbf{x})$:

$$
\begin{aligned}
I(\mathbf{x}) &= A(\mathbf{x}) + N(\mathbf{x}; 0, \sigma^2(\mathbf{x})) \\
&= A(\mathbf{x}) + \sigma(\mathbf{x}) \cdot N(\mathbf{x}; 0, 1).
\end{aligned}
\tag{13}
$$

Our purpose is to estimate $\sigma(\mathbf{x})$ from the final image $I(\mathbf{x})$. To that aim, we use a homomorphic filtering that will extract the spatially variant pattern of noise.

Let us assume that the variance of noise $\sigma^2(\mathbf{x})$ slowly varies across the image, i.e. it is a low frequency signal. We remove the mean of the image to avoid any contribution of $A(\mathbf{x})$:

$$
I_n(\mathbf{x}) = I(\mathbf{x}) - E\{I(\mathbf{x})\} = \sigma(\mathbf{x}) \cdot N(\mathbf{x}).
$$

where $E\{I(\mathbf{x})\}$ denotes the expectation value in each point of the image, i.e. the *local* mean, so that $E\{I(\mathbf{x})\} = A(\mathbf{x})$. Next, we separate signals $\sigma(\mathbf{x})$ and $N(\mathbf{x})$ by applying the logarithm:

$$
\log |I_n(\mathbf{x})| = \underbrace{\log \sigma(\mathbf{x})}_{\text{low frequency}} + \underbrace{\log |N(\mathbf{x})|}_{\text{higher frequency}}.
$$

The noise term $\log |N(\mathbf{x})|$ has its energy distributed along all frequencies, while the term $\log \sigma(\mathbf{x})$ is a low frequency signal by hypothesis. The latest can be recovered using a low pass filtering of $\log |I_n(\mathbf{x})|$:

$$
\text{LPF}\{\log |I_n(\mathbf{x})|\} \approx \log \sigma(\mathbf{x}) + \delta_N
\tag{14}
$$

with δ_N being a low pass residue of $\log |N(\mathbf{x})|$. This residue must be calculated to remove it from the estimation. If we assume that the LPF has a small passband, the LPF $\{\log |N(\mathbf{x})|\}$ is a good estimator of the local mean. By hypothesis, $N(\mathbf{x})$ is stationary, and therefore the mean is the same for all pixels. Thus, we can consider the LPF as a good approximation of the mean of the signal:

$$
\text{LPF}\{\log |N(\mathbf{x})|\} \approx E\{\log |N(\mathbf{x})|\}.
$$

Since we know that $N(\mathbf{x})$ follows a Gaussian distribution (with zero mean and unitary variance), then $|N(\mathbf{x})|$ follows a *half-normal* distribution, and the mean of $\log |N(\mathbf{x})|$ can be written as
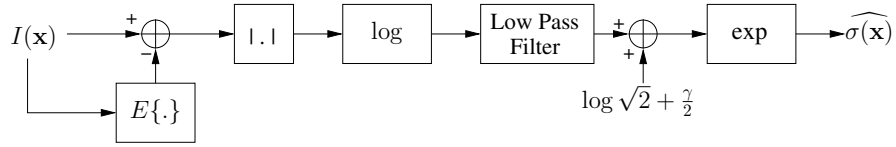
Fig. 1 Pipeline of $\sigma(\mathbf{x})$ estimation assuming Gaussian noise. $I(\mathbf{x})$ is the original image, $E\{.\}$ is the local expected value of the signal and $|.|$ is the absolute value.

$$E\left\{\log|N(\mathbf{x};0,\sigma^2)|\right\} = \int_0^\infty \log(x)\frac{\sqrt{2}}{\sigma\sqrt{\pi}}e^{-\frac{x^2}{2\sigma^2}}\,dx$$
$$= \log\sigma - \log\sqrt{2} - \frac{\gamma}{2} \tag{15}$$

were γ is the Euler-Mascheroni constant. With this solution, and with $\delta_N = E\left\{\log|N(\mathbf{x};0,1)|\right\}$, eq. (8) becomes:

$$\text{LPF}\left\{\log|I_n(\mathbf{x})|\right\} \approx \log\sigma(\mathbf{x}) - \log\sqrt{2} - \frac{\gamma}{2}. \tag{16}$$

And taking the exponential of each term:

$$e^{\text{LPF}\{\log|I_n(\mathbf{x})|\}} \approx \sigma(\mathbf{x})\frac{e^{-\gamma/2}}{\sqrt{2}}. \tag{17}$$

Thus, we can define an estimator for $\sigma(\mathbf{x})$ as

$$\widehat{\sigma(\mathbf{x})} = \sqrt{2}e^{\text{LPF}\{\log|I(\mathbf{x})-E\{I(\mathbf{x})\}|\}+\gamma/2}. \tag{18}$$

The whole estimation pipeline for the Gaussian case is depicted in Fig. 1. Note that a practical problem may arise when estimating the local mean. This estimation is usually carried out by local sample moments under the assumption of local stationarity. This assumption is not valid in regions with more than one tissues, particularly on the edges, and therefore the estimation can be biased. The proposed methodology overcomes this problem, even when the estimation of $E\{I(x)\}$ is not perfectly achieved. Note that the edges within the image are high frequency areas. The low pass filtering used for the homomorphic separation will remove the effect of edges in the calculation of local moments.

3.2 The Rayleigh case

We know that in those areas of MRI data where the signal is absent, under certain conditions, the noise is Rayleigh distributed ([79, 47]). The Rayleigh distribution in the background of the acquisitions has traditionally been used for noise estimation

in the stationary case. However, once $\sigma(\mathbf{x})$ becomes $\mathbf{x}$-dependent, the estimation over the background might not be related to the estimation over the signal areas. Nevertheless, for the sake of completeness, we add the Rayleigh case here as a previous step for the Rician case. In addition, note that it can be also used in a calibration step or to design coil configuration attending to the generated noise map. Furthermore, note that the spatially variable noise here proposed is similar to some speckle models in literature, and results can be easily extrapolated.

Let us assume a complex Gaussian noise with zero mean and spatially-dependent variance $\sigma^2(\mathbf{x})$:

$$N_0(\mathbf{x}) = N_r(\mathbf{x}; 0, \sigma^2(\mathbf{x})) + j \cdot N_i(\mathbf{x}; 0, \sigma^2(\mathbf{x})). \tag{19}$$

The module of $N_0(\mathbf{x})$ follows a Rayleigh distribution

$$
\begin{aligned}
R(\mathbf{x}; \sigma(\mathbf{x})) &= |N_0(\mathbf{x})| \\
&= \sigma(\mathbf{x}) \cdot \sqrt{N_r^2(\mathbf{x}; 0, 1) + N_i^2(\mathbf{x}; 0, 1)} \\
&= \sigma(\mathbf{x}) \cdot R_1(\mathbf{x}; 1)
\end{aligned}
\tag{20}
$$

As in the Gaussian case, our purpose is to estimate $\sigma(\mathbf{x})$ from the Rayleigh noise $R(\mathbf{x}; \sigma(\mathbf{x}))$. To that aim, we use again a homomorphic filtering:

$$\log|R(\mathbf{x}; \sigma(\mathbf{x}))| = \underbrace{\log \sigma(\mathbf{x})}_{\text{low frequency}} + \underbrace{\log R_1(\mathbf{x}; 1)}_{\text{higher frequency}}.$$

The term $\log \sigma(\mathbf{x})$ is a signal with lower frequency components than $\log R_1(\mathbf{x}; 1)$. We apply the low pass filtering:

$$\text{LPF}\left\{\log R(\mathbf{x}; \sigma(\mathbf{x}))\right\} \approx \log \sigma(\mathbf{x}) + \text{LPF}\left\{\log R_1(\mathbf{x}; 1)\right\}. \tag{21}$$

Let us assume again that the LPF is equivalent to a local averaging:

$$\text{LPF}\left\{\log R(\mathbf{x}; \sigma(\mathbf{x}))\right\} \approx \log \sigma(\mathbf{x}) + E\left\{\log R_1(\mathbf{x}; 1)\right\}.$$

$R_1(\mathbf{x}; 1)$ follows a Rayleigh distribution and the mean of $\log R_1(\mathbf{x}; \sigma)$ can be written as

$$
\begin{aligned}
E\left\{\log R_1(\mathbf{x}; \sigma)\right\} &= \int_0^\infty \log(x) \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \, dx \\
&= \log \sigma + \log \sqrt{2} - \frac{\gamma}{2}
\end{aligned}
\tag{22}
$$

Eq. (21) then becomes:

$$\text{LPF}\left\{\log R(\mathbf{x}; \sigma(\mathbf{x}))\right\} \approx \log \sigma(\mathbf{x}) + \log \sqrt{2} - \frac{\gamma}{2}. \tag{23}$$

Taking the exponential of each term, we can define an estimator for $\sigma(\mathbf{x})$ as
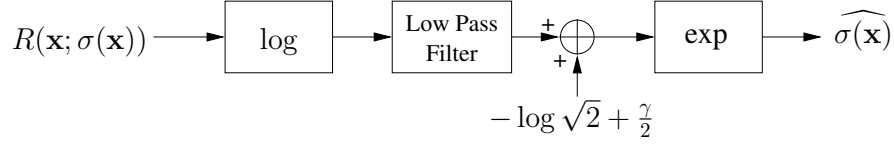
Fig. 2 Pipeline of $\sigma(\mathbf{x})$ estimation assuming Rayleigh noise. $R(\mathbf{x}; \sigma(\mathbf{x}))$ is the original Rayleigh image.

$$\widehat{\sigma(\mathbf{x})} = \frac{1}{\sqrt{2}} e^{\mathrm{LPF}\{\log R(\mathbf{x})\} + \gamma/2}. \tag{24}$$

The whole estimation pipeline for the Rayleigh case is depicted in Fig. 2.

3.3 The Rician case

Finally, we consider the Rician case. Let us assume a signal A($\mathbf{x}$) corrupted with complex Gaussian noise with zero mean and spatially-dependent variance $\sigma^2(\mathbf{x})$, whose module follows a non-stationary Rician distribution:

$$I(\mathbf{x}; A(\mathbf{x}), \sigma(\mathbf{x})) = |A(\mathbf{x}) + N_r(\mathbf{x}; 0, \sigma^2(\mathbf{x})) + j \cdot N_i(\mathbf{x}; 0, \sigma^2(\mathbf{x}))| \tag{25}$$
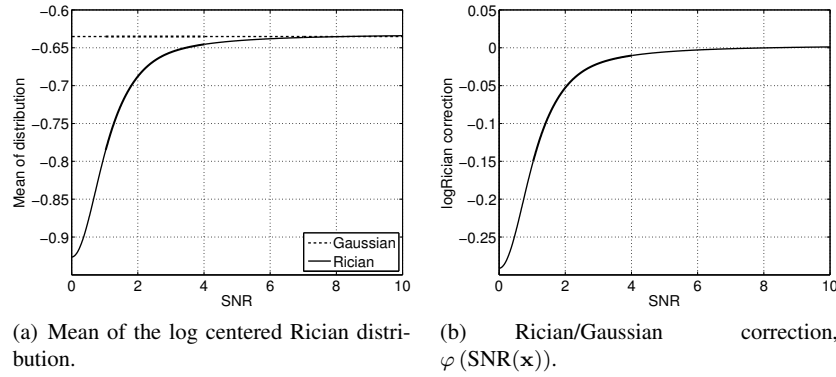
For the sake of simplicity, the dependence of $I(\mathbf{x})$ with $A(\mathbf{x})$ and $\sigma(\mathbf{x})$ will be removed. As in the Gaussian case, our purpose is to estimate $\sigma(\mathbf{x})$ from the Rician signal $I(\mathbf{x})$. However, in this case, the signal and noise are not totally separable. Nevertheless, following the assumption of slow varying $\sigma(\mathbf{x})$, the homomorphic filter can be used to extract the spatially variant pattern of noise.

First, the data is centered by substracting the local mean of the image.

$$
\begin{aligned}
I_n(\mathbf{x}) &= I(\mathbf{x}) - E\{I(\mathbf{x})\} \\
&= I(\mathbf{x}) - \sigma(\mathbf{x})\sqrt{\frac{\pi}{2}} L_{1/2}\left(-\frac{A^2(\mathbf{x})}{2\sigma^2(\mathbf{x})}\right) \\
&= \sigma(\mathbf{x}) \cdot |s_0(\mathbf{x}) + N_r(\mathbf{x}; 0, 1) + j \cdot N_i(\mathbf{x}; 0, 1)| - \sigma(\mathbf{x})\sqrt{\frac{\pi}{2}} L_{1/2}\left(-\frac{s_0^2(\mathbf{x})}{2}\right) \\
&= \sigma(\mathbf{x}) \cdot \mathcal{G}(s_0(\mathbf{x})),
\end{aligned}
$$

where $s_0(\mathbf{x}) = \frac{A(\mathbf{x})}{\sigma(\mathbf{x})}$ is the SNR at each point and $\mathcal{G}(s_0(\mathbf{x}))$ is a function of the SNR:

$$
\begin{aligned}
\mathcal{G}(s_0(\mathbf{x})) &= I(\mathbf{x}; s_0(\mathbf{x}), 1) - E\{I(\mathbf{x}; s_0(\mathbf{x}), 1)\} \\
&= |s_0(\mathbf{x}) + N_r(\mathbf{x}; 0, 1) + j \cdot N_i(\mathbf{x}; 0, 1)| - \sqrt{\frac{\pi}{2}} L_{1/2}\left(-\frac{s_0^2(\mathbf{x})}{2}\right) \tag{26}
\end{aligned}
$$

(a) Mean of the log centered Rician distribution.

(b) Rician/Gaussian correction, $\varphi\left(\text{SNR}(\mathbf{x})\right)$.

Fig. 3 Mean of the logRice distribution as a function of the SNR.

For large values of $s_0(\mathbf{x})$, the Rician distribution tends to a Gaussian and, therefore,

$$\lim_{s_0(\mathbf{x})\to\infty} \mathcal{G}(s_0(\mathbf{x})) = N_r(\mathbf{x}; 0, 1).$$

So, for high SNR, the Rician case converges to the Gaussian case previously studied.

To carry out the homomorphic filtering, we first apply the logarithm:

$$\log |I_n(\mathbf{x})| = \underbrace{\log \sigma(\mathbf{x})}_{\text{low frequency}} + \underbrace{\log |\mathcal{G}(s_0(\mathbf{x}))|}_{\text{higher frequency}}.$$

and afterwards a low pass filtering:

$$\text{LPF}\left\{\log |I_n(\mathbf{x})|\right\} \approx \log \sigma(\mathbf{x}) + \delta_R \tag{27}$$

with δ_R being a low pass residue of $\log |\mathcal{G}(s_0(\mathbf{x}))|$. Again, the LPF behaves as the expected value of the signal:

$$\text{LPF}\left\{\log |\mathcal{G}(s_0(\mathbf{x}))|\right\} \approx E\left\{\log |\mathcal{G}(s_0(\mathbf{x}))|\right\}.$$

To derive an expression of $\sigma(\mathbf{x})$ from eq. (26), assuming a generic Rician random variable $I_R(\mathbf{x}; A, \sigma)$, the expected value of $\log |I_R(\mathbf{x}) - a_1|$ must be studied:

$$E\left\{\log |I_R(\mathbf{x}) - a_1|\right\} = \frac{1}{\sigma^2} \int_0^\infty \log |x - a_1|\, x\, e^{-\frac{x^2+A^2}{2\sigma^2}} I_0\left(\frac{Ax}{\sigma^2}\right) dx \tag{28}$$

The integral has been numerically solved for $A = s_0(\mathbf{x})$, $\sigma = 1$ and $a_1 = \sqrt{\frac{\pi}{2}} L_{1/2}\left(-\frac{s_0^2(\mathbf{x})}{2}\right)$ and depicted in Fig. 3-(a), together with the mean of the log Gaussian in eq. (15). Note that the value of the mean in eq. (28) depends on the SNR of the signal and, for larger values of SNR, the value approximates to the
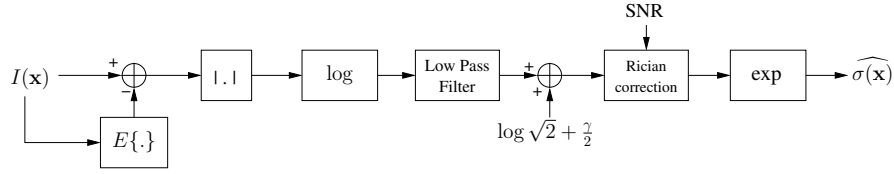
Fig. 4 Pipeline of $\sigma(\mathbf{x})$ estimation assuming Rician noise. $I(\mathbf{x})$ is the original image, $E\{.\}$ is the local expected value of the signal and $|.|$ is the absolute value.

(a) MR slices from Brainweb
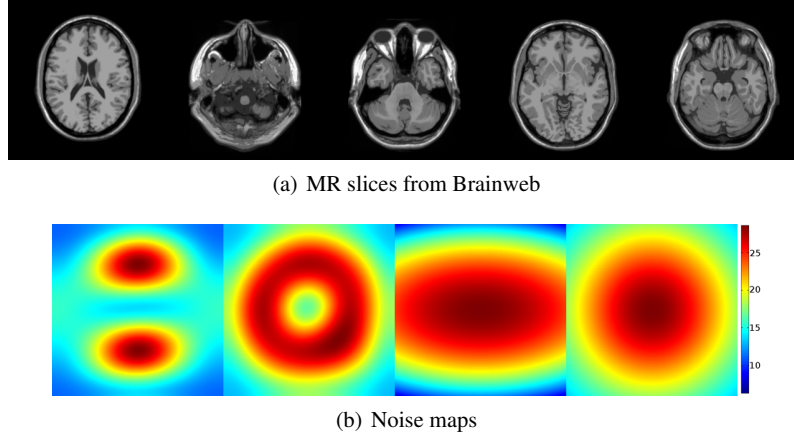
(b) Noise maps

Fig. 5 Synthetic data considered for the experiments. (a) Five different MR images slices from the Brainweb data base. (b) Synthetic $\mathbf{x}-$ dependent noise maps.

Gaussian case. In this case, the expected value can be approximated as the mean value obtained for the Gaussian case plus a correction factor as follows:

$$\text{LPF}\{\log|I_n(\mathbf{x})|\} \approx \log\sigma(\mathbf{x}) - \log\sqrt{2} - \frac{\gamma}{2} + \varphi\left(s_0(\mathbf{x})\right). \qquad (29)$$

with $\varphi\left(s_0(\mathbf{x})\right)$ a Rician/Gaussian correction function, depicted in Fig. 3-(b). Thus, the estimator for $\sigma(\mathbf{x})$ can be defined as

$$\widehat{\sigma(\mathbf{x})} = \sqrt{2}e^{\text{LPF}\{\log|I(\mathbf{x})-E\{I(\mathbf{x})\}|\}}e^{\frac{\gamma}{2}-\varphi(s_0(\mathbf{x}))}. \qquad (30)$$

Note that an estimate of the SNR in each pixel is necessary. This requirement is common to other estimators in literature like [12, 68].

The whole estimation pipeline for the Rician case is depicted in Fig. 4.

4 Materials and methods

4.1 Materials

For the validation of the proposed estimation methodology, the following data sets are considered:

1. *MR synthetic Images*: Five different slices from the Brainweb database ([53]) are used. These slices are shown in Fig. 5-(a). All the images are noise free and the background value is zero.
2. *Noise maps*: The noise is artificially added to the previous MR images. Four different spatially variant patterns are considered, as depicted in Fig. 5-(b). These patterns are coherent with noise patterns that can be found in real acquisitions. The first pattern is derived from a real GRAPPA acquisition ([70]); the second is a linear combination of the first pattern with a rotation of itself; the third pattern is a second order polynomial and the fourth is an isotropic Gaussian function.
3. *Real SENSE acquisitions:* A doped ball phantom is scanned at a 3T Philips Achieva scanner, TFE Pulse Sequence, $224 \times 224 \times 59$, TR/TE=5.264/2.569, slice thickness 3.20mm, SENSE with acceleration 2x (reduction factor $r = 2$). 20 repetitions of the same volume were acquired.

4.2 Methods

For the sake of comparison, the proposed method will be compared to some of the proposals in the literature for the estimation of non-stationary noise in MRI. Only those capable to carry out the estimation over a single image are considered:

1. [71]: originally proposed for Gaussian noise. A good behavior for high SNR is expected, when the Gaussian assumption becomes reasonable. The estimator is defined as:

$$\widehat{\sigma^2(\mathbf{x})} = \langle \left(I^{(1,\mathrm{HH})}(\mathbf{x}) \right)^2 \rangle_{\mathbf{x}} \tag{31}$$

with $I^{(1,\mathrm{HH})}(\mathbf{x})$ the high-high subband coefficients of the stationary wavelet transform (SWT) of the image $I(\mathbf{x})$ at the scale $s = 1$. The operator $\langle I(\mathbf{x}) \rangle_{\mathbf{x}}$ stands for the local sample estimator of the mean:

$$\langle I(\mathbf{x}) \rangle_{\mathbf{x}} = \frac{1}{|\eta(\mathbf{x})|} \sum_{\mathbf{p} \in \eta(\mathbf{x})} I(\mathbf{p}),$$

with $\eta(\mathbf{x})$ a neighborhood centered in $\mathbf{x}$.
2. [72]:

$$\widehat{\sigma^2(\mathbf{x})} = \left(2 - \frac{\pi}{2} \right)^{-1} \left[\langle \widetilde{I}^2(\mathbf{x}) \rangle_{\mathbf{x}} - \left(\langle \widetilde{I}(\mathbf{x}) \rangle_{\mathbf{x}} \right)^2 \right] \tag{32}$$

where $\widetilde{I}(\mathbf{x})$ is the image after removing low frequencies (low-low subband) by using a SWT.

3. [74]:

$$\widehat{\sigma_R(\mathbf{x})} = 1.4826 \cdot \mathrm{MAD}_{\mathbf{x}}\left(I^{(1,HH)}(\mathbf{x})\right) \tag{33}$$

where $I^{(1,HH)}(\mathbf{x})$ is again the high-high subband coefficients of the SWT of $I(\mathbf{x})$ and $\mathrm{MAD}_{\mathbf{x}}(.)$ is the (local) median absolute deviation defined as:

$$\mathrm{MAD}_{\mathbf{x}}\left(I(\mathbf{x})\right) = \underset{\mathbf{p}\in\eta(\mathbf{x})}{\mathrm{median}}\left|I(\mathbf{p}) - \underset{\mathbf{q}\in\eta(\mathbf{x})}{\mathrm{median}}\left(I(\mathbf{q})\right)\right|.$$

For low SNR, a correction is needed:

$$\widehat{\sigma(\mathbf{x})} = \frac{\widehat{\sigma_R(\mathbf{x})}}{\sqrt{\xi(\theta)}} \tag{34}$$

where function $\xi(\theta)$ is defined in [48] and SNR parameter θ is estimated iteratively:

$$\theta_{k+1} = \sqrt{\xi(\theta_k)\left(1 + \frac{\langle I(\mathbf{x})\rangle_{\mathbf{x}}^2}{\widehat{\sigma_R^2(\mathbf{x})}}\right) - 2}.$$

4. [68]: the Gausssian estimator is defined as

$$\widehat{\sigma_R(\mathbf{x})} = 1.4826 \cdot \mathrm{MAD}_{\mathbf{x}}\left(I(\mathbf{x})\right). \tag{35}$$

For low SNR, the correction in eq. (34) is applied.

5. [100]:

$$\widehat{\sigma_R^2(\mathbf{x})} = \min_{\mathbf{p}\in\eta(\mathbf{x}):\, \mathbf{p}\neq\mathbf{x}} \|R(\mathbf{x}) - R(\mathbf{p})\|_2^2 \tag{36}$$

where $R(\mathbf{x}) = I(\mathbf{x}) - \psi(I(\mathbf{x}))$ and $\psi(I(\mathbf{x}))$ is the low-pass filtered data. For low SNR, the correction in eq. (34) is also applied.

6. [69] for one single image:

$$\widehat{\sigma(\mathbf{x})} = Q_n\left(\{\mathbf{p}\in\eta(\mathbf{x}):\, \epsilon(\mathbf{p})\}\right) \tag{37}$$

where ϵ is the difference (residual) between noisy data and a biophysical model projection onto data, see [69] for further details. For the sake of simplicity, and due to the lack of an available biophysical model, we considered the sample mean:

$$\epsilon(\mathbf{x}) = |I(\mathbf{x}) - \langle I(\mathbf{x})\rangle_{\mathbf{x}}|.$$

The Q_n scale estimator is defined by [161] as

$$Q_n(\{x_1,\ldots,x_n\}) = 2.2219 \cdot \{|x_i - x_j|; i < j\}_{(k)}$$

where symbol $\{\cdot\}_{(k)}$ denotes k-th element in the ascending ordered data (order statistics) and here $k = \binom{[n/2]+1}{2}/4$. To mitigate the impact of the outliers, Landman proposes the re-estimation of the noise map by removing observations with lower SNR value than an adaptive computed threshold:

$$t(\mathbf{x}) = \min \left\{ 5, \underset{\mathbf{p} \in \eta(\mathbf{x})}{\text{median}} \left\{ I(\mathbf{p}) / \widehat{\sigma(\mathbf{x})} \right\} - 3 \right\}.$$

7. The EM estimator in [98]: The estimator is initially defined for working with multiple samples. To adapt it for a single image, we replace the estimation *along samples* by a local estimation:

$$\widehat{A_{k+1}}(\mathbf{x}) = \left\langle \frac{I_1 \left(\frac{\widehat{A_k(\mathbf{x})} I(\mathbf{x})}{\widehat{\sigma_k^2(\mathbf{x})}} \right)}{I_0 \left(\frac{\widehat{A_k(\mathbf{x})} I(\mathbf{x})}{\widehat{\sigma_k^2(\mathbf{x})}} \right)} I(\mathbf{x}) \right\rangle_{\mathbf{x}} \tag{38}$$

$$\widehat{\sigma_{k+1}^2}(\mathbf{x}) = \max \left\{ \frac{1}{2} \left\langle I^2(\mathbf{x}) \right\rangle_{\mathbf{x}} - \frac{(\widehat{A_k}(\mathbf{x})=^2}{2}, 0 \right\} \tag{39}$$

where $I_n(.)$ is the modified Bessel function of the first kind and n-th order. The initialization process of the EM algorithm is obtained by the method of the moments

$$\widehat{A_0}(\mathbf{x}) = \left(2 \langle I^2(\mathbf{x}) \rangle_{\mathbf{x}}^2 - \langle I^4(\mathbf{x}) \rangle_{\mathbf{x}} \right)^{\frac{1}{4}}$$

$$\widehat{\sigma_0^2}(\mathbf{x}) = \frac{1}{2} \left(\langle I^2(\mathbf{x}) \rangle_{\mathbf{x}} - (\widehat{A_0}(\mathbf{x}))^2 \right).$$

For all the methods a 5×5 window has been considered for the local operators, except for the EM, where the optimal behavior was empirically observed for 3×3 neighborhoods. In all the cases where the SWT is needed, the Daubechies (db7) wavelet was used.

The proposed homomorphic algorithm was implemented using the following methods:

1. For the initial estimation of the SNR, the output of the local EM estimator for signal and noise in eq. (38) and eq. (39) has been considered, using a 3×3 estimation window:

$$\text{SNR}(\mathbf{x}) = \frac{\widehat{A_k(\mathbf{x})}}{\widehat{\sigma_k(\mathbf{x})}}$$

with $k = 10$ iterations.
2. The low-pass filter was designed as a Gaussian filter in the Fourier domain with $H(0,0) = 1$ and standard deviation σ_f. In all the experiments considered, σ_f is always in the range $3 - 5$. To avoid the border effects due to the implicit

periodicity of the DFT, the equivalent filter in the Discrete Cosine Transform domain was used.

5 Experiments and results

For the sake of validation of the proposed methodology, some illustrative experiments are carried out.

5.1 Noise separability

Before testing the estimation methods, we check the initial assumption that the non-stationary noise can be separated into two different components using a homomorphic filtering. To that end, we consider two simple but illustrative examples for the Gaussian and Rayleigh scenarios. In this experiment we want to show that the proposed methodology can be affectively applied to separate the stationary component of noise from the spatially variant component.

First, we assume a 256×256 stationary Gaussian noise image, $N(\mathbf{x}; 0, 1)$, with zero mean and unitary variance. The image is multiplied by the first noise map from Fig. 5, obtaining a non-stationary noise image $N(\mathbf{x}; 0, \sigma(\mathbf{x})) = \sigma(\mathbf{x}) \cdot N(\mathbf{x}; 0, 1)$, where $\sigma(\mathbf{x})$ is the considered noise map. The image is then processed using the pipeline in Fig. 1 but using two different filters: a low pass filter to obtain the noise map estimation and its high pass counterpart to obtain the stationary noise component. The high pass filter is defined in the frequency domain as

$$H_{\text{HPF}}(\mathbf{f}) = 1 - H_{\text{LPF}}(\mathbf{f}).$$

This way, the output of the low pass filter provides the estimator $\widehat{\sigma(\mathbf{x})}$, while the high pass filter should provide the stationary component of noise $N_{\text{HP}}(\mathbf{x})$. Due to the absolute value in the pipeline, $|N(\mathbf{x}; 0, 1)|$ follows a half normal distribution and, therefore, $N_{\text{HP}}(\mathbf{x})$ should also follow that distribution. In order to check this assumption, we consider the transformation $N_{\text{HP}}^2(\mathbf{x})$, which follows a χ square distribution with 1 degree of freedom. This transformation leads to easily check the hypothesis of stationarity by means of a Chi-squared goodness-of-fit test[2], where the null hypothesis is: the data follows a χ square distribution with 1 degree of freedom.

---

[2] The Chi-squared goodness-of-fit test evaluates the discrepancies between the observed frequency distribution and a particular theoretical distribution. The p-value obtained from this test is interpreted based on the significance level (we consider 0.05 in this study) in such a way that if $p \geq 0.05$ the null hypothesis is accepted. i.e. there are no statistically significant differences between the frequency distribution and the theoretical distribution.

(a) Gaussian case   (b) Rayleigh case
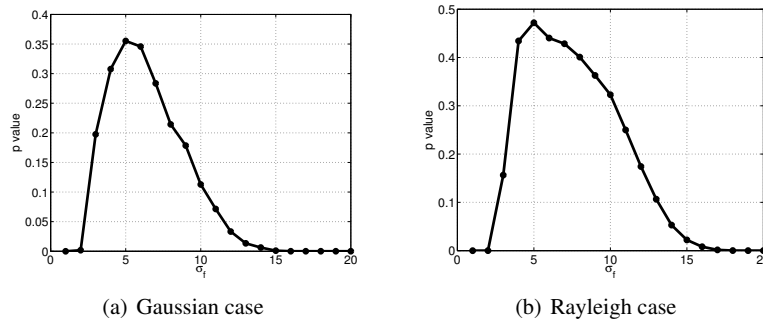
Fig. 6 P-values of the Chi-square goodness-of-fit test for the output distributions as a function of the filter bandwidth. The average of 100 experiments is considered.

All the aforementioned assumptions were checked throughout the pipeline shown in Fig. 1 for a set of 100 repetitions. First, we tested the Normality assumption of $N(\mathbf{x}; 0, 1)$ and $N(\mathbf{x}; 0, \sigma(\mathbf{x}))$ where the null hypothesis is the stationary Gaussian distribution. As was expected the goodness-of-fit test for $N(\mathbf{x}; 0, 1)$ obtained an average of 0.4851 for the p-value, and the stationarity of the Gaussian is accepted for all cases. In the case of $N(\mathbf{x}; 0, \sigma(\mathbf{x}))$, the stationarity of data is discarded in all repetitions, where a p-value close to 0 was obtained for every case. This result shows that the noise map applied from Fig. 5 causes a statistically significant non-stationarity in the data. The goodness-of-fit test performed over $N_{\mathrm{HP}}^2(\mathbf{x})$ was applied for different values of the filter bandwidth σ_f to check its influence in the stationarity. The averages of p-values for each σ_f are depicted in Fig. 6-(a), where the null hypothesis is accepted, i.e. the stationarity of the high pass component of noise is accepted. Note that these results confirm that the output noise $N_{\mathrm{HP}}(\mathbf{x})$ is stationary for a bandwidth between 3 and 11. These results show the suitability of the proposed methodology for separating the stationary and non-stationary components of noise in the Gaussian scenario.

The same experiment is repeated for the Rayleigh scenario. We assume now a 256×256 Rayleigh image, $R(\mathbf{x}; 1)$, with unitary parameter. The image is also multiplied by the first noise map from Fig. 5, obtaining the non-stationary noise image $R(\mathbf{x}; \sigma(\mathbf{x})) = \sigma(\mathbf{x}) \cdot R(\mathbf{x}; 1)$. The image is then processed using the pipeline in Fig. 2 with a high pass filter in order to obtain the high pass noise component. The output of the low pass filter is again considered as the estimator $\widehat{\sigma(\mathbf{x})}$ and the high pass filter output is supposed to be stationary noise $R_{\mathrm{HP}}(\mathbf{x})$. If the separation is successfully performed, this noise must follow a Rayleigh distribution with unitary parameter.

The goodness-of-fit test was performed for 100 repetitions of the experiment. We first tested the stationarity of $R(\mathbf{x}; 1)$ and $R(\mathbf{x}; \sigma(\mathbf{x}))$ as in the previous scenario. Results obtained for $R(\mathbf{x}; \sigma(\mathbf{x}))$ showed an average p-value of 0.5183, and the null hypothesis (stationarity of the Rayleigh distribution) was confirmed for all the cases. The test for the non-stationary noise $R(\mathbf{x}; \sigma(\mathbf{x}))$ discarded the null hypothesis for
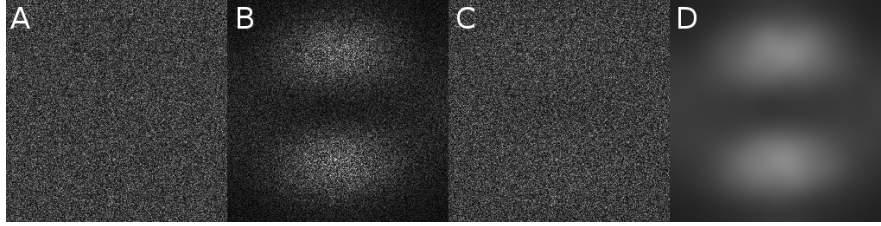
Fig. 7 Example of separability of noise components: a) Stationary Rayleigh noise, $R(\mathbf{x}; 1)$; b) Non-stationary Rayleigh noise, $R(\mathbf{x}; \sigma(\mathbf{x}))$; c) High pass component of the homomorphic filtering, $R_{\text{HP}}(\mathbf{x})$; d) Low pass component of the homomorphic filtering, $\widehat{\sigma(\mathbf{x})}$. All images has been normalized for display purposes.

all the cases, obtaining a negligible p-value for all of them. This result also confirms that the non-stationary noise map applied is statistically significant. The separated high pass noise $R_{\text{HP}}(\mathbf{x})$ obtained after the homomorphic filtering was also tested for different values of the filter bandwidth σ_f. The averages of p-values with respect σ_f is depicted in Fig. 6-(b) where the stationarity hypothesis is accepted for a similar range as in the Gaussian scenario. In the light of these results we can conclude that the proposed methodology also separates the stationary and non-stationary components of noise in the Rayleigh scenario.

For the sake of illustration, the separation process of the non-stationary Rayleigh noise is depicted in Fig. 7. Note that, due to the presence of a spatial-dependent $\sigma(\mathbf{x})$, noise image $R(\mathbf{x}; \sigma(\mathbf{x}))$ in Fig. 7-(b) shows a spatially variant pattern. The homomorphic filtering process is able to recover the two original components of noise: in Fig. 7-(c) a stationary noisy pattern $R_{\text{HP}}(\mathbf{x})$, and in Fig. 7-(d) the noise maps corresponding to $\widehat{\sigma(\mathbf{x})}$.

5.2 Synthetic experiments

First, synthetic images and synthetic noise maps shown in Fig. 5 were used. The noise is generated as a Rician noise with $\mathbf{x}-$dependent variance $\sigma^2(\mathbf{x})$ following the generic model in eq. (7) and eq. (8). Two different cases are considered: (a) one single image (first slice in Fig. 5) and one single noise map (first one in Fig. 5); and (b) all combinations of the five images and the four noise maps (20 different combinations). All the noise and signals were normalized to have the same maximum SNR. In order to generate different SNR values, in each case, the noise-free signal is divided by a constant and the noise is generated using the original map. A set of 100 repetitions of each configuration is considered. Noise is estimated using the aforementioned state-of-the-art methods and the relative error for each pixel is calculated as:

$$\text{error}(\mathbf{x}) = \frac{|\sigma(\mathbf{x}) - \widehat{\sigma(\mathbf{x})}|}{\sigma(\mathbf{x})} \tag{40}$$
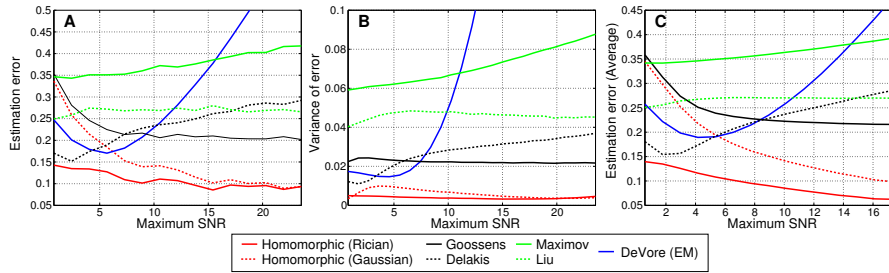
Fig. 8 Error of estimation for parameter $\sigma(\mathbf{x})$ using synthetic non-stationary Rician noise. The average of 100 experiments is considered for each configuration. a) Estimation error for a single image; b) Variance of error for a single image; c) Average of estimation error for 20 different images.

The average of all the points, repetitions and configurations is considered to achieve a single value for each SNR. For the single image case, the variance of the error along the samples is also considered. Results are depicted in Fig. 8. For the sample moments estimation a window of 5×5 size is used. The homomorphic methods use a low pass filter set to $\sigma_f = 3.4$. This value has been selected following results in Fig. 6.

Although a single value is not representative for the estimation of the spatial pattern, this experiment gives a quantitative insight of the behavior of some of the methods in literature and the proposed one. Note that, for the whole considered SNR range, the proposed method clearly outperforms those in literature. For high SNR values, Gaussian and Rician models converge, as expected. Goossens' method shows also a behavior that improves with the SNR, since this method assumes a Gaussian distribution of the data in the wavelet domain. However, the Gaussian homomorphic should have the same limitation, and it shows a much better performance for all the values.

It is also interesting the results of the variance of the error. The proposed methods are those with a minimum variance (among those tested), a fact that improves their suitability as estimators. In addition, note that the Rician homomorphic scheme shows almost a constant variance for all the SNR values. The EM and the Maximov estimators, on the other hand, show an increase of the variance for high SNR values. This effect is due to the dependence of the estimators on the signal: for high SNR the level of noise is low and the estimators mismatch signal details with noise. Goossens' and Liu's show both consistent variance values for the whole SNR range, although with higher values than the proposed methods.

In a whole, results do not differ when one single configuration or the 20 possible configurations are considered. Thus, from the numerical results in this experiment, the proposed homomorphic method is the one which always shows the lowest estimation error and the lowest variance. In addition, note that inside the tissues, the SNR is usually high enough to use the Gaussian approximation, which is simpler and does not require a SNR calculation. The error committed in the areas of interest
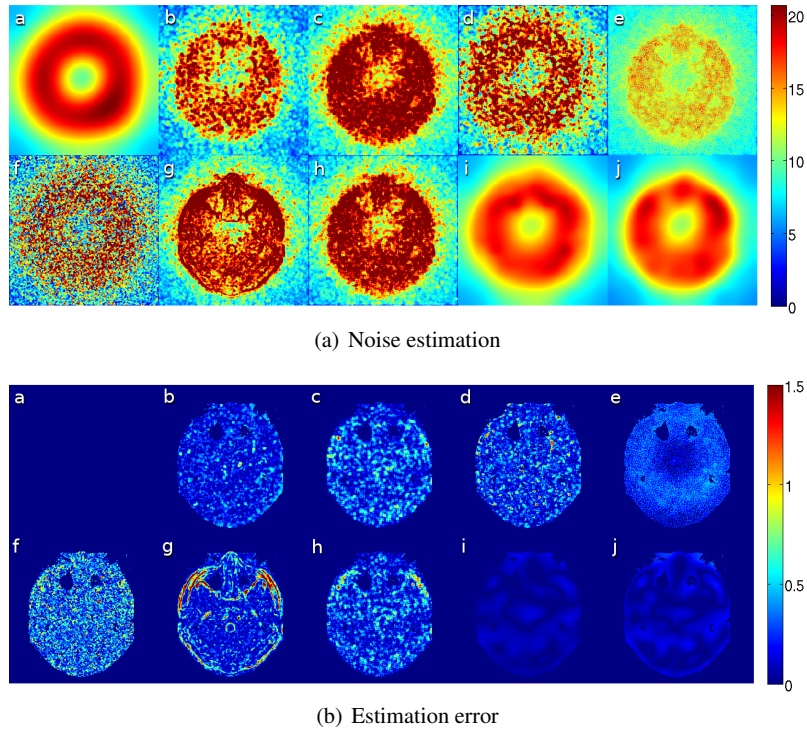
(a) Noise estimation
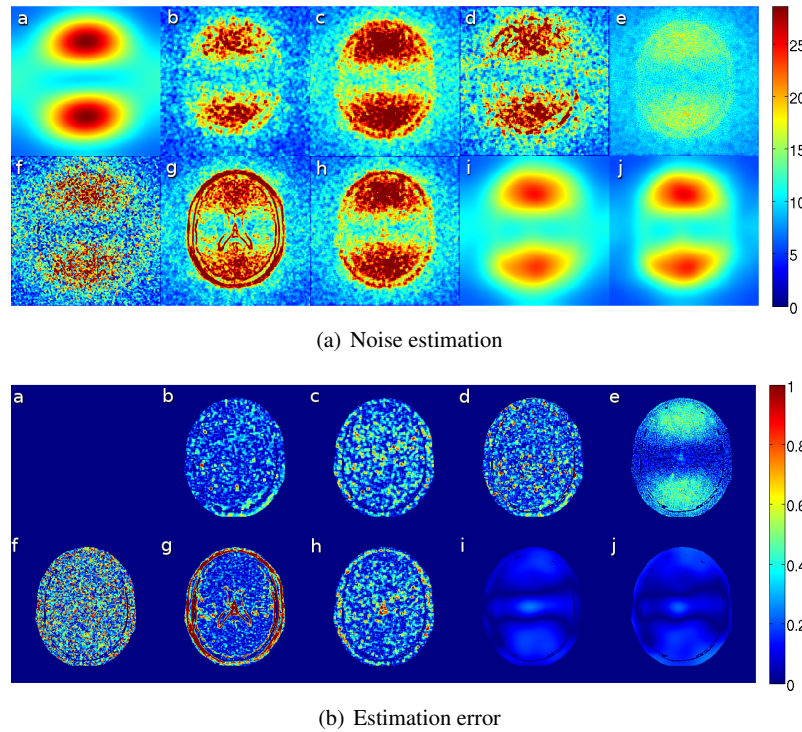
(b) Estimation error

Fig. 9 Noise estimation (top) and estimation error (bottom), using first image and first noise map. (a) Original noise map; (b) Goossens; (c) Delakis; (d) Liu; (e) Manjón; (f) Maximov; (g) EM; (h) Landman; (i) Rician Homomorphic; (j) Gaussian Homomorphic.

is very similar to the Rician and, therefore, the Gaussian estimator can be successfully used in that case.

In the **second** experiment we search for visual comparison of the different estimation methods. The first image and the first noise map from Fig. 5 were used to generate a single MRI slice corrupted by non-stationary Rician noise. The noise map $\sigma(\mathbf{x})$ is estimated from the magnitude image. Visual results for the considered methods are depicted in Fig. 9, together with the estimation error, calculated by using eq. (40). To ease the visualization, the image has been saturated to the maximum value of the original map. The maximum value of the error was also limited to 1, although some of the methods showed much greater errors.

This experiment clearly shows the suitable behavior of the homomorphic approach when compared to local estimators. The use of a LPF to extract the noise map leads to avoid the granular pattern and to eliminate the influence of the edges. Note that those methods based on local estimation precisely show such a granular pattern due to the use of local neighborhoods with small size. This implies that the number of points used for estimation is low and, conversely, the variance of the estimation is high. Another problem with local estimators is related to the edges of

(a) Noise estimation

(b) Estimation error

Fig. 10 Noise estimation (top) and estimation error (bottom), using second image and second noise map. (a) Original noise map; (b) Goossens; (c) Delakis; (d) Liu; (e) Manjón; (f) Maximov; (g) EM; (h) Landman; (i) Rician Homomorphic; (j) Gaussian Homomorphic.

the image. In those areas where two different tissues or regions lay inside the same window, the estimation becomes highly biased and causes the miss-estimation of local moments. The proposed methodology is able to overcome these local problems by properly estimating a smooth and reliable noisy pattern. Even when the Rician hypothesis is relaxed, as in the Gaussian assumption, the method still outperforms the other methods.

In Fig. 10, we repeat the experiment for the second image and the second noise maps of Fig. 5. Results are totally coherent with the previous experiment. Once more, the proposed method shows the best estimation with both assumptions (Rice and Gaussian). Note that this time, all the results show some influence of the the high signal areas over the estimated noise map, though it is more subtle with the homomorphic approach, Fig.10-(i) and 10-(j).
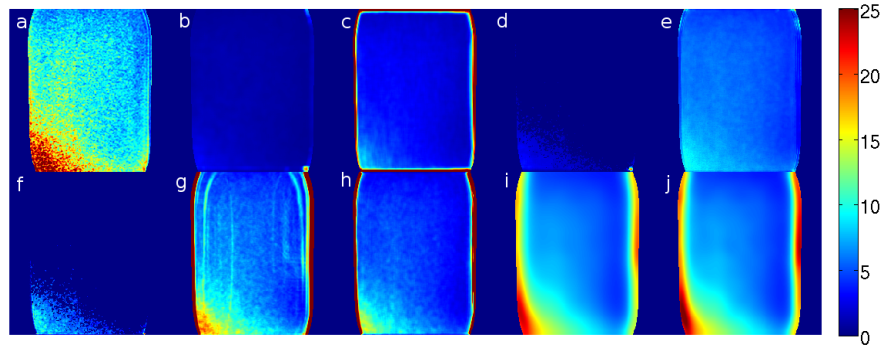
Fig. 11 Noise estimation using a real SENSE acquisition. (a) EM along 20 samples (Silver standard); (b) Goossens; (c) Delakis; (d) Liu; (e) Manjón; (f) Maximov; (g) Local EM; (h) Landman; (i) Rician Homomorphic; (j) Gaussian Homomorphic.

5.3 Real MR acquisition

Finally, an experiment with real acquisitions is carried out. Since a golden standard is not available, we consider the original (multi-sample) EM estimation of $\sigma(\mathbf{x})$ using the 20 samples as a reference (*silver standard*). We compare it to the average of the estimation of the map of noise in each of the acquisitions by using the different estimators. Results are gathered in Fig. 11. Note that only the local EM scheme, the Landman's method and the homomorphic approaches are able to follow the variation pattern in the map of noise. Note also that some of the methods based on local estimation show a border effect in the boundaries of the slice. This effect is caused by the larger dimension of the object compared to the FOV, which is not the case of brain acquisitions, but it is in other areas such as imaging of body or knee. Thus, it must be taken into consideration.

The proposed methods show a good estimation of the map of noise, coherent with that considered as silver standard. The local EM approach in Fig. 11-(g) shows here a very good behavior, which differs from results of previous experiments, where the worse results were related to those areas with borders and transitions between different tissues. In this case, the phantom has not inner edges and, thus, the EM estimator is able to provide a proper estimate of noise in all the signal area. This different behavior evidences the influence of transitions between different tissues in local estimation. A problem that the proposed methodology successfully overcomes.

6 Conclusions

A new methodology to estimate spatially variant noise in MR has been presented. It is based on the homomorphic approach that allows to separate the contribution of

the pattern of noise and the noise itself. The estimator has proved to be robust, easy to implement and it does not require additional information about the acquisition, nor multiple repetitions to carry out the estimation. It also reduces the influence of borders in the final estimated map, and avoids the traditional granular pattern shown by methods relying on local estimation. Additionally, results have shown that the proposed methodology presents the lowest error variance and its independence to the SNR. This result evidences the suitability of the proposed estimators.

Three different estimators for $\sigma(\mathbf{x})$ were presented: for the Gaussian, Rayleigh and Rician cases. While the latest would be the most suitable for MRI, the first one presents a good alternative, with the advantage of not requiring a prior estimation of the SNR. The Gaussian estimator could be used as an approximation of the Rician case that will fail in the low SNR areas: mostly the background, where a proper estimation is not so relevant. Additionally, it also presents a solid alternative for a built-in automatic estimator inside the scanning software, which just needs to access to the final complex image in the processing pipeline just before the absolute value is taken. Using this complex image, the variable noise pattern can be accurately estimated without the need of any additional information or SNR estimation.

The proposed methodology shows also a great potential when is jointly used with other estimation methods. In this paper, we have carried out the estimation over the $\mathbf{x}-$space using a simple low-pass filtering. However, the homomorphic separation of noise can also be applied to other methods in the literature, particularly those based on wavelets. With the proper adjustments, these other methods can improve in robustness and presumably more accurate results can be achieved.

Finally, in this paper we have only focused on the Rician noise, leaving aside the nc-χ model, so popular lately in the literature. It was done under the assumption that nowadays most acquisitions are based on single coil systems or in pMRI-SENSE reconstruction. Even those scanning softwares using GRAPPA or similar algorithms (like Siemens) give also the option of a matched filter reconstruction that lately will produce Rician data. Nevertheless, the Gaussian approach for high SNR is also valid for non-stationary nc-χ data, and a correction like the one here proposed for Rician can be easily derived.

Acknowledgements

Blind estimation of spatially variant noise in GRAPPA MRI

Santiago Aja-Fernández, Gonzalo Vegas-Sánchez-Ferrero[*]

Abstract The reconstruction process in multiple coil MRI scanners makes the noise features in the final magnitude image become non-stationary, i.e. the variance of noise becomes position-dependent. Therefore, most noise estimators proposed in the literature cannot be used in multiple-coil acquisitions. This effect is augmented when parallel imaging methods, such as GRAPPA, are used to increase the acquisition rate. In this work we propose a new technique that allows the estimation of the spatially variant maps of noise from the GRAPPA reconstructed signal when only one single image is available and no additional information is provided. Other estimators in the literature need extra information that is not always available, which has supposed an important limitation in the usage of noise models for GRAPPA. The proposed approach uses a homomorphic separation of the spatially variant noise in two terms: a stationary noise term and one low frequency signal that correspond to the **x**-dependent variance of noise. The non-stationary variance of noise is estimated by a low pass filtering. The noise term is obtained via prior wavelet decomposition. Results in real and synthetic experiments evidence the suitability of the simplification used and the good performance of the proposed methodology.

1 Introduction

Noise is one of the main sources of quality deterioration in Magnetic Resonance (MR) data. It is produced by the stochastic motion of free electrons in the RF coil, which is a conductor, and by eddy current losses in the patient, which are inductively coupled to the RF coil. The presence of noisy patterns on the acquired MR signal

---

is a problem that affects not only the visual quality of the images, but also may interfere with further processing techniques.

The stationary Rician distribution has been widely accepted as a suitable model for noise in MR magnitude images. The main assumption for this model is that a single value of σ characterizes the whole data set. However, this premise will mostly fail when considering modern scanners with multiple-coil antennae. Parallel MRI techniques extended the applicability of these systems by increasing the image acquisition rate via subsampling of the **k**–space data. In order to correct the artifacts created by the subsampling, reconstruction methods such as GeneRalized Autocalibrated Partially Parallel Acquisitions (GRAPPA) [33] are employed. When GRAPPA is used for reconstruction, noise in the reconstructed image can be approximated by a non-stationary nc-χ distribution [42]. Nowadays, GRAPPA has become a *de facto* standard in the acquisition software of commercial scanners. However, many processing techniques still assume the stationary Rician distribution, forgetting about the non-stationarity of the data. In this context, most noise estimators proposed in literature are based on the assumption of a single σ value for all the pixels in the image [47, 84]. There are also some proposals in literature to estimate noise out of GRAPPA data, such as [70]. However, they all require extra information about the reconstruction process.

In this paper we propose a new technique to estimate the spatially variant maps of noise $\sigma(\mathbf{x})$ from the magnitude signal under GRAPPA reconstructions. The proposal only needs one single magnitude image and no additional information is required. To that end, the analysis of noise in GRAPPA is carefully studied in order to properly relax the constraints to provide a blind estimation in realistic scenarios (no information about coils sensitivities, only one image is provided). For this purpose, two initial assumptions observed in real cases are considered: (**1**) The signal-to-noise ratio (SNR) in each of the coils is high. This assumption allows to relax the noise characterization to a non-stationary Gaussian model. We will show that this relaxation is suitable for real MR acquisitions; (**2**) The variability of the map of noise is smaller than the variability of the noise itself, i.e., $\sigma(\mathbf{x})$ is a low frequency signal when compared to the noise, which is also a rational assumption in MRI acquisitions. Both sources of variability can therefore be separated by using a homomorphic transformation [159]. This separation of noise components was previously described by authors in [75] for SENSE acquisitions.

2 Simplified noise model in GRAPPA

The GRAPPA [33] method reconstructs the full **k**-space in each coil from a sub–sampled **k**-space acquisition. The reconstructed lines are estimated through a linear combination of the existing samples. According to [51, 42], this reconstruction can be written in the **x**–space as a multiplication:

$$S_l^{\mathcal{R}}(\mathbf{x}) = |\Omega| \sum_{m=1}^{L} S_m^{\mathcal{S}}(\mathbf{x}) \times W_m(l, \mathbf{x}) \tag{1}$$

where $S_l^{\mathcal{S}}(\mathbf{x})$ is the subsampled data in coil l-th, with $l = 1, \cdots, L$, and $S_l^{\mathcal{R}}(\mathbf{x})$ is the reconstructed data in each coil. $W_m(l, \mathbf{x})$ are the GRAPPA reconstruction coefficients in the $\mathbf{x}$-space and $|\Omega|$ the cardinal of the image in each coil. The reconstructed image in each coil can be seen as a non-stationary Gaussian image:

$$S_l^{\mathcal{R}}(\mathbf{x}) = A_l^{\mathcal{R}}(\mathbf{x}) + N_l^{\mathcal{R}}(\mathbf{x} : 0, \sigma_l^R(\mathbf{x})) \tag{2}$$

with $A_l^{\mathcal{R}}(\mathbf{x})$ being the reconstructed signal and $N_l^{\mathcal{R}}(\mathbf{x})$ a complex Gaussian noise with zero mean and spatially variant variance. The composite magnitude signal (CMS) is obtained using the sum-of-squares (SoS) of the reconstructed signal in each coil:

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^{L} |S_l^{\mathcal{R}}(\mathbf{x})|^2}. \tag{3}$$

In [42] authors pointed out that the resultant distribution of the CMS is not strictly a nc-χ, but its behavior will be very similar and can be efficiently modeled as such with a small approximation error. However, the reconstruction method will highly increase the correlations between the reconstructed signals in each coil, which leads to a decrease of the number of degrees of freedom of the distribution. As a consequence, the final distribution will show a (reduced) *effective number of coils* L_{eff} and an (increased) *effective variance of noise* σ_{eff}^2:

$$L_{\text{eff}}(\mathbf{x}) = \frac{|\mathbf{A}_{\mathcal{R}}|^2 \, \text{tr}\,(\mathbf{C}_X) + (\text{tr}\,(\mathbf{C}_X))^2}{\mathbf{A}_{\mathcal{R}}^* \mathbf{C}_X \mathbf{A}_{\mathcal{R}} + ||\mathbf{C}_X||_F^2}; \tag{4}$$

$$\sigma_{\text{eff}}^2(\mathbf{x}) = \frac{\text{tr}\,(\mathbf{C}_X)}{L_{\text{eff}}}. \tag{5}$$

where $\mathbf{C}_X(\mathbf{x}) = \mathbf{W}\mathbf{\Sigma}\mathbf{W}^*$ is the covariance matrix of the interpolated data at each spatial location, $\mathbf{A}_{\mathcal{R}}(\mathbf{x}) = [A_1^{\mathcal{R}}, \cdots, A_L^{\mathcal{R}}]^T$ is the noise-free reconstructed signal, $||.||_F$ is the Frobenius norm, $\mathbf{\Sigma}$ is the covariance matrix of the original data and $\mathbf{W}(\mathbf{x})$ is the $L \times L$ GRAPPA interpolation matrix for each $\mathbf{x}$. Although the nc-χ model is feasible for GRAPPA, the resulting distribution is non-stationary since the effective parameters are spatially dependent.

The main inconvenience of this noise model is that not only $\sigma_{\text{eff}}^2(\mathbf{x})$ has to be estimated for each point of the image, but also $L_{\text{eff}}(\mathbf{x})$. In addition, the product $\sigma_{\text{eff}}^2(\mathbf{x}) \cdot L_{\text{eff}}(\mathbf{x})$ is also spatially dependent, which complicates any blind parameters estimation.

If we assume a high SNR, the non-stationary nc-χ model can be simplified: if $A_l^{\mathcal{R}}(\mathbf{x}) >> \sigma_l^R(\mathbf{x}))$ in eq. (2), then the SoS in eq. (3) can be approximated using a Taylor series as:

$$M_L(\mathbf{x}) \approx A_T(\mathbf{x}) + N^{\mathcal{R}}(\mathbf{x}; 0, \sigma_T^2(\mathbf{x})), \tag{6}$$
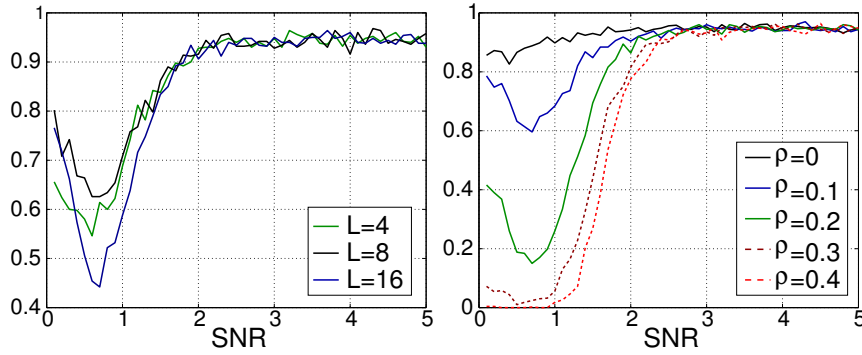
Fig. 1 Ratio of experiments in which the Gaussian distribution hypothesis is accepted, using a Pearson goodness-of-fit test. Left: $\rho = 0.1$. Right: $L = 8$.

where $A_T^2(\mathbf{x}) = \mathbf{A}_\mathcal{R}^* \mathbf{A}_\mathcal{R}$ and $N^\mathcal{R}(\mathbf{x}; 0, \sigma_T^2(\mathbf{x}))$ is a non-stationary Gaussian noise with zero mean and variance:

$$\sigma_T^2(\mathbf{x}) = |\Omega|^2 \frac{\mathbf{A}_\mathcal{R}^* \mathbf{C}_X \mathbf{A}_\mathcal{R}}{\mathbf{A}_\mathcal{R}^* \mathbf{A}_\mathcal{R}} = |\Omega|^2 \frac{\mathbf{A}_\mathcal{S}^* \mathbf{W}^* \mathbf{W} \mathbf{\Sigma} \mathbf{W}^* \mathbf{W} \mathbf{A}_\mathcal{S}}{\mathbf{A}_\mathcal{S}^* \mathbf{W}^* \mathbf{W} \mathbf{A}_\mathcal{S}}. \tag{7}$$

Matrix $\mathbf{A}_\mathcal{S} = [A_1^\mathcal{S}, \cdots, A_L^\mathcal{S}]^T$ is the original sampled signal (without noise), and $\mathbf{A}_\mathcal{R} = \mathbf{W} \cdot \mathbf{A}_\mathcal{S}$. The variance of noise depends on the position, the GRAPPA reconstruction coefficients, the original covariance matrix and the noise-free signals.

Finally, note that the variance of noise for the simplified model in eq. (7) is consistent to the formulation of effective values proposed in [42] and eq. (5).

The suitability of the Gaussian simplification for nc-χ data was validated by means of a Pearson goodness-to-fit test. The p-value obtained from this test is interpreted based on a significance level of 0.05: if $p \geq 0.05$ the null hypothesis (the data follows a Gaussian distribution) is accepted, i.e. there are no statistically significant differences between the frequency distribution and the theoretical distribution. The validation was performed by generating a series of discrete correlated complex Gaussian random variables (500 samples per variable). Different values for SNR in each coil, L (number of coils) and ρ (correlation coefficient) were used. L variables were combined using the SoS, see eq. (3). 1000 experiments of each configuration were considered. The ratio of experiments in which the hypothesis is accepted is depicted in Fig. 1. For the majority of the configurations, for an SNR> 2 the Gaussian assumption can be accepted in around the 95% of the experiments, which means that, in most of the practical situations in MRI, this assumption will be valid.

3 Blind Noise estimation for grappa

In what follows the simplified non-stationary Gaussian model of eq. (6) will be assumed for GRAPPA, where an image $A_T(\mathbf{x})$ is corrupted with additive Gaussian

noise with zero mean and spatially-dependent variance $\sigma_T^2(\mathbf{x})$. Eq. (6) can be written as

$$M_L(\mathbf{x}) = A_T(\mathbf{x}) + \sigma_T(\mathbf{x}) \cdot N(\mathbf{x}; 0, 1).$$

Our purpose here is to estimate $\sigma_T(\mathbf{x})$ from the final image $M_L(\mathbf{x})$ with no additional information assumed (blind estimation). To that aim, we propose a homomorphic separation of the noise components to extract the spatially variant pattern of noise.

Let us assume that the variance of noise $\sigma_T^2(\mathbf{x})$ slowly varies across the image, i.e. it is a low frequency signal. To remove the contribution of the signal $A_T(\mathbf{x})$, we will use the stationary wavelet transform (SWT). We define $I^{(1,\mathrm{HH})}(\mathbf{x})$ as the high-high subband coefficients of the SWT of the image $M_L(\mathbf{x})$ at the scale $s = 1$. Typically, this subband is considered to be an *only noise* component when the data follows a Gaussian distribution [71]:

$$I^{(1,\mathrm{HH})}(\mathbf{x}) \approx \sigma_T(\mathbf{x}) \cdot N(\mathbf{x}; 0, 1).$$

Next, we separate signals $\sigma_T(\mathbf{x})$ and $N(\mathbf{x})$ by applying the logarithm:

$$\log |I^{(1,\mathrm{HH})}(\mathbf{x})| = \underbrace{\log \sigma_T(\mathbf{x})}_{\text{low frequency}} + \underbrace{\log |N(\mathbf{x})|}_{\text{higher frequency}}.$$

The noise term $\log |N(\mathbf{x})|$ has its energy distributed along all frequencies, while the term $\log \sigma_T(\mathbf{x})$ is a low frequency signal by hypothesis. The later can be recovered using a low pass filtering (LPF):

$$\mathrm{LPF} \left\{ \log |I^{(1,\mathrm{HH})}(\mathbf{x})| \right\} \approx \log \sigma_T(\mathbf{x}) + \delta_N \tag{8}$$

with δ_N being a low pass residue of $\log |N(\mathbf{x})|$. This residue must be calculated to remove it from the estimation. If we assume that the LPF has a small passband, the LPF $\{\log |N(\mathbf{x})|\}$ is a good estimator of the local mean. By hypothesis, $N(\mathbf{x})$ is stationary, and therefore the local mean is the same for all pixels. Thus, we can consider the LPF as a good approximation of the mean of the signal:

$$\mathrm{LPF} \left\{ \log |N(\mathbf{x})| \right\} \approx E \left\{ \log |N(\mathbf{x})| \right\}.$$

$|N(\mathbf{x})|$ follows a *half-normal* distribution with expectation

$$E \left\{ \log |N(\mathbf{x}; 0, 1)| \right\} = -\log \sqrt{2} - \frac{\gamma}{2}$$

where γ is the Euler-Mascheroni constant. With this solution, eq. (8) becomes:

$$\mathrm{LPF} \left\{ \log |I^{(1,\mathrm{HH})}(\mathbf{x})| \right\} \approx \log \sigma_T(\mathbf{x}) - \log \sqrt{2} - \frac{\gamma}{2}. \tag{9}$$

Thus, we can define an estimator for $\sigma_T(\mathbf{x})$ as
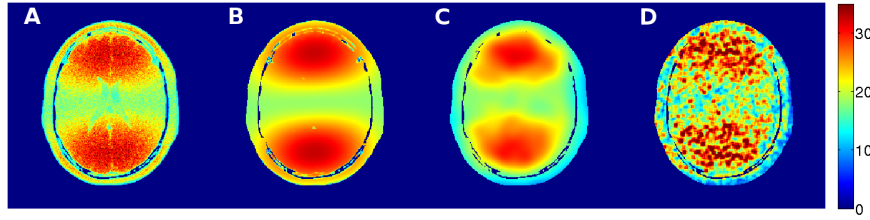
Fig. 2 Estimation of $\sigma_T(\mathbf{x})$. a) Std. of 100 samples; b) Theoretical value; c) Estimation over 1 sample (homomorphic); d) Estimation over 1 sample (Goossens).

$$\widehat{\sigma_T(\mathbf{x})} = \sqrt{2}e^{\text{LPF}\{\log|I^{(1,\text{HH})}(\mathbf{x})|\}+\gamma/2}. \tag{10}$$

4 Results

For the sake of validation, first, a synthetic experiment is carried out using a 2D synthetic slice from the BrainWeb dataset, with intensity values in $[0-255]$. 8-coil systems where simulated using a realistic sensitivity map. The image in each coil is corrupted with Gaussian noise with variance $\sigma^2 = 100$ and correlation between coils $\rho = 0.1$. The $\mathbf{k}$–space is uniformly subsampled by a factor of 2, keeping 32 ACS lines. The CMS is reconstructed using GRAPPA and SoS. 100 repetitions of the experiment are considered to obtain statistical measures.

Before carrying out the estimation, a Pearson goodness-of-fit test is carried out for each point for the 100 repetitions where the null hypothesis is again that the data follows a Gaussian distribution. In the 91.85% of the area of interest (the background is not considered for the experiments) the null hypothesis is accepted. So, since the Gaussian assumption can be applied for this configuration, we can use the proposed methodology to estimate the noise.

For the sake of comparison, the estimation of the GRAPPA noise maps is done using the following methods: 1) The sample variance for each location when 100 realizations of the same experiments are considered; 2) the theoretical value of the simplified model in eq. (7); 3) the proposed homomorfic estimation over one sample; 4) the method proposed by Goossens in [71] over one sample. This method also uses the SWT to estimate the spatially variant pattern of noise. In all the cases where the SWT is needed, the Daubechies (db7) wavelet is used. Results are shown in Fig.2. The homomorphic scheme matches the noise maps predicted by the theoretical value in Fig. 2-b. The experiment clearly shows the suitable behavior of the homomorphic approach when compared to methods that use local estimators, like Goossens'. The use of a LPF to extract the noise map leads to avoid the granular pattern and to eliminate the influence of the edges.

For the last experiment, real acquisitions are considered. 100 repetitions of the same slice of a phantom, scanned in an 8-channel head coil on a GE Signa 1.5T EXCITE 12m4 scanner with FGRE Pulse Sequence to generate low SNR, see Fig. 3.
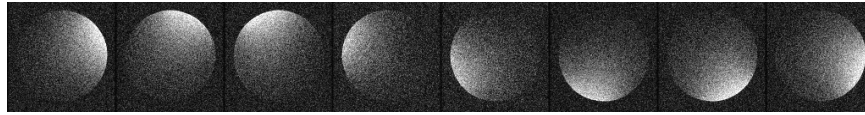
Fig. 3 Slice of an 8-coil acquisition of a doped ball phantom.
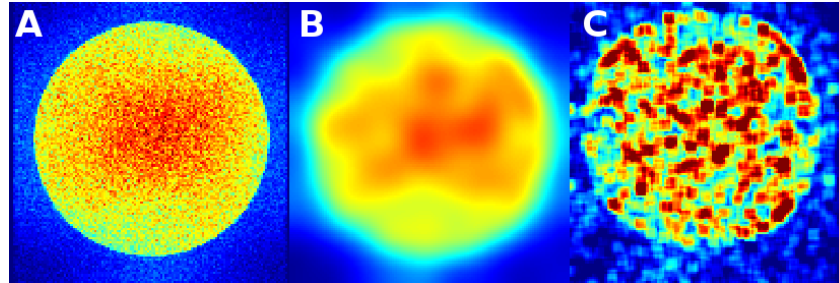
Fig. 4 Noise estimation over the phantom. a) Standard deviation of 100 samples; b) Homomorphic estimation of one sample; c) Estimation using Goossens' method.

Matrix size 128×128, TR/TE 8.6/3.38 ms, FOV 21×21cm, slice thickness 1mm. All the 100 samples are 2× subsampled. The GRAPPA reconstruction coefficients are derived from one sample, using 16 ACS lines, and used for interpolation in all samples. The CMS is obtained by SoS.

First, the Pearson goodness-of-fit test is repeated for each point for the 100 repetitions. For the 94.02% of points inside the signal area the null hypothesis is accepted (91.1% if the whole image is considered). Again, the Gaussian assumption holds, even when the original images may look *vey noisy*. Results of the estimation are in Fig. 4. Once more, the homomorphic approach succeeds in following the spatial pattern of the variance of noise.

Additionally, in order to compare the capability to successfully estimate the non-stationary variance of noise with just one single image, we calculated the ratio of pixels laying in the 95% confidence interval provided by the estimation of the standard deviation with different number of samples. In Fig. 5 this ratio is depicted for both methods. Note that the homomorphic approach outperforms the Goosens' method in all the cases, and more than 90% of pixels lay within the confidence interval obtained for 20 different acquisitions. This result evidences that the proposed method provides with just one acquisition similar results to the ones obtained with a higher number of acquisitions.

5 Conclusions

A new methodology to estimate spatially variant noise in GRAPPA-reconstructed MRI without any additional knowledge has been presented. It is based on the homo-
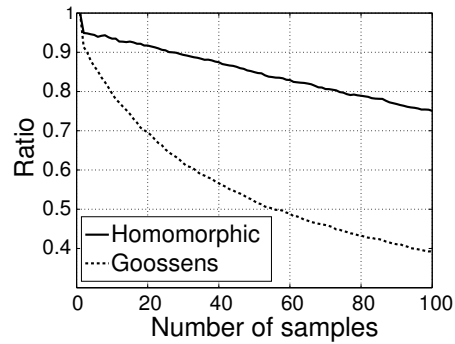
Fig. 5 Ratio of pixels whose variance estimates calculated with one image lay within the 95% confidence interval of the non-stationary variance of noise for an increasing number of acquisitions.

morphic approach that allows separating the contribution of the pattern of noise and the noise itself after a wavelet decomposition. The estimator has proved to be robust, easy to implement and it does not require additional information about the acquisition, nor multiple repetitions to carry out the estimation. Visual results show that it succeeds in extracting the non-stationary component of noise, while it avoids the traditional granular pattern shown by methods relying on local estimation. For the whole methodology, a simplified model of noise for GRAPPA has been adopted. Although this Gaussian model might seem inaccurate, some initial experiments point out that it can be a suitable model for typical GRAPPA acquisitions. On this paper, only synthetic data and a real phantom have been used to test the Gaussianity assumption. In the future, we expect to extend the evaluation of this model over multiple MRI acquisitions with different protocols and over different areas of the body. We are confident that the trend holds. Finally, the estimator here proposed has been focused on GRAPPA but it can be easily extrapolated to non-accelerated multiple coil acquisitions with SoS reconstruction.

Acknowledgements

References

1. G. Kruger, G. H. Glover, Physiological noise in oxygenation-sensitive magnetic resonance imaging, Magn. Reson. Med. 46 (2001) 1631–637.
2. N. Petridou, A. Schäfer, P. Gowland, R. Bowtell, Phase vs. magnitude information in functional magnetic resonance imaging time series: toward understanding the noise, Magn. Reson. Imag. 27 (2009) 1046–1057.
3. A. Counter, A. Olofsson, E. Borg, B. Bjelke, A. Häggström, H. Grahn, Analysis of magnetic resonance imaging acoustic noise generated by a 4.7 t experimental system, Acta Oto-Laryngologica 120 (6) (2000) 739–743.
4. G. McGibney, M. Smith, Unbiased signal-to-noise ratio measure for magnetic resonance images, Med. Phys. 20 (4) (1993) 1077–1078.
5. H. Gudbjartsson, S. Patz, The Rician distribution of noisy MRI data, Magn. Reson. Med. 34 (6) (1995) 910–914.
6. S. Aja-Fernández, M. Niethammer, M. Kubicki, M. E. Shenton, C.-F. Westin, Restoration of DWI data using a Rician LMMSE estimator, IEEE Trans. Med. Imag. 27 (10) (2008) 1389–1403.
7. J. Sijbers, A. J. den Dekker, D. Van Dyck, E. Raman, Estimation of signal and noise from Rician distributed data, in: Proc. of the Int. Conf. on Signal Proc. and Comm., Las Palmas de Gran Canaria, Spain, 1998, pp. 140–142.
8. S. Aja-Fernández, C. Alberola-López, C.-F. Westin, Noise and signal estimation in magnitude MRI and Rician distributed images: A LMMSE approach, IEEE Trans. Image Process. 17 (8) (2008) 1383–1398.
9. A. Tristán-Vega, S. Aja-Fernández, DWI filtering using joint information for DTI and HARDI, Med. Imag. Anal. 14 (2) (2010) 205–218.
10. V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. LeBihan, F. Poupon, Parallel MRI noise correction: an extension of the LMMSE to non central Chi distributions, in: MICCAI 2011, Vol. 6891 of Lecture Notes in Computer Science, 2011, pp. 217–224.
11. N. Wiest-Daesslé, S. Prima, P. Coupé, S. Morrissey, C. Barillot, Rician noise removal by Non–Local Means filtering for low signal–to–noise ratio MRI: Applications to DT–MRI, in: Medical Image Computing and Computer-Assisted Intervention, Vol. 5242 of Lecture Notes in Computer Science, Springer–Verlag, 2008, pp. 171–179.
12. J. Manjón, J. Carbonell-Caballero, J. Lull, G. García-Martí, L. Martı-Bonmatı, M. Robles, MRI denoising using Non-Local Means, Med. Imag. Anal. 12 (2008) 514–523.
13. A. Tristán-Vega, V. García-Pérez, S. Aja-Fernández, C.-F. Westin, Efficient and robust Non-local Means denoising of MR data based on salient features matching, Comp. Meth. Prog. in Biom. 105 (2012) 131–144.

14. R. Salvador, A. Peña, D. K. Menon, T. Carpenter, J. Pickard, E. T. Bullmore, Formal charac-
    terization and extension of the linearized diffusion tensor model, Human brain mapping 24
    (2005) 144–155.
15. A. Tristán-Vega, S. Aja-Fernández, C.-F. Westin, Least squares for diffusion tensor estima-
    tion revisited: Propagation of uncertainty with Rician and non-Rician signals, NeuroImage
    59 (2012) 4032–4043.
16. J. Andersson, Maximum a posteriori estimation of diffusion tensor parameters using a Rician
    noise model: why, how and but, NeuroImage 42 (4) (2008) 1340–1356.
17. B. Landman, P.-L. Bazin, J. Prince, J. Hopkins, Diffusion tensor estimation by maximizing
    Rician likelihood, in: Procs. of IEEE 11th International Conference on Computer Vision,
    2007, pp. 1–8.
18. C. Poupon, A. Roche, J. Dubois, J.-F. Mangin, F. Poupon, Real-time MR diffusion tensor
    and Q-Ball imaging using Kalman filtering, Med. Imag. Anal. 12 (5) (2008) 527–534.
19. P. Casaseca-de-la-Higuera, A. Tristán-Vega, S. Aja-Fernández, C. Alberola-López, C. F.
    Westin, R. San-José-Estépar, Optimal real-time estimation in Diffusion Tensor Imaging,
    Magn. Reson. Imag. 30 (4) (2012) 506–517.
20. X. Wu, S. Bricq, C. Collet, Brain MRI segmentation and lesion detection using generalized
    Gaussian and Rician modeling, in: B. M. Dawant, D. R. Haynor (Eds.), Medical Imaging
    2011: Image Processing, Vol. 7962, SPIE, 2011.
21. S. Roy, A. Carass, P.-L. Bazin, S. Resnick, J. L. Prince, Consistent segmentation using a
    Rician classifier, Med. Imag. Anal. 16 (2) (2012) 524–535.
22. O. Michailovich, Y. Rathi, S. Dolui, Spatially regularized compressed sensing for high angu-
    lar resolution diffusion imaging, IEEE Trans. Med. Imag. 30 (5) (2011) 1100–1115.
23. S. Dolui, A. Kuurstra, O. V. Michailovich, Rician compressed sensing for fast and stable
    signal reconstruction in diffusion MRI, in: D. R. Haynor, S. Ourselin (Eds.), Medical Imaging
    2012: Image Processing, Vol. 8314, SPIE, 2012, pp. –.
24. J. Noh, V. Solo, Rician distributed functional MRI: Asymptotic power analysis of likelihood
    ratio tests for activation detection, in: Acoustics Speech and Signal Processing (ICASSP),
    2010 IEEE International Conference on, 2010, pp. 477–480.
25. J. Noh, V. Solo, Rician distributed fMRI: Asymptotic power analysis and Cramer–Rao lower
    bounds, IEEE Trans. Signal Proccess. 59 (3) (2011) 1322–1328.
26. M. Welvaert, J. Durnez, B. Moerkerke, G. Verdoolaege, Y. Rosseel, neuRosim: An R package
    for generating fMRI data, Journal of Statistical Software 44 (10).
27. C. Constantinides, E. Atalar, E. McVeigh, Signal-to-noise measurements in magnitude im-
    ages from NMR based arrays, Magn. Reson. Med. 38 (1997) 852–857.
28. M. Stehling, R. Turner, P. Mansfield, Echo-planar imaging: magnetic resonance imaging in
    a fraction of a second, Science 254 (5028) (1991) 43–50.
29. R. L. DeLaPaz, Echo-planar imaging, Radiographics 14 (1994) 1045–1058.
30. W. S. Hoge, D. H. Brooks, B. Madore, W. E. Kyriakos, A tour of accelerated parallel MR
    imaging from a linear systems perspective, Concepts Magn Reson Part A 27A (1) (2005)
    17–37.
31. D. J. Larkman, R. G. Nunes, Parallel magnetic resonance imaging, Phys. Med. Biol. 52
    (2007) 15–55, invited Topical Review.
32. K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, SENSE: Sensitivity encoding
    for fast MRI, Magn. Reson. Med. 42 (5) (1999) 952–62.
33. M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer,
    A. Haase, Generalized autocalibrating partially parallel acquisitions (GRAPPA), Magn. Re-
    son. Med. 47 (6) (2002) 1202–1210.
34. M. Blaimer, F. Breuer, M. Mueller, R. Heidemann, M. Griswold, P. Jakob, SMASH, SENSE,
    PILS, GRAPPA: how to choose the optimal method, Top Magn Reson Imag 15 (4) (2004)
    223–36.
35. J. Yu, H. Agarwal, M. Stuber, M. Schär, Practical signal-to-noise ratio quantification for
    sensitivity encoding: application to coronary MR angiography, J. Magn. Reson. 33 (6) (2011)
    1330–1340.

36. R. Henkelman, Measurement of signal intensities in the presence of noise in MR images, Med. Phys. 12 (2) (1985) 232–233.

37. M. Brummer, R. Mersereau, R. Eisner, R. Lewine, Automatic detection of brain contours in MRI data sets, IEEE Trans. Med. Imag. 12 (2) (1993) 153–166.

38. P. Roemer, W. Edelstein, C. Hayes, S. Souza, O. Mueller, The NMR phased array, Magn. Reson. Med. 16 (1990) 192–225.

39. C. Hayes, P. B. Roemer, Noise correlations in data simultaneously acquired form multiple surface coil arrays, Magn. Reson. Med. 16 (1990) 181–191.

40. M. Harpen, Noise correlations exist for independent RF coils, Magn. Reson. Med. 23 (1992) 394–397.

41. R. Brown, Y. wang, P. Spincemaille, R. F. Lee, On the noise correlation matrix for multiple radio frequency coils, Magn. Reson. Med. 58 (2004) 218–224.

42. S. Aja-Fernández, A. Tristán-Vega, W. S. Hoge, Statistical noise analysis in GRAPPA using a parametrized non-central chi approximation model, Magn. Reson. Med. 65 (4) (2011) 1195–1206.

43. P. Thünberg, P. Zetterberg, Noise distribution in SENSE- and GRAPPA- reconstructed images: a computer simulation study, Magn. Reson. Imag. 25 (2007) 1089–94.

44. D. Drumheller, General expressions for Rician density and distribution functions, IEEE Trans. on Aerospace and Electronic Systems 29 (2) (1993) 580–588.

45. M. K. Simon, Probability distributions involving Gaussian random variables, Kluwer Academic Publishers, 2002.

46. O. Dietrich, *et al*, Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics, Magn. Reson. Imag. 26 (2008) 754–762.

47. S. Aja-Fernández, A. Tristán-Vega, C. Alberola-López, Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models, Magn. Reson. Imag. 27 (2009) 1397–1409.

48. C. G. Koay, P. J. Basser, Analytically exact correction scheme for signal extraction from noisy magnitude MR signals, J. Magn. Reson. 179 (2006) 317–322.

49. S. Aja-Fernández, A. Tristán-Vega, Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction, Magn. Reson. Med. 67 (2) (2012) 580–585.

50. S. Aja-Fernández, A. Tristán-Vega, V. Brion, Effective noise estimation and filtering from correlated multiple-coil mr data, Magn. Reson. Imag. 31 (2) (2013) 272–285.

51. F. A. Breuer, S. A. Kannengiesser, M. Blaimer, N. Seiberlich, P. M. Jakob, M. A. Griswold, General formulation for quantitative G-factor calculation in GRAPPA reconstructions, Magn. Reson. Med. 62 (3) (2009) 739–746.

52. P. Robson, A. Grant, A. Madhuranthakam, R. Lattanzi, D. Sodickson, C. McKenzie, Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions, Magn. Reson. Med. 60 (2008) 895.

53. D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, A. Evans, Design and construction of a realistic digital brain phantom, IEEE Trans. Med. Imag. 17 (3) (1998) 463–468.

54. D. Jones, P. Basser, Squashing peanuts and smashing pumpkins: how noise distorts diffusion weighted MR data, Magn. Reson. Med. 52 (2004) 979–993.

55. R. Clarke, P. Scifo, G. Rizzo, F. Dell'Acqua, G. Scotti, F. Fazio, Noise correction on Rician distributed data for fibre orientation estimators, IEEE Trans. Med. Imag. 27 (9) (2008) 1242–1251.

56. O. Friman, A. Hennemuth, H.-O. Peitgen, A Rician-Gaussian mixture model for segmenting delayed enhancement MRI images, in: Proceedings of the 16th Scientific Meeting of the ISMRM, Vol. 3, 2008, p. 1040.

57. A. Hennemuth, A. Seeger, O. Friman, S. Miller, B. Klumpp, S. Oeltze, H.-O. Peitgen, A comprehensive approach to the analysis of contrast enhanced cardiac MR images, IEEE Trans. Med. Imag. 27 (11) (2008) 1592–1610.

58. V. J. Schmid, B. Whitcher, A. R. Padhani, N. J. Taylor, G.-Z. Yang, A Bayesian hierarchical model for the analysis of a longitudinal dynamic contrast-enhanced MRI oncology study, Magn. Reson. Med. 61 (1) (2009) 163–174.

59. M. Martín-Fernandez, C. Aberola-López, J. Ruiz-Alzola, C.-F. Westin, Sequential anisotropic Wiener filtering applied to 3D MRI data, Magn. Reson. Imag. 25 (2007) 278–292.

60. C. G. Koay, E. Özarslan, P. J. Basser, A signal transformational framework for breaking the noise floor and its applications in MRI, J. Magn. Reson. 197 (2) (2009) 108–119.

61. M. Martín-Fernández, E. Muñoz-Moreno, L. Cammoun, J.-P. Thiran, C.-F. Westin, C. Alberola-López, Sequential anisotropic multichannel Wiener filtering with Rician bias correction applied to 3D regularization of DWI data, Med. Imag. Anal. 13 (1) (2009) 19–35.

62. K. Krissian, S. Aja-Fernández, Noise-driven anisotropic diffusion filtering of MRI, IEEE Trans. Image Process. 18 (10) (2009) 2265 –2274.

63. S. Aja-Fernandez, G. Vegas-Sanchez-Ferrero, A. Tristan-Vega, Noise estimation in MR GRAPPA reconstructed data, in: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, 2011, pp. 1815 –1818.

64. H.-E. Assemlal, D. Tschumperlé, L. Brun, K. Siddiqui, Recent advances in diffusion MRI modeling: Angular and radial reconstruction, Med. Imag. Anal. 15 (4) (2011) 369–396.

65. J. Rajan, D. Poot, J. Juntu, J. Sijbers, Noise measurement from magnitude MRI using local estimates of variance and skewness, Physics in Medicine and Biology 55 (16) (2010) N441.

66. R. Maitra, D. Faden, Noise estimation in magnitude MR datasets, IEEE Trans. Med. Imag. 28 (10) (2009) 1615–1622.

67. J. Veraart, J. Rajan, R. R. Peeters, A. Leemans, S. Sunaert, J. Sijbers, Comprehensive framework for accurate diffusion MRI parameter estimation, Magn. Reson. Med. 70 (4) (2013) 972–984.

68. I. I. Maximov, E. Farrher, F. Grinberg, N. Jon Shah, Spatially variable Rician noise in magnetic resonance imaging, Med. Imag. Anal. 16 (2) (2012) 536–548.

69. B. A. Landman, P.-L. Bazin, S. A. Smith, J. L. Prince, Robust estimation of spatially variable noise fields, Magn. Reson. Med. 62 (2) (2009) 500–509.

70. S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, A. Tristán-Vega, Noise estimation in parallel MRI: GRAPPA and SENSE, Magn. Reson. Imag. 32 (3) (2014) 281–290.

71. B. Goossens, A. Pizurica, W. Philips, Wavelet Domain Image Denoising for Non-Stationary Noise and Signal-Dependent Noise, in: IEEE International Conference on Image Processing, 2006, pp. 1425–1428.

72. I. Delakis, O. Hammad, R. I. Kitney, Wavelet-based de-noising algorithm for images acquired with parallel magnetic resonance imaging (MRI), Phys. Med. Biol. 52 (13) (2007) 3741.

73. J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, M. Robles, New methods for MRI denoising based on sparseness and self-similarity, Med. Imag. Anal. 16 (1) (2012) 18–27.

74. R. W. Liu, L. Shi, W. Huang, J. Xu, S. C. H. Yu, D. Wang, Generalized total variation-based MRI Rician denoising model with spatially adaptive regularization parameters, Magn. Reson. Imag. 32 (6) (2014) 702–720.

75. S. Aja-Fernández, T. Pieciak, G. Vegas-Sanchez-Ferrero, Spatially variant noise estimation in mri: A homomorphic approach, Med. Imag. Anal. 20 (2015) 184 – 197.

76. S. Aja-Fernández, G. Vegas-Sanchez-Ferrero, Blind estimation of spatially variant noise in GRAPPA MRI, in: 2015 IEEE International Symposium on Biomedical Imaging, ISBI, Brooklyn, NY, 2015.

77. K. Tabelow, H. U. Voss, J. Polzehl, Local estimation of the noise level in MRI using structural adaptation, Med. Imag. Anal. 20 (1) (2015) 76 – 86.

78. P. Coupé, J. V. Manjón, E. Gedamu, D. Arnold, M. Robles, D. L. Collins, Robust Rician noise estimation for MR images, Med. Imag. Anal. 14 (4) (2010) 483–493.

79. A. Macovski, Noise in MRI, Magn. Reson. Med. 36 (1996) 494–497.

80. J. Sijbers, A. den Dekker, P. Scheunders, D. Van Dyck, Maximum-likelihood estimation of Rician distribution parameters, IEEE Trans. Med. Imag. 17 (3) (1998) 357–361.

81. J. Sijbers, A. den Dekker, J. Van Audekerke, M. Verhoye, D. Van Dyck, Estimation of the noise in magnitude MR images, Magn. Reson. Imag. 16 (1) (1998) 87–90.

82. J. Sijbers, A. J. den Dekker, Maximum likelihood estimation of signal amplitude and noise variance form MR data, Magn. Reson. Imag. 51 (2004) 586–594.

83. J. Sijbers, A. J. den Dekker, D. Poot, M. Verhoye, N. Van Camp, A. Van der Linden, Robust estimation of the noise variance from background MR data, in: Proc. of SPIE. Medical Imaging 2006: Image Processing, Vol. 6144, 2006, pp. 2018–2028.

84. J. Sijbers, *et al*, Automatic estimation of the noise variance from the histogram of a magnetic resonance image, in: Physics in Medicine and Biology, Vol. 52, 2007, pp. 1335–1348.

85. P. Getreuer, M. Tong, L. Vese, A variational model for the restoration of MR images corrupted by blur and rician noise, in: G. Bebis, R. Boyle, B. Parvin, D. Koracin, S. Wang, K. Kyungnam, B. Benes, K. Moreland, C. Borst, S. DiVerdi, C. Yi-Jen, J. Ming (Eds.), Advances in Visual Computing, Vol. 6938 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2011, pp. 686–698.

86. R. Nowak, Wavelet-based Rician noise removal for Magnetic Resonance Imaging, IEEE Trans. Image Process. 8 (10) (1999) 1408–1419.

87. S. Basu, T. Fletcher, R. Whitaker, Rician noise removal in diffusion tensor MRI, in: Proceedings of MICCAI, Vol. 1, 2006, pp. 117–125.

88. P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, C.Barillot, An optimized blockwise non local means denoising filter for 3D magnetic resonance images, IEEE Trans. Med. Imag. 27 (2008) 425–441.

89. T. Marzetta, EM algorithm for estimating the parameters of multivariate complex Rician density for polarimetric SAR, in: Proc. of ICASSP, Vol. 5, 1995, pp. 3651–3654.

90. L. He, I. Greenshields, A nonlocal Maximum Likelihood Estimation method for Rician noise reduction in MR images, IEEE Trans. Med. Imag. 28 (2) (2009) 165–172.

91. H. Liu, C. Yang, N. Pan, E. Song, R. Green, Denoising 3D MR images by the enhanced non-local means filter for Rician noise, Magn. Reson. Imag. 28 (10) (2010) 1485 – 1496.

92. R. Riji, J. Rajan, J. Sijbers, M. Nair, Iterative bilateral filter for rician noise reduction in mr images, Signal, Image and Video Processing (2014) 1–6.

93. A. C. S. Chung, J. A. Noble, Statistical 3D vessel segmentation using a Rician distribution, in: Proceedings of the 2nd MICCAI, Vol. 1679 of Lecture Notes in Computer Science, Springer, 1999, pp. 82–89.

94. C. G. Koay, E. Özarslan, C. Pierpaoli, Probabilistic identification and estimation of noise (PIESNO): A self-consistent approach and its applications in MRI, J. Magn. Reson. 199 (1) (2009) 94–103.

95. A. Martn, E. Schiavi, Noise modelling in parallel magnetic resonance imaging: A variational approach, in: A. Campilho, M. Kamel (Eds.), Image Analysis and Recognition, Lecture Notes in Computer Science, Springer International Publishing, 2014, pp. 121–128.

96. G. Lohmann, S. Bohn, K. Müller, R. Trampel, R. Turner, Image restoration and spatial resolution in 7-Tesla magnetic resonance imaging, Magn. Reson. Med. 64 (1) (2010) 15–22.

97. J. Rajan, J. Veraart, J. V. Audekerke, M. Verhoye, J. Sijbers, Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images, Magn. Reson. Imag. 20.

98. M. D. DeVore, A. D. Lanterman, J. A. O'Sullivan, ATR performance of a Rician model for SAR images, in: Proc. of SPIE 2000, Orlando, 2000, pp. 34–37.

99. W. Guo, F. Huang, Adaptive Total Variation Based Filtering for MRI Images With Spatially Inhomogeneous Noise and Artifacts, in: Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on, 2009, pp. 101–104.

100. J. V. Manjón, P. Coupé, L. Marti-Bonmati, D. L. Collins, M. Robles, Adaptive Non-Local Means Denoising of MR Images With Spatially Varying Noise Levels, J. Magn. Reson. Imag. 31 (1) (2010) 192–203.

101. Y. Ding, Y.-C. Chung, O. P. Simonetti, A method to assess spatially variant noise in dynamic MR image series, Magn. Reson. Med. 63 (3) (2010) 782–789.

102. A. A. Samsonov, C. R. Johnson, Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels, Magn. Reson. Med. 52 (4) (2004) 798–806.

103. Y. Chang, D. Liang, L. Ying, Nonlinear GRAPPA: A kernel approach to parallel MRI reconstruction, Magn. Reson. Med.

104. M. Lustig, D. Donoho, J. Santos, J. Pauly, Compressed sensing MRI, IEEE Signal Processing Magazine 25 (2) (2008) 72–82.

105. J. Aelterman, H. Q. Luong, B. Goossens, A. Pizurica, W. Philips, Compass: a joint framework for parallel imaging and compressive sensing in MRI, in: Proc. International Conference on Image Processing (ICIP'10), IEEE, 2010, pp. 1653–1656.

106. M. Lustig, J. Pauly, SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary $k-$space, Magn. Reson. Med. 64 (2) (2010) 457–471.

107. S. Aja-Fernández, G. Vegas-Sanchez-Ferrero, A. Tristán-Vega, About the background distribution in MR data: a local variance study, Magn. Reson. Imag. 28 (2010) 739–752.

108. S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, M. Martín-Fernández, C. Alberola-López, Automatic noise estimation in images using local statistics. Additive and multiplicative cases, Imag. Vision Comp. 27 (6) (2009) 756–770.

109. A. Tristán-Vega, C.-F. Westin, S. Aja-Fernández, Bias of least squares approaches for diffusion tensor estimation from array coils in DT–MRI, in: MICCAI 2009, Vol. 5761 of Lecture Notes in Computer Science, 2009, pp. 919–926.

110. P. Fillard, X. Pennec, V. Arsigny, N. Ayache, Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics, IEEE Trans. Med. Imag. 26 (11) (2007) 1472–1482.

111. Z.-P. Liang, P. C. Lauterbur, Principles of Magnetic Resonance Imaging. A signal processing perspective, IEEE Press, New York, 2000.

112. J. X. Ji, J. B. Son, S. D. Rane, PULSAR: A matlab toolbox for parallel magnetic resonance imaging using array coils and multiple channel receivers, Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering 31B (1) (2007) 24–36.

113. E. M. Haacke, R. W. Brown, M. R. Thompson, R. Venkatesan, Magnetic Resonance Imaging: Physical Principles and Sequence Design, Wiley-Liss, 1999.

114. L. Jiang, W. Yang, Adaptive magnetic resonance image denoising using mixture model and wavelet shrinkage, in: Proc. VIIth digital image computing: techniques and applications, Sydney, Australia, 2003, pp. 831–838.

115. A. Pižurica, W. Philips, I. Lemahieu, M. Acheroy, A versatile Wavelet domain noise filtration technique for medical imaging, IEEE Trans. Med. Imag. 22 (3) (2003) 323–331.

116. M. Lysaker, A. Lundervold, X. Tai, Noise removal using fourth order partial differential equations with applications to medical magnetic resonance imaging in space and time, IEEE Trans. Image Process. 12 (12).

117. T. McGraw, B. C. Vemuri, Y. Chen, M. Rao, T. Mareci, DT-MRI denoising and neuronal fiber tracking, Med. Imag. Anal. 8 (2004) 95–111.

118. C. B. Ahn, Y. C. Song, D. J. Park, Adaptive template filtering for signal-to-noise ratio enhancement in magnetic resonance imaging, IEEE Trans. Med. Imag. 18 (6) (1999) 549–556.

119. S. Awate, R. W. RT, Nonparametric neighborhood statistics for MRI denoising, in: Proc. Int. Conf. Inf. Process. Med. Imaging, Vol. 3565, 2005, pp. 677–688.

120. J. Hu, N. C. Beaulieu, Accurate closed-form approximations to Ricean sum distributions and densities, IEEE Communications letters 9 (2) (2005) 133–135.

121. S. M. Kay, Fundamentals of statistical signal processing: Estimation theory, Prentice Hall, New Jersey, 1993.

122. A. Papoulis, Probability, Random Variables and Stochastic Processes, Mc-Graw Hill, New York, NY, 1991.

123. N. C. Beaulieu, An infinite series for the computation of the complementary probability distribution function of a sum of independent ramdom variables and its application to the sum of Rayleigh random variables, IEEE Trans. on Comm. 38 (9) (1990) 1463–1473.

124. M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, ninth Edition, Dover, New York, 1964.

125. J. S. Lim, Two Dimensional Signal and Image Processing, Prentice Hall, Englewood Cliffs, NJ, 1990.

126. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

127. S. Aja-Fernández, R. San-José-Estépar, C. Alberola-López, C.-F. Westin, Image quality assessment based on local variance, in: Proc of the 28th IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC), New York, 2006, pp. 4815–4818.

128. Live database.
     URL http://live.ece.utexas.edu/research/quality/subjective.htm

129. S. Aja-Fernández, C. Alberola-López, C.-F. Westin, Signal LMMSE estimation from multiple samples in MRI and DT-MRI, in: Proceedings of MICCAI 07. Lecture Notes on Computer Science, Vol. 4792, 2007, pp. 368–375.

130. S. Aja-Fernández, K. Krissian, An unbiased non-local means scheme for DWI filtering, in: Proc. of the Workshop on Computational Diffusion MRI, MICCAI 2008, 2008, pp. 277–284.

131. D. Donoho, I. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (1994) 425–455.

132. J.-L. Starck, F. Murtagh, Automatic noise estimation from the multiresolution support, Publications of the Astronomical Society of the Pacific 110 (1998) 193–199.

133. K. Konstantinides, B. Natarajan, G. S. Yovanof, Noise estimation and filtering using block-based singular value decomposition, IEEE Trans. Image Process. 6 (3) (1997) 479–483.

134. M. Salmeri, A. Mencattini, E. Ricci, A. Salsano, Noise estimation in digital images using fuzzy processing, in: Proc. of the IEEE Int.l Conf. on Image Processing, Vol. 1, Thessaloniki (Greece), 2001, pp. 517–520.

135. D.-H. Shin, R.-H. Park, S. Yang, J.-H. Jung, Noise estimation using adaptive Gaussian filtering, IEEE Trans. on Consumer Electronics 51 (1) (2005) 218–226.

136. G. M. P. van Kempen, L. J. van Vliet, Background estimation in nonlinear image restoration, J. Opt. Soc. Am. A 17 (3) (2000) 425–433.

137. L.-C. Chang, G. K. Rohde, C. Pierpaoli, An automatic method for estimating noise-induced signal variance in magnitude-reconstructed magnetic resonance images, Medical Imaging 2005: Image Processing 5747 (1) (2005) 1136–1142.

138. N. Otsu, A threshold selection method from gray-scale histogram, IEEE Trans. on System, Man and Cybernetics 9 (1) (1979) 62–66.

139. T. Kurita, N. Otsu, N. Abdelmalek, Maximum likelihood thresholding based on population mixture models, Pattern Recognition 25 (10) (1992) 1231–1240.

140. A. Buades, B. Coll, J. Morel, A review of image denoising algorithms, with a new one, Multiscale Modeling and Simulation 4 (2) (2005) 490–530.

141. Y. Gal, A. Mehnert, A. Bradley, K. McMahon, D. Kennedy, S. Crozier, Denoising of dynamic contrast-enhanced MR images using dynamic nonlocal means, IEEE Trans. Med. Imag. 29 (2) (2010) 302 –310.

142. M. Caan, G. Khedoe, D. Poot, A. den Dekker, S. Olabarriaga, K. Grimbergen, L. van Vliet, F. Vos, Adaptive noise filtering for accurate and precise diffusion estimation in fiber crossings, in: T. Jiang, N. Navab, J. Pluim, M. Viergever (Eds.), Medical Image Computing and Computer-Assisted Intervention, MICCAI 2010, Vol. 6361 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2010, pp. 167–174.

143. Y. Rathi, M. Kubicki, S. Bouix, C.-F. Westin, J. Goldstein, L. Seidman, R. Mesholam-Gately, R. W. McCarley, M. E. Shenton, Statistical analysis of fiber bundles using multi-tensor tractography: application to first-episode schizophrenia, Magn. Reson. Imag. 29 (4) (2011) 507 – 515.

144. A. Tristán-Vega, C.-F. Westin, S. Aja-Fernández, Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging, NeuroImage 47 (2) (2009) 638 – 650.

145. T. Sumanaweera, G. Glover, S. Song, S. N. John Adler and, Quantifying MRI geometric distortion in tissue, Magn. Reson. Med. 31 (1) (2005) 40–47.

146. S. J. Kisner, T. M. Talavage, Testing the distribution of nonstationary MRI data, in: Proc. of the IEEE EMBS, San Francisco (CA), 2004, pp. 1888–1891.

147. L. Sendur, I. Selesnick, Bivariate shrinkage with local variance estimation, IEEE Signal Processing Letters 9 (12) (2002) 438–441.

148. S. Aja-Fernández, C. Alberola-López, On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering, IEEE Trans. Image Process. 15 (9) (2006) 2694–2701.

149. N. Coops, D. Culvenor, Utilizing local variance of simulated high spatial resolution imagery to predict spatial pattern of forest stands, Remote sensing of environment 71 (3) (2000) 248–260.

150. T. Royen, Exact distribution of the sample variance from a gamma parent distribution, Math.ST (Eprint: arXiv/0704.1415) (2007) 1–16.

151. G. S. Mudholkarand, C. T. Madhusudan, A gaussian approximation to the distribution of the sample variance for nonnormal populations, J. Amer. Statist. Assoc. 76 (1981) 479–485.

152. H. Solomon, M. A. Stephens, An approximation to the distribution of the sample variance, Canadian Journal of Statistics 11 (2) (1983) 149–154.

153. V. Bentkus, M. V. Zuijlen, Bounds for tail probabilities of the sample variance, Journal of Inequalities and Applications (2009) 1–20.

154. S. Hanson, B. Bly, The distribution of BOLD susceptibility effects in the brain is non-Gaussian, Neuroreport 12 (9) (2001) 1971–1977.

155. C.-C. Chena, C. Tylera, H. A. Baselera, Statistical properties of BOLD magnetic resonance activity in the human brain, NeuroImage 20 (2) (2003) 1096–1109.

156. A. Wink, J. B. T. M. Roerdink, BOLD noise assumptions in fMRI, International Journal of Biomedical Imaging (2006) 1–11.

157. H. Huang, J. Zhang, P. C. van Zijl, S. Mori, Analysis of noise effects on DTI-based tractography using the brute-force and multi-roi approach, Magn. Reson. Med. 52 (3) (2004) 559–565.

158. Y. Wang, Description of parallel imaging in MRI using multiple coils, Magn. Reson. Med. 44 (2000) 495–499.

159. A. Oppeinhein, R. Schafer, T. Stockham, Nonlinear filtering of multiplied and convolved signals, Proc. IEEE USA 56 (8) (1968) 1–264.

160. C. A. McKenzie, E. N. Yeh, M. A. Ohliger, M. D. Price, D. K. Sodickson, Self-calibrating parallel imaging with automatic coil sensitivity extraction, Magn. Reson. Med. 47 (3) (2002) 529–538.

161. P. J. Rousseeuw, C. Croux, Alternatives to the Median Absolute Deviation, J. Amer. Statist. Assoc. 88 (424) (1993) 1273–1283.