Original contribution

Scalar diffusion-MRI measures invariant to acquisition parameters: A first step towards imaging biomarkers

Santiago Aja-Fernández[a,*], Tomasz Pieciak[b], Antonio Tristán-Vega[c],
Gonzalo Vegas-Sánchez-Ferrero[d,e], Vicente Molina[f], Rodrigo de Luis-García[a]

[a] LPI, ETSI Telecomunicación, Universidad de Valladolid, Spain
[b] AGH University of Science and Technology, Kraków, Poland
[c] LPI, ETSI Industriales, Universidad de Valladolid, Spain
[d] ACIL, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[e] Biomedical Image Tech Lab, ETSI Telecomunicación, Universidad Politécnica de Madrid, Spain
[f] Psychiatry Service, Hospital Clínico Universitario, Valladolid, Spain

ARTICLE INFO

ABSTRACT

An imaging biomarker is a biologic feature in an image that is relevant to a patient's diagnosis or prognosis. In order to qualify as a biomarker, a measure must be robust and reproducible. However, the usual scalar measures derived from diffusion tensor imaging are known to be highly dependent on the variation of the acquisition parameters, which prevents their possible use as biomarkers. In this work, we propose a new set of quantitative measures based on diffusion magnetic resonance imaging from single-shell acquisitions that are designed to be robust to the variations of several acquisition parameters (number of gradient directions, *b*-value and SNR) while keeping a high discrimination power on differences in the diffusion characteristics of the tissue.

These new scalar measures are analytically obtained from a generic diffusion function that does not require the calculation of a diffusion tensor. This way, on one hand, we avoid the use of a specific diffusion model and, on the other hand, we make easier the statistical characterization of the measures. Accordingly, the analysis of the measures bias is carried out and it is used to minimize their dependency with respect to the acquisition noise for different SNRs. The robustness and discrimination power of the measures are tested for different number of gradients, *b*-values and SNRs using a realistic phantom and three real datasets: (1) 13 control subjects and different acquisition parameters; (2) a public data set from a single subject acquired using multiple shells and (3) 32 schizophrenia patients and 32 age and sex-matched healthy controls with a varying number of gradient directions.

The proposed quantitative measures exhibit low variability to the changes of the acquisition parameters, while at the same time they preserve a discrimination power that is able to detect significant changes in the anisotropy of the diffusion.

1. Introduction

Diffusion magnetic resonance imaging (DMRI) is an imaging technique that allows the quantification of the diffusivity of water molecules within the tissue *in vivo*. With different acquisition schemes and employing different models, it has been commonly used for the study of the properties of diffusion within tissues all throughout the body, with special relevance in abdominal organs and brain studies [1]. In the brain, particularly, DMRI has attracted extraordinary interest among the scientific community over recent years due to the relationships found between a number of neurological and neurosurgical pathologies

and alterations in the white matter as revealed by an increasing number of studies [2-4].

From the clinical research point of view, white matter studies using DMRI often rely on the comparison of scalar measures that describe the diffusion within a voxel. The most common measures based on the diffusion tensor (DT) model are the fractional anisotropy (FA), which measures to what extent a diffusion direction is dominant over the others, and the mean diffusivity (MD), which quantifies the total amount of diffusion. Other important scalar measures include the axial and radial diffusivity (AD, RD). However, and despite their popularity, these scalar measures show a high dependency with the acquisition

---

parameters that may produce different output values under different acquisition conditions. This effect makes them unfit to be used as a reliable clinical biomarkers.

An *imaging biomarker* is a biologic feature detectable in an image that is relevant to a patient's diagnosis. Opposite to traditional radiological diagnosis based on a qualitative assessment via visual inspection, Quantitative Imaging Biomarkers (QIBs) are quantifiable (numerical) features extracted from medical images which are able to assess the severity, change, or status of a disease, injury, or chronic condition relative to normal [5]. The key point is that a QIB is an objectively measurable characteristic, in contrast to a qualitative observation.

To qualify as a valid biomarker, a measure must fulfill the following requirements [5-7]:

1. *Accuracy*: it must show a small estimation bias or, at least, a constant bias for all measures.
2. *Precision*: the measurement must be repeatable for the same subject and between subjects, in the same session and between sessions.
3. *Robustness*: the measurement must be insensitive to acquisition parameters, platform and estimation algorithms.
4. *Reproducibility*: it must show low variability across sites and platforms.
5. *Clinical utility*: it must demonstrate an important role in diagnosis or treatment monitoring.

Following these criteria, we can state that the usual measures derived from the DT in DMRI present some serious drawbacks that limit their use as a QIB:

1. The inaccuracy of the diffusion model: the Gaussian assumption for the diffusion is unable to describe situations such as fiber crossings or fanning, which are known to be present, to a great extent, in the white matter.
2. Confounding and external factors, including changes in the acquisition parameters, noise, motion, etc.
3. Estimation and analysis techniques: the variability of the algorithms used for estimation [8], postprocessing [9] and analysis [10-13] is a factor in the variability of the results.

Among these issues, in this paper we will focus on the dependency of DT-derived measures with the acquisition parameters. It is well-known that the variation of the acquisition conditions produces different values on the estimated scalar measures. The parameters that most importantly affect the variability of the scalar measures are as follows:

1. The number of scan repetitions (NEX): an increase on the NEX produces an augmentation on the signal-to-noise ratio (SNR). Decreases in the FA values have been found in this situation [14]. The variation of the measures due to the reduction of the SNR has been reported in [15] and [16].
2. The diffusion sensitivity *b*-value: an increase of the *b*-value significantly decreases MD, AD and RD [17-20]. Small variations of the FA have also been reported [17, 21], especially in low SNR acquisitions [22].
3. Resolution: a smaller voxel size increases the values of FA [20, 23, 24]. Note that a smaller voxel size produces a lower SNR and also influences the population of fiber bundles that are present within the voxel.
4. Number of acquired gradient directions: decreases of the FA values have been reported when increasing the number of gradients [15,

16, 20, 24], together with a decrease in their variance [20, 25].

Other factors that can be a source of intra-session variability are the MR signal variation, subject physiological noise, motion and positioning [14, 26], the nonlinearity of diffusion gradients [27] or even the number of shells in a multi-shell acquisition [28]. In [20], we showed that, in addition, there is also a dependency with the area of the brain (bundle) and the specific analysis carried out.

Different approaches have been proposed to cope with this variability assuming an underlying tensor model. Vollmar et al. [29] proposed the use a global scaling factor to reduce the coefficient of variation inter- and intra-site, after parcellation. Similarly, in [30] and [31] the authors proposed the weighing of the datasets acquired in different sessions and/or scanners based on their variability. In [32], it was found that inter-scanner and field-strength effects over the scalar measures can be reduced with linear correction factors specific to the different regions of interest. The study in [33] employed statistical models that include as covariate factors some sources of variability, such as scanner manufacturer, magnetic field strength and number of gradient directions. Pohl et al. [34] also proposed a harmonization of the measures to cope with inter-scanner variability based on human-phantom data. Finally, an alternative solution to multi-source data harmonization is the representation of the diffusion signal an alternative space, like a Spherical Harmonics (SH) basis. Using the SH coefficients, Mirzaalian et al. [35, 36] compute a set of rotation invariant features that can be used to estimate a specific linear mapping between the signal from different scanners.

As previously stated, one important source of variability is the inaccuracy of the diffusion model adopted. The search for appropriate models of diffusion beyond tensor representations has been a very active research field the last decade. The number of degrees of freedom to represent for these models is far larger than the traditional 6 free components of the diffusion tensor, so a great number of diffusion gradients needs to be acquired. Some of the most popular approaches are diffusion spectrum imaging (DSI) [37] (which aims to completely characterize the diffusion process by means of the sparse sampling for all possible orientations and magnitudes of the diffusion sensitizing gradients) and high angular resolution diffusion imaging (HARDI), which characterizes all possible directions of the gradients for one given magnitude using different approaches like multi-tensor models [38], generalized tensor models [39], Q-balls [37, 40]. These techniques, although rather experimental, have been successfully used to resolve complex architectures of the white matter, including fiber crossing, bending, and kissing. Many different scalar measures can be derived from HARDI and DSI information, which can add relevant structural information and meaningful descriptive maps of the white matter. Among these measures some of the most popular are the generalized anisotropy (GA) [41], generalized fractional anisotropy (GFA) [42], return-to-origin probability (RTOP) [43,44] and the return-to-plane and return-to-axis probabilities (RTPP, RTAP) [43]. Most of these measures require a dense sampling of the **q**-space or sampling at high *b*-values.

In the present paper, we will stick to *standard* diffusion tensor imaging (DTI) acquisitions: *b*-values around $1000 \, s/mm^2$ and low to moderate number of gradients. We propose a new set of measures that describe the characteristics of the diffusion, while at the same time they are robust to the acquisition parameters. For the sake of simplicity, we will confine ourselves to only three sources of variability: (1) the number of acquired gradients; (2) the variation of the *b*-value and (3) changes in the SNR. The measures proposed here are more robust to the variation of these parameters than FA and MD, while they keep a discrimination power similar to them. To show their potential, the proposed new measures are validated over various experiments on a

pseudo-synthetic phantom and three real datasets.

2. Theory

2.1. The diffusion signal

In DMRI, the probability density function (PDF) of the displacement of water molecules in a given direction, $P(\mathbf{R}|\Delta)$, is related to the signal provided by the MRI scanner at each image location $E(\mathbf{q})$ through the Fourier transform

$$P(\mathbf{R}|\Delta) = \mathfrak{F}\{|E(\mathbf{q})|\}(\mathbf{R}) = \int_V E(\mathbf{q})\exp(-2\pi j\mathbf{q}^T\mathbf{R})d\mathbf{q}. \tag{1}$$

The inference of the exact information on the $\mathbf{R}$-space would require the sampling of the whole $\mathbf{q}$-space to use the Fourier relationship between both spaces. In order to obtain a closed form solution from a reduced number of diffusion-weighted imaging (DWI) datasets, a Gaussian model for the diffusion is usually adopted. A Gaussian diffusion propagator can be assumed in case all fiber bundles within the voxel resolution are nearly identically distributed. In this case, $P(\mathbf{R}|\Delta)$ is a mixture of independent and (nearly) identically distributed bounded cylinder statistics and, by virtue of the central limit theorem, their superposition is Gaussian distributed:

$$P(\mathbf{R}|\Delta) = \frac{1}{\sqrt{|\mathscr{D}|(4\pi\Delta)^3}} \exp\left(\frac{-\mathbf{R}^T\mathscr{D}^{-1}\mathbf{R}}{4\Delta}\right). \tag{2}$$

The measured signal in the $\mathbf{q}$-space is the (inverse) Fourier transform of the PDF in Eq. (2):

$$E(\mathbf{q}) = \mathfrak{F}^{-1}\{P(\mathbf{R}|\Delta)\}(\mathbf{q}) = \exp(-4\pi^2\tau\mathbf{q}^T\mathscr{D}\mathbf{q}), \tag{3}$$

which is the well-known Stejskal–Tanner equation [45]. The diffusion tensor $\mathscr{D}$ is the anisotropic covariance matrix of $P(\mathbf{R}|\Delta)$, and therefore is a symmetric positive–definite matrix, with positive eigenvalues and orthonormal eigenvectors. Eq. (3) is usually re-written in terms of the b-value, under certain assumptions:

$$E(\mathbf{q}) = \exp(-b\mathbf{g}^T\mathscr{D}\mathbf{g}) \tag{4}$$

$$b = 4\pi^2\tau\|\mathbf{q}\|^2 = \tau\gamma^2\delta^2\|\mathbf{G}\|^2, \tag{5}$$

where $\mathbf{G} = \frac{2\pi}{\gamma\delta}\mathbf{q}$ and $\mathbf{g} = \mathbf{G}/\|\mathbf{G}\| = \mathbf{q}/\|\mathbf{q}\|$ is the gradient direction. For more complex approximations to diffusion analysis beyond the DT, a more general expression for $E(\mathbf{q})$ can be used ($q_0 = \|\mathbf{q}\|$, θ, ϕ are the angular coordinates in the spherical system):

$$E(\mathbf{q}) = \exp(-4\pi^2\tau q_0^2 D(q_0, \theta, \phi)), \tag{6}$$

where the positive function D is the apparent diffusion coefficient (ADC). It is common to consider that it does not depend on $\|\mathbf{q}\|$, i.e. $D(\theta,\phi)$ and therefore

$$E(\mathbf{q}) = \exp(-4\pi^2\tau q_0^2 D(\theta, \phi)). \tag{7}$$

This assumption is implicitly done when considering the diffusion tensor model, and it is also common in HARDI models [39, 40, 46] where a marginalization along the radial component is necessary. The acquired signal in the scanner is not directly $E(\mathbf{q})$, but a scaled version of it, $S(\mathbf{g})$:

$$S(\mathbf{g}) = S(0)\cdot E(\mathbf{q}), \tag{8}$$

where $\mathbf{g} = \mathbf{q}/\|\mathbf{q}\|$ and $S(0)$ is the so-called baseline image, i.e., the acquired signal when no diffusion gradients are applied.

2.2. Scalar diffusion measures

Traditional scalar measures derived from DTs suffer from variability due to their dependence on the acquisition parameters. The FA, for instance, is known to highly depend on the number of gradients, on the voxel size and on the SNR [20]. We can explicitly write this dependence as

$$FA = FA(\mathbf{x}; \Theta), \tag{9}$$

where Θ stands for the different acquisition parameters, i.e.

$$\Theta = \{N_g, b, SNR, r, \Upsilon\},$$

where N_g is the number of gradient directions, b is the b-value, r is the resolution (voxel size) and Υ represents other unaccounted factors such as the scanner manufacturer or the strength of the magnetic field. As discussed in Section 1, here we will restrict ourselves to three parameters: the number of gradients, b and the SNR:

$$FA = FA(\mathbf{x}; N_g, b, SNR).$$

Other scalar measures also depend on the acquisition parameters, although in a possibly completely different manner. For instance, changes in the MD are considerable when the b-value varies, but not so noticeable for changes in the number of gradients [20].

2.3. Robust scalar diffusion measures

Our purpose is to define a series of new diffusion measures that (a) do not depend on the acquisition parameters (number of gradients, b and SNR) and (b) that are able to encode the diffusion characteristics in such a way that differences induced by pathologies can be found. To that end we propose a series of scalar measures designed with the following desirable properties in mind:

1. They can be calculated directly from the DWIs, without relying on the tensor model.
2. They must be robust to the changes of the considered parameters.
3. The bias of the measure must not depend on the considered parameters.
4. The measures must be able to describe the amount and/or anisotropy of diffusion.
5. They must show a discriminant power similar to that of FA and MD: differences in white matter group studies found using FA and MD must also be found with the new measures.

The proposed scalar measures are the following:

1) **Diffusion volume (DV):** the value of $P(\mathbf{R}|\Delta)$ in the origin is related to the volume of $E(\mathbf{q})$:

$$P(\mathbf{0}|\Delta) = \int_V E(\mathbf{q})d\mathbf{q}. \tag{10}$$

This value is the probability density of zero displacement (also known as RTOP [43]), which is related to the probability density of water molecules that minimally diffuse within the diffusion time Δ, a measure known to provide relevant information about the white matter structure [47-50]. Assuming a Gaussian diffusion model, we can write Eq. (10) as

$$P(\mathbf{0}|\Delta) = \frac{1}{\sqrt{|\mathscr{D}|(4\pi\Delta)^3}} \propto \frac{1}{b^{3/2}\cdot|\mathscr{D}|^{1/2}}. \tag{11}$$

The determinant of the tensor $|\mathscr{D}|$ can be seen as the volume of the ellipsoid that represents the diffusion, and thus it can be used as a

measure of the diffusion. Assuming that the diffusion signal $D(\mathbf{q})$ does not depend on the radial direction, we can approximate the integral by the summation of the values sampled in a particular shell (see Appendix A):

$$\int_V E(\mathbf{q})d\mathbf{q} \approx \sum_{i=1}^{N_g} \frac{\sqrt{\pi}}{2c_0^{3/2}D_i^{3/2}}\Delta S, \tag{12}$$

where ΔS is an increment of a sphere, c_0 is a constant and D_i are the sampled diffusion values in a particular direction:

$$D_i = \frac{-\log E(\mathbf{q}_i)}{b}, \quad i = 1, ..., N_g. \tag{13}$$

We define the diffusion volume proportional to $\sqrt{|\mathscr{D}|}$ and thus related to the volume of the ellipsoid as

$$\begin{aligned} \mathrm{DV}(\mathbf{x}) &= \frac{1}{b^3}\frac{1}{N_g}\sum_{i=1}^{N_g}(-\log E(\mathbf{q}_i))^{3/2} \\ &= \frac{1}{b^3}\left\langle (-\log E(\mathbf{q}_i))^{3/2} \right\rangle, \end{aligned} \tag{14}$$

where $\langle . \rangle$ is the averaging operator:

$$\left\langle I(\mathbf{x}_i) \right\rangle = \frac{1}{N_g}\sum_{i=1}^{N_g} I(\mathbf{x}_i). \tag{15}$$

Note that the DV would depend on the parameter b. To make the measure robust to changes in that parameter we redefine it as

$$\mathrm{DV}(\mathbf{x}) = \frac{1}{b^{3/2}}\langle (-\log E(\mathbf{q}_i))^{3/2} \rangle, \tag{16}$$

Although this measure was initially derived using a Gaussian assumption, note that the final equation does not consider that specific model, but a generic diffusion.

2) **Average sample diffusion (ASD):** In order to obtain a measure of the *amount* of diffusion we propose to calculate the integral of the diffusion in the surface of a sphere of radius q_0:

$$\int_{S_{q_0}} D(\theta, \phi)dS, \tag{17}$$

assuming that the diffusion function does not depend on the radial coordinate. Similar to the DV, we can approximate Eq. (17) as

$$\int_{S_{q_0}} D(\theta, \phi)dS \approx 2\sum_{i=1}^{N_g} D_i \Delta S.$$

Considering a uniform sampling of the sphere, i.e. $\Delta S \propto 1/N_g$, we have

$$\int_{S_{q_0}} D(\theta, \phi)dS \propto \frac{1}{N_g}\sum_{i=1}^{N_g} D_i = \frac{-1}{b}\frac{1}{N_g}\sum_{i=1}^{N_g}\log E(\mathbf{q}_i).$$

This value can be seen as an averaging of the sampled values of the diffusion on a particular shell, i.e., the average sampled diffusion:

$$\mathrm{ASD}(\mathbf{x}) = \frac{1}{b}\langle -\log E(\mathbf{q}_i) \rangle. \tag{18}$$

3) **Diffusion energy:** we can define the energy of the diffusion signal in a specific shell:

$$\int_{S_{q_0}} |D(\theta, \phi)|^2 dS.$$

The discrete version of the measure is obtained as follows:

$$\begin{aligned} \int_{S_{q_0}} |D(\theta, \phi)|^2 dS &\approx 2\sum_{i=1}^{N_g} D_i^2 \Delta S \propto \frac{1}{N_g}\sum_{i=1}^{N_g} D_i^2 \\ &= \frac{1}{b^2}\frac{1}{N_g}\sum_{i=1}^{N_g}(\log E(\mathbf{q}_i))^2. \end{aligned}$$

This value can be seen as the sampled second order moment of the diffusion on a shell:

$$\mathrm{SMD}_2(\mathbf{x}) = \frac{1}{b^2}\langle (\log E(\mathbf{q}_i))^2 \rangle. \tag{19}$$

4) **Coefficient of variation of the diffusion (CVD):** the FA can be seen as the square root of the variance of the eigenvectors of the DT divided by their second-order moment. Accordingly, the same idea can be applied to the diffusion signal[1]:

$$\mathrm{CVD}(\mathbf{x}) = \sqrt{\frac{\mathcal{V}(\log E(\mathbf{q}_i))}{\langle (-\log E(\mathbf{q}_i))^2 \rangle}}, \tag{20}$$

where $\mathcal{V}(.)$ is the sample variance defined as

$$\mathcal{V}(I(\mathbf{x}_i)) = \frac{N_g}{N_g - 1}[\langle I(\mathbf{x}_i)^2 \rangle - \langle I(\mathbf{x}_i)\rangle^2].$$

This measure can also be seen as an alternative implementation of the generalized anisotropy [41, 51]. An overview of all the proposed measures is presented in Table 1. Note that DV, ASD and SMD_2 are designed to mainly quantify the amount of diffusion, while CVD quantifies the anisotropy of the diffusion.

2.4. Bias and variance analysis

The measures previously defined will be subject to acquisition artifacts, and they will be corrupted with noise. The presence of noise will introduce a bias in the estimator that depends on the *amount of noise* and may also depend on other acquisition parameters. Thus, in order to obtain robust measures, this bias must be properly corrected.

The acquired signal $E(\mathbf{q}_i)$ is defined as an acquired T2 signal weighted by some oriented gradient. If we define the acquired signal as $S_i(\mathbf{x})$, we can define $E(\mathbf{q}_i)$ as

$$E(\mathbf{q}_i) = E_i(\mathbf{x}) = \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})},$$

where $S_0(\mathbf{x})$ is the baseline acquisition. For the sake of simplicity, let us assume that the acquired signals $S_i(\mathbf{x})$ and $S_0(\mathbf{x})$ are corrupted with Rician noise [52]:

$$S_i(\mathbf{x}) = |A_i(\mathbf{x}) + N(\mathbf{x}; 0, \sigma^2)|, \quad i = 0, \cdots, N_g,$$

where $A_i(\mathbf{x})$ is the original signal if no noise is present and $N(\mathbf{x};0,\sigma^2)$ is a complex additive Gaussian noise with zero mean and variance σ^2. This is a common assumption in MRI acquisitions, valid for single-coil acquisitions and multi-coil parallel imaging reconstructed with a spatial matched filter, like SENSE, for instance. In the latter, noise can become *non-stationary*, i.e., the variance of noise will depend on the position and σ must be replaced by $\sigma(\mathbf{x})$, which does not affect to the following study.

The analytical study of the bias and variance of the proposed measures is described in Appendix B. Results are summarized in Table 2. Note that the ASD is an unbiased measure, while the bias of the other measures depends on the b-value, the variance of noise σ^2 and the original value of the baseline $A_0(\mathbf{x})$ and DWIs $A_i(\mathbf{x})$. Consequently, a change in the b-value or the SNR may introduce an undesired different bias to the measure. Luckily, this bias can be easily corrected with the proper estimation of some of the parameters: σ can be estimated from the baseline [52] and the value of $A_0(\mathbf{x})$ can be obtained from a filtered version of $S_0(\mathbf{x})$ or from several acquired repetitions of the baseline. To estimate the averaged value of $A_i^2(\mathbf{x})$ we can assume a high SNR

---

[1] Note that the CV is usually defined as $\mathrm{CV} = \frac{\sqrt{\mathrm{Var}\{X\}}}{\mathbb{E}\{X\}} = \sqrt{\frac{\mathbb{E}\{X^2\}}{\mathbb{E}\{X\}^2} - 1}$. We have considered an alternative definition: $\mathrm{CV} = \frac{\sqrt{\mathrm{Var}\{X\}}}{\sqrt{\mathbb{E}\{X^2\}}} = \sqrt{1 - \frac{\mathbb{E}\{X\}^2}{\mathbb{E}\{X^2\}}}$ to make the measure similar to the FA.

Table 1

Survey of the proposed diffusion measures.

Measure	Formula	Equation
Diffusion volume	$\mathrm{DV}(\mathbf{x}) = \frac{1}{b^{3/2}} \langle (-\log E(\mathbf{q}_i))^{3/2} \rangle$	(16)
Average of sampled diffusion	$\mathrm{ASD}(\mathbf{x}) = \frac{1}{b} \langle -\log E(\mathbf{q}_i) \rangle$	(18)
Diffusion energy	$\mathrm{SMD}_2(\mathbf{x}) = \frac{1}{b^2} \langle (\log E(\mathbf{q}_i))^2 \rangle$	(19)
Diffusion coefficient of variation	$\mathrm{CVD}(\mathbf{x}) = \sqrt{\frac{\mathcal{V}(\log E(\mathbf{q}_i))}{\langle (-\log E(\mathbf{q}_i))^2 \rangle}}$	(20)

Table 2

Bias and variance of the proposed measures, assuming a Rician distribution for the composite magnitude signal.

Measure	Bias	Variance
$\mathrm{DV}(\mathbf{x})$	$\frac{3}{8} \frac{\sigma^2}{b^{3/2}} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \left(\log \frac{A_0}{A_i} \right)^{-\frac{1}{2}} \right\rangle$	$\frac{1}{b^3} \frac{9\sigma^2}{4N_g} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \log \frac{A_0}{A_i} \right\rangle + \frac{1}{b^3} \frac{9\sigma^4}{64 N_g} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right)^2 \left(\log \frac{A_0}{A_i} \right)^{-1} \right\rangle$
$\mathrm{ASD}(\mathbf{x})$	0	$\frac{1}{b^2} \frac{\sigma^2}{N_g} \left(\left\langle \frac{1}{A_i^2} \right\rangle + \frac{1}{A_0^2} \right)$
$\mathrm{SMD}_2(\mathbf{x})$	$\frac{\sigma^2}{b^2} \left(\left\langle \frac{1}{A_i^2} \right\rangle + \frac{1}{A_0^2} \right)$	$\frac{1}{b^4} \frac{4\sigma^2}{N_g} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \left(\log \frac{A_i}{A_0} \right)^2 \right\rangle - \frac{1}{b^4} \frac{\sigma^4}{N_g} \left\langle \frac{1}{A_i^2 A_0^2} \right\rangle$
$\mathrm{CVD}^2(\mathbf{x})$	No closed expression	$\frac{4\sigma^2}{N_g} \frac{\left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \left(\log \frac{A_i}{A_0} \right)^2 \right\rangle - \frac{\sigma^2}{4} \left\langle \frac{1}{A_i^2 A_0^2} \right\rangle}{\left(\left\langle \left(\log \frac{A_i}{A_0} \right)^2 \right\rangle + \left\langle \frac{\sigma^2}{A_i^2} \right\rangle + \frac{\sigma^2}{A_0^2} \right)^2} + \frac{4\sigma^2}{N_g} \frac{4 \left\langle \log \frac{A_i}{A_0} \right\rangle^2 \left(\left\langle \frac{1}{A_i^2} \right\rangle + \frac{1}{A_0^2} \right)}{\left(\left\langle \left(\log \frac{A_i}{A_0} \right)^2 \right\rangle + \left\langle \frac{\sigma^2}{A_i^2} \right\rangle + \frac{\sigma^2}{A_0^2} \right)^2}$

scenario, restrict ourselves to a second order expansion and use the approximation

$$\langle A_i^{-2}(\mathbf{x}) \rangle \approx \frac{1}{\langle A_i^2(\mathbf{x}) \rangle}.$$

Using the second order moment of a Rician random variable [52], we can write

$$\langle A_i^2(\mathbf{x}) \rangle = \langle S_i^2(\mathbf{x}) \rangle - 2\sigma^2.$$

More accurate unbiased estimators could be achieved by using more terms in the approximations considered.

Finally, note that the bias of the measures does not depend on the number of gradient directions. On the other hand, their variance is proportional to the level of noise σ^2, as expected, and is inversely proportional to the number of gradients. Thus, employing more gradient directions will not affect the bias of the measures, but it will decrease the variance of such measures.

3. Methods

In order to validate the proposed methods, a series of experiments have been carried out using pseudo-synthetic data and real acquisitions.

3.1. DWI phantom

For the sake of comparison to a gold standard, a realistic DWI phantom is used, specifically the one proposed in [53, 54]. The phantom was generated from a real DWI dataset from a SENSE EPI acquisition scanned with a 3T GE system. It consists of a $256 \times 256 \times 81$ volume with 8 baseline volumes and 51 gradient directions and a resolution of $0.94 \times 0.94 \times 1.7\,\mathrm{mm}^3$. A denoising/regularization process was carried out in order to obtain a noise-free volume:

1. Unbiased non-local means denoising for each DWI channel and baseline average using [54].
2. Normalization of the intensities so that the baseline shows gray values in the range [0–255].
3. Regularization of the DWI dataset across the gradient directions. To that end, the diffusion signal $D(\mathbf{q})$ is expressed in the basis of SH and reconstructed for an arbitrary gradient directions.
4. Different b-values are simulated assuming that the diffusion signal does not depend on the radial component and therefore $E(b_1) = [E(b_0)]^{b_1/b_0}$, with $b_0 = 1200$. This approximation is only valid for moderate variations of b, in the range where the Gaussian
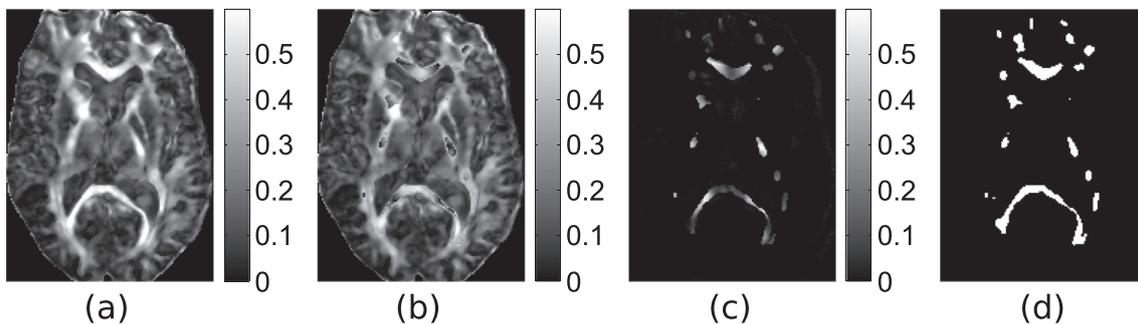
Fig. 1. Original and modified FA maps obtained from 51-gradient volume: (a) original FA, (b) modified FA, (c) absolute difference of (a) and (b), and (d) mask used to modify the anisotropic diffusion in certain ROIs of the image.

Table 3
Demographic data for the employed dataset shown as mean (standard deviation).

	Schizophrenia	Healthy controls
	N = 32	N = 32
Age	36.62 (8.86)	35.00 (10.44)
Sex (M:F)	22:10	21:11

diffusion holds.

5. In order to change the SNR, the final data is corrupted with Rician noise [52] with different standard deviation of noise σ.

This pseudo-synthetic phantom is used in two experiments:

1. *DWI phantom - robustness test*: In order to assess the robustness of measures to the change of parameters, 50 repetitions of the volume with individual noise realization will be considered for each set of parameters (number of gradients, b-value and σ of noise). A T-test is carried out pointwise to detect differences between the same volume with different acquisition parameters for each measure. To quantify the differences, we define the following ratio:

$$R(p_0) = \frac{\text{Number of points with } p < p_0}{\text{Number of points}}. \tag{21}$$

Only those voxels belonging to the white matter are considered. A value of $R(p_0)$ close to 1 indicates that most of the points in the image are considered to be significantly different, while a value closer to 0 indicates that differences cannot be assured. For the experiments, we will consider $p_0 = 0.01$.

2. *DWI phantom - discriminant power test*: So as to test the discriminant power of the methods, a mask delineating several white matter areas with high FA was created. Then, the value of coefficients 2 to 4 of the SH decomposition was reduced in those ROIs, thus decreasing the anisotropy of the diffusion (see Fig. 1). Two different data sets were made available this way, one with the original diffusion and another with a reduced anisotropy. Different acquisition parameters were considered for each set, and differences are evaluated using an ANOVA test, similar to the previous case.

3.2. Real data

Next, three real datasets are employed to verify two desired properties of the proposed diffusion measures: their robustness to changes in the acquisition parameters and their ability to detect differences in white matter group studies:

1. *Real dataset 1*: MR images were acquired from thirteen healthy male adults, aged between 23 and 31 (average age 27 years), as detailed in [20]. Images were acquired in a Philips 1.5T unit. DWIs were acquired using a multi-shot pseudo-3D double spin-echo echo-planar imaging (SE-EPI) sequence with $b = 800 \, \text{s/mm}^2$, spatial resolution $2 \times 2 \times 2 \, \text{mm}^3$, matrix size 256×256, a total of 66 sections, TE = 1.6 ms, TR = 8 ms. 61 gradient directions and one baseline volume were acquired. The gradient directions were specifically designed so that they can be subsampled to 40, 21 or 6 gradient directions while remaining equally spaced for each configuration. This subsampling technique allows the measurement of the effect of using different number of gradients with only one acquisition.

2. *Real dataset 2*: This is a publicly available dataset acquired from one single volunteer using a Siemens Trio 3T unit [55]. Acquisition parameters included the following: double spin echo DW EPI sequence, 33 gradient directions, 15 b-values (only three will be considered for this experiment: $b = (800, 1000 \text{ and } 1200) \, \text{s/mm}^2$),

19 consecutive slices, isotropic resolution of 2.5 mm, matrix size 96×96, SNR = 39 at baseline.

3. *Real dataset 3*: MRI data were obtained from 32 schizophrenia patients (SZ) and 32 age and sex-matched healthy controls (HC; Table 3). Acquisitions were carried out using a Philips Achieva 3T unit at the MRI facility at University of Valladolid, including T1-weighted and diffusion-weighted images. For the anatomical T1-weighted images, acquisition parameters included the following: TFE sequence, 256×256 matrix size, $1 \times 1 \times 1 \, \text{mm}^3$ of spatial resolution and 160 slices covering the whole brain. With regard to the DWIs, the acquisition parameters were 61 gradient directions, one baseline volume, b-value = $1000 \, \text{s/mm}^2$, $2 \times 2 \times 2 \, \text{mm}^3$ of voxel size, 128×128 matrix and 34 slices covering the entire brain. The scheme of 61 gradient directions is designed in such a way that it can be subsampled into a set of valid 40 gradient directions or a set of 21 valid gradient directions (same scheme used for Dataset 1). More information about the clinical details of participants, as well as inclusion/exclusion criteria, can be found in [56].

Three different experiments were carried out:

1. *Real dataset 1 - robustness to number of gradients*: Several regions of interest within the white matter were automatically delineated by projecting these regions from a white matter atlas to the native space of each acquisition, using *flirt* and *fnirt* utilities from FSL[2]. The regions of interest considered were genu of corpus callosum, body of corpus callosum, splenium of corpus callosum, right Cingulum (cingulate gyrus) and left Cingulum (cingulate gyrus). The average values of the different scalar measures (FA, MD and the proposed measures) were computed for each ROI using different configurations of gradient directions and compared.

2. *Real dataset 2 - robustness to b-value*: Three different shells from the Kurtosis acquisition were considered. From them, the different scalar measures were calculated for a single slice and an error measure was computed as

$$\text{Error}\left(b_i, b_j\right) = \frac{|M(b_i) - M(b_j)|}{M(b_{1000})},$$

where $M(b_i)$ is the considered measure for b-value b_i, and $M(b_{1000})$ is the measure for a b-value equal to $1000 \, \text{s/mm}^2$, which is considered as a reference.

3. *Real dataset 3 - discrimination power and robustness to number of gradients*: data set 3 is used to test the capability of the measures to find differences between groups. The rationale of the experiment is the following: differences between SZ and HC using DMRI have been extensively found in the literature [2, 56, 57], showing a widespread reduction in FA and an increase in MD in patients with respect to controls. Therefore, a comparison performed using a certain common acquisition scheme should find these differences. However, if the FA and MD are not robust with respect to changes in the number of gradient directions, comparisons where the number of gradient directions changes in different groups could yield misleading results, overestimating the differences, underestimating them or even finding differences in the opposite direction. On the contrary, if the proposed new scalar measures are robust with respect to these changes, then the results will be stable across different comparisons. On top of that, for the new measures to be useful it is also necessary that they are sensitive to pathological conditions, that is, they are able to find differences between patients and healthy controls.
Following this idea, a group study was performed using TBSS (tract-based spatial statistics) [58, 59] in order to discover differences between the SZ and HC groups. Data with different number of gradients were employed in each group (SZ and HC), yielding several

---

[2] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki.

Table 4
The combinations of gradient directions used for experiments on real data.

Schizophrenia	Healthy controls
40	40
21	40
40	21
21	21
21	61
61	21

combinations: for instance, using 61 gradient directions for the subjects in the SZ group and 21 gradient directions for the subjects in the HC group, using 40 gradient directions for the subjects in the SZ

group and 61 gradient directions for the subjects in the HC group, and so on. The combination of 61 gradient directions for both the SZ and HC groups was considered to be the golden standard, and the other combinations (included in Table 4) were compared to this one. In order to focus on the performance of the diffusion measures and not in other factors such as the registration process performed in TBSS, registrations to a common template were performed only for the golden standard and then applied to other combinations.

TBSS was the method of choice in this experiment because of two reasons. First, for being a popular and commonly used white matter analysis method. Second, most of other alternatives, such as tractography-based methods, for instance, inherently depend on multiple factors that we wish to avoid as much as possible. The implications and possible limitations of the choice of TBSS are discussed in Section 5.
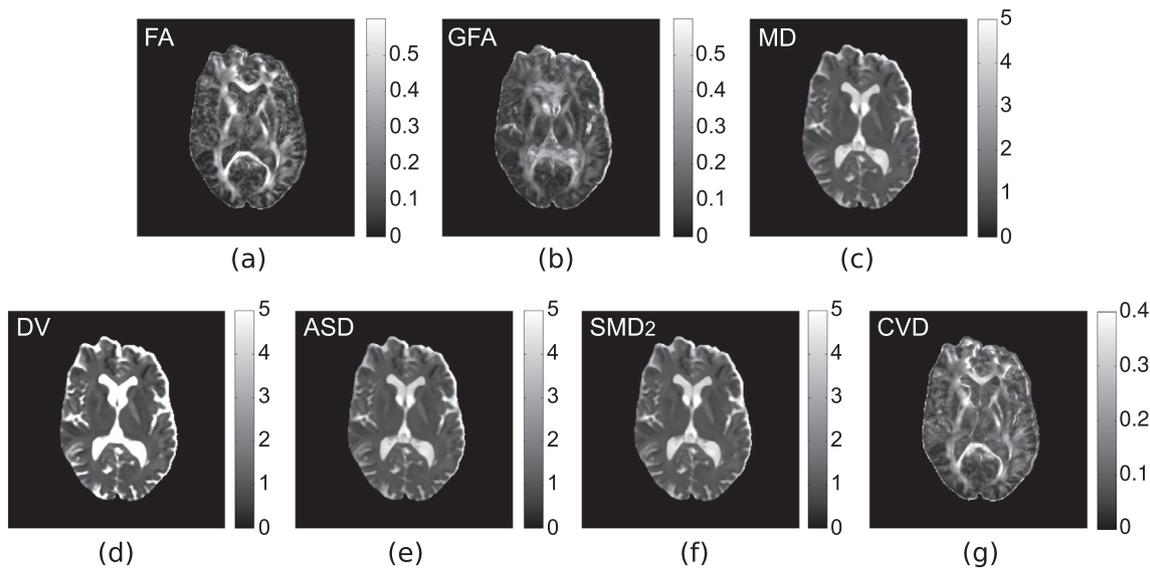
Fig. 2. Diffusion scalar measures obtained from axial slice of a realistic phantom with 51 gradients: (a) FA, (b) GFA, (c) $MD \times b$, (d) $DV \times b^{3/2}$, (e) $ASD \times b$, (f) $(SMD_2)^{1/2} \times b$, and (g) CVD.
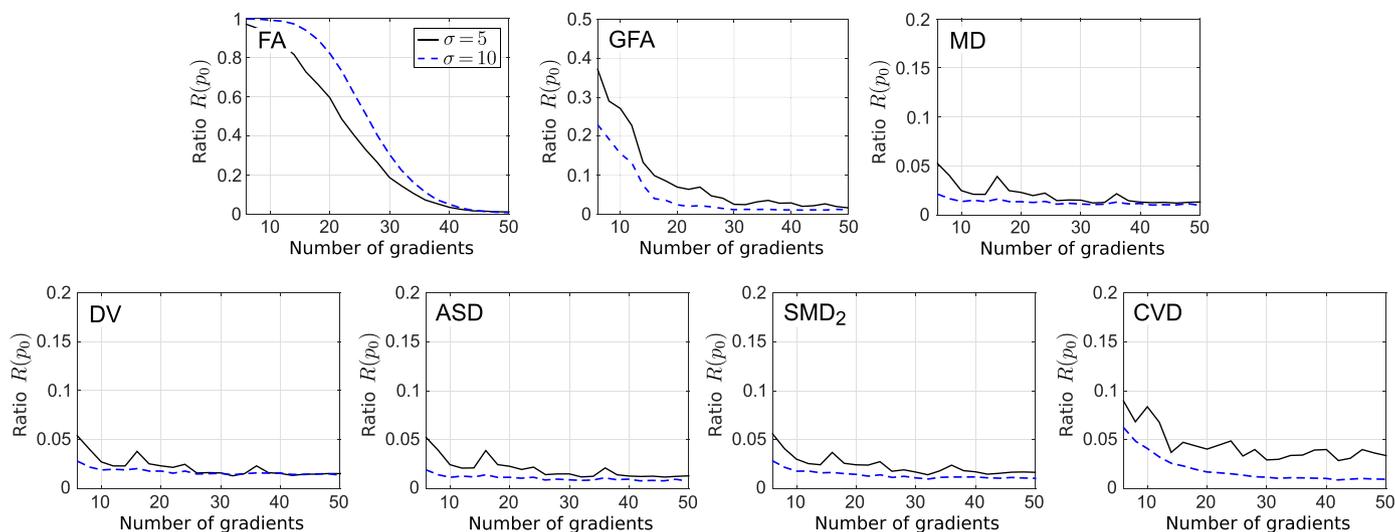
Fig. 3. Ratio $R(p_0)$ of points assumed to be different in the pseudo-synthetic phantom for variable number of gradients when compared to the Golden Standard (51 gradient directions). Two levels of noise are used: $\sigma = 5$ (black) and $\sigma = 10$ (blue). For all the cases $b = 1200$ s/mm². (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
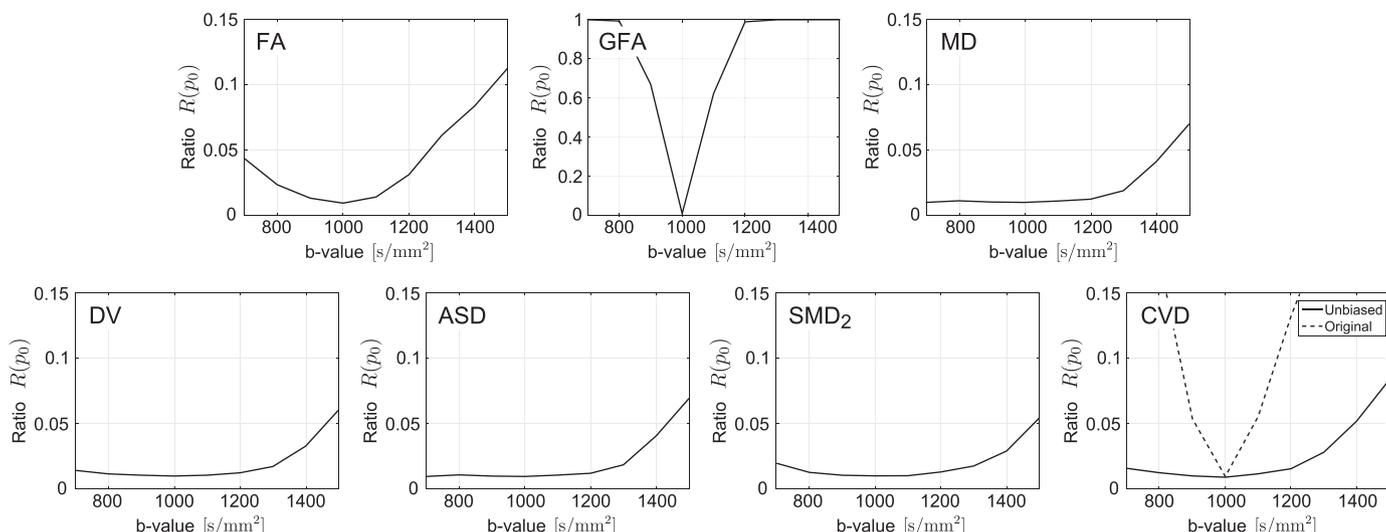
Fig. 4. Ratio $R(p_0)$ of points assumed to be different in the pseudo-synthetic phantom for variable b-value when compared to the Golden Standard ($b = 1000$ s/mm^2). For CVD the biased and unbiased versions are depicted. In all the cases 41 gradient directions have been used and $\sigma = 3$.
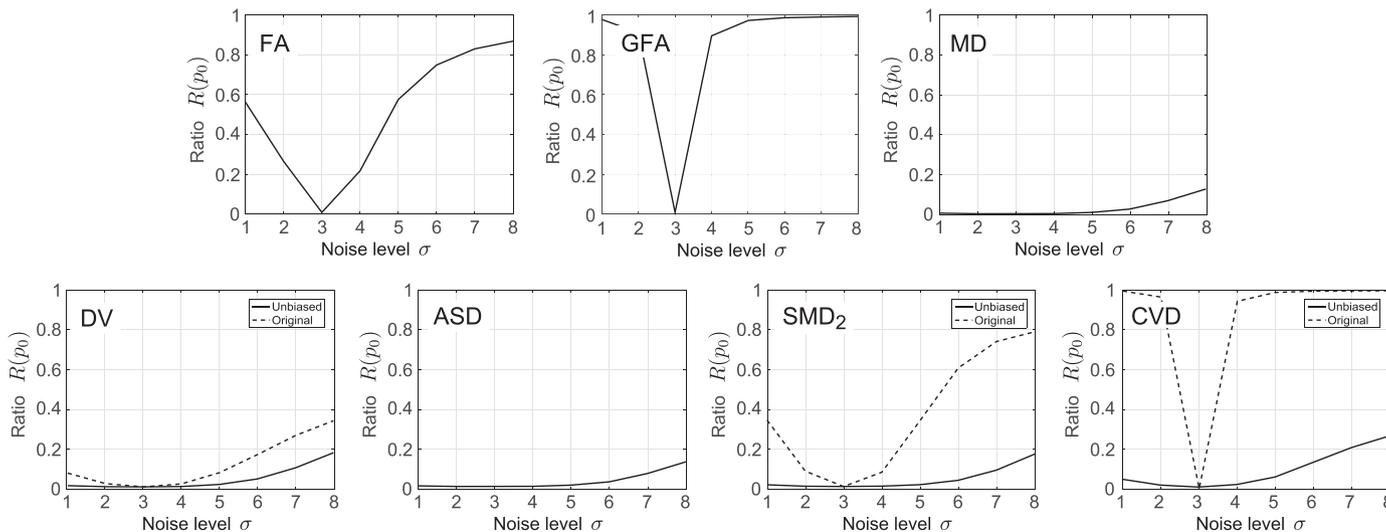
Fig. 5. Ratio $R(p_0)$ of points assumed to be different in the pseudo-synthetic phantom for variable SNR, expressed in terms of σ, when compared to the Golden Standard ($\sigma = 3$). For DV, SMD$_2$, and CVD the biased and unbiased versions are depicted. In all the cases 41 gradient directions have been used and $b = 1000$ s/mm^2.

4. Results

4.1. Pseudo-synthetic phantom

For the sake of visual understanding of the proposed measures, those are depicted in Fig. 2 together with the FA, GFA and MD. We have added FA and MD for being the most used scalar values in DTI analysis. We have also added a non-tensor-based measure, GFA, which calculates the anisotropy using the SH components. The measures in Fig. 2 have been scaled to make them visually comparable. Note that the DV and the sample moments of the diffusion produce results visually similar to the MD, while the CVD shows the anisotropy structure, as expected.

Results of the *DWI phantom - robustness test* (robustness of the proposed measures with respect to changes in the parameters using the pseudo-synthetic phantom) in terms of the ratio $R(p_0)$ defined in Eq. (21), are presented in Figs. 3 (number of gradients), 4 (b-value), and 5 (SNR).

With regard to the traditional scalar measures, results show that the FA is greatly variant with respect to the three parameters, most significantly when changing the number of gradients. GFA shows a greater performance than FA, but it still shows a great variation for small number of gradients. MD, on the other hand, is rather stable for changes in the number of gradients, but not that robust when the b-value changes.

Results in Fig. 3 indicate that the new measures are significantly more robust than FA and GFA. Even the worst results (the CVD with 6 gradient directions and $\sigma = 5$ for different number of gradients) show a ratio $R(p_0) = 0.09$, whereas the FA results reach almost 1 and GFA 0.38.

It is necessary to use 38 gradient directions in the FA or 18 in the GFA to achieve similar results to the CVD with only 6 gradient directions. On the other hand, DV and ASD show a similar behavior than MD for this comparison.

As shown in Fig. 4, FA presents a heavy variability with respect to the *b*-value. Although more robust, all the other measures also present some variation. In any case, the variation is similar to the one provided by the MD and, in some cases (SMD$_2$) slightly smaller for high *b*-values. It is important to point out that the CVD, since it takes into account the variance of the diffusion, is inherently greatly influenced by the *b*-value, as also happens with FA and GFA. Therefore, it is imperative to employ the unbiased version (shown in Fig. 4 together with the original versions). Note that GFA shows a similar behavior to the biased CVD. It is precisely the use of the unbiased version that improves the behavior of CVD.

Finally, in Fig. 5, FA shows a major dependency to changes in the SNR. The unbiased versions of CVD, although also encoding anisotropy information, present a superior performance. MD, on the other hand, shows a relatively low variation, which is also achieved by DV, ASD and SMD$_2$ (note that ASD and SMD$_2$ also convey some anisotropy

information). This last figure shows one important feature of the new measures: due to their formulation, the bias analysis is feasible and noise-related bias can be easily removed. Once more, the GFA shows a poor result, very similar to that of CVD but without the capability to unbiased the measure.

All in all, these three experiments show that, in general, the new measures show a more robust behavior with respect to the variation of acquisition parameters than the FA and GFA, and similar to that of MD.

We next show results for experiment *DWI phantom - discriminant power test*, where we analyze another fundamental feature of a good diffusion feature: its ability to discriminate differences in the diffusion. In this experiment, these differences have been designed to mainly reflect a change in the anisotropy of the diffusion. Figs. 6 and 7 show qualitative results of the T-test carried out over two sets, considering 50 realizations of each set. Different acquisition parameters are considered for each test: different number of gradients (Fig. 6(a)–(b)), different *b*-values (Fig. 6(c)–(d)), different SNR (Fig. 7(a)), and variation of more than one parameter (Fig. 7(b)–(d)). For each experiment, significant differences detected with $p < 0.01$ are highlighted. The top figures (labeled as TRUE) show areas in which differences in diffusion are
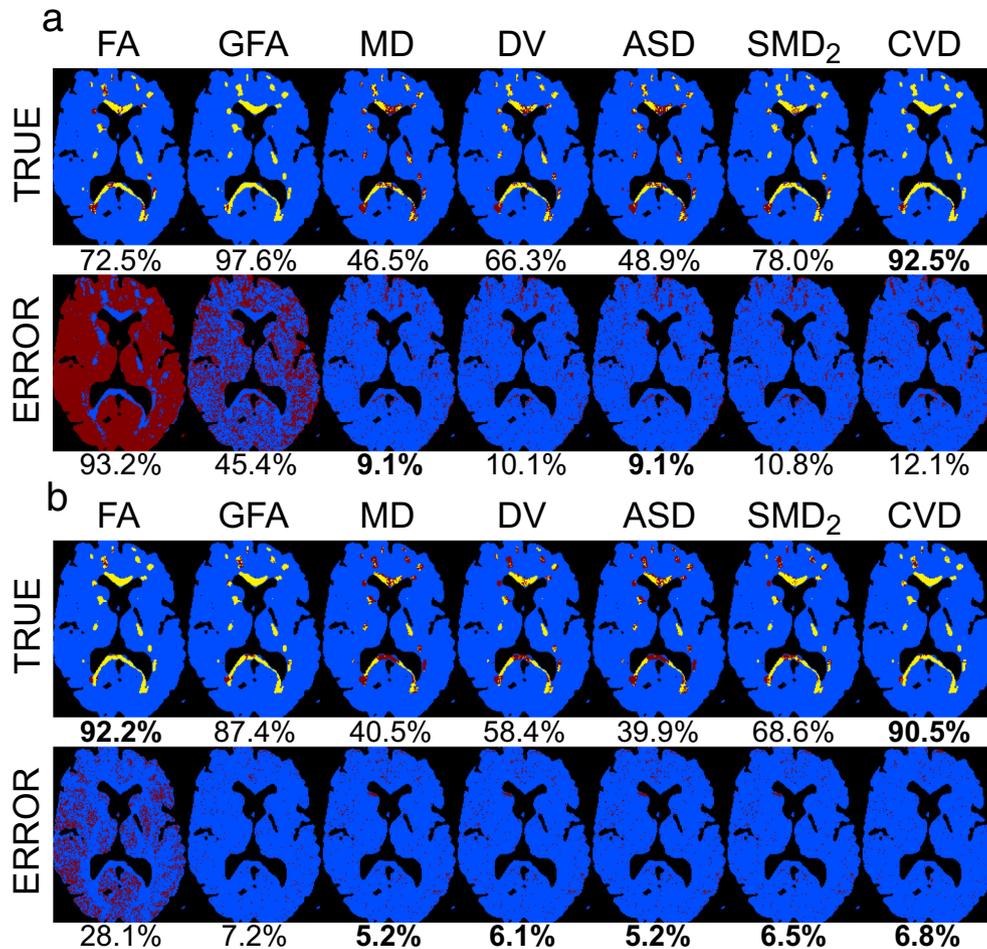
Fig. 6. Points of a single slice where the T-test (50 realizations) detects significant differences with $p < 0.01$ for various scalar measures and different configurations: (a) $N_g = \{10, 41\}$, $b = \{1000, 1000\}$ s/mm^2, $\sigma = \{4, 4\}$, (b) $N_g = \{30, 45\}$, $b = \{1000, 1000\}$ s/mm^2, $\sigma = \{4, 4\}$, (c) $N_g = \{30, 30\}$, $b = \{\mathbf{1000}, \mathbf{800}\}$ s/mm^2, $\sigma = \{3, 3\}$, and (d) $N_g = \{30, 30\}$, $b = \{\mathbf{1500}, \mathbf{800}\}$ s/mm^2, $\sigma = \{3, 3\}$. Top figures (TRUE): yellow color denotes the area of modified diffusion that is detected by the T-test and red color denotes the area of modified FA that is not detected (and it should). Bottom figures (ERROR): red color denotes the area of the non-modified white matter FA that is detected as *different* by the T-test (and it should not). The percentage written below the row labeled as TRUE illustrates the ratio of correctly detected differences, while below the row labeled as ERROR indicates the ratio of points incorrectly detected as different. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. (*continued*)

present. Yellow color denotes the area of modified diffusion that is correctly detected by the T-test, and the red color indicates the area of modified diffusion that is not detected but should have been. The percentage written below the images indicates the ratio of correctly detected differences (100% would be a perfect score). Bottom figures, labeled as ERROR, show the areas in which there are no differences, and thus the test should not detect them. Red color denotes points in which, wrongly, differences have been detected. Here, the percentage under the images indicates the ratio of points incorrectly detected as different (the lower, the best).

These results indicate the FA to be specially sensitive to changes in the number of gradient directions and in the SNR, as expected in sight of the results from the previous experiment. Even for variations of the *b*-value, its performance is not as good as that of the other measures. On the contrary, the MD can be seen as an upper bound for robustness, showing always the lowest errors. However, it is not able to properly detect the differences between groups. This is also expected, as the experiment consisted of an alteration of the anisotropy of the diffusion, for which the MD is very insensitive. The new measures, on the other hand, show a compromise between detecting the differences in anisotropy and keeping a reduced error outside the modified area. For a reduction in the number of gradients (Fig. 6(a)–(b)), CVD shows a great discrimination power (over 85%) while having a moderately small error. For a change in the *b*-value (Fig. 6(c)–(d)), they show similar

discrimination to the FA with similar error. SMD_2 shows also some robustness in this case. Moreover, when more than one parameter is changed, the FA behaves erratically, wrongly detecting differences in most of the white matter. See, for instance Fig. 7(b), where the number of gradients and the *b*-value are changed. The FA shows almost a 90% of error, while the proposed measures are in the range of the MD level while showing nearly a 90% of discriminant power. In addition, note that the GFA performs better than the FA (and worse than CVD) for different number of gradients, but it is highly affected by changes on the *b*-value and the SNR.

Finally, note that we have restricted ourselves to a minimum number of gradients of 6, to make the measures compatible with those derived from the DT. However, there is no restriction in the number of gradients for the measures in Table 1. On the other hand, according to the results in Table 2, the variance of the proposed metrics decreases with the number of gradients. A small number of directions will imply a very high variance in the measures, specially in CVD. As an illustration, the measures previously depicted in Fig. 2 for 51 gradient directions are now recalculated for 3 directions in Fig. 8. Note that the anisotropy structure of the brain can still be found with CVD using only 3 gradients. In addition, results for DV, ASD and SMD_2 are very similar to the 51 counterpart.

4.2. Real datasets

Results for experiment *real dataset 1 - robustness to number of gradients* are depicted in Fig. 9. As expected, there is a great variation in the FA values when the number of gradients changes. For this dataset, the average FA decays from 0.6 for 6 gradient directions to under 0.4 for 61 directions. On the other hand, MD, DV, ASD and SMD$_2$ show a robust behavior when changing the number of gradients. On the other hand, the CVD confirms the results in the synthetic experiments showing a small variation (0.43 to 0.45) with a smaller range than the FA and even better than the GFA.

With regard to experiment *real dataset 2 - robustness to b-value*, results are presented in Fig. 10. This figure shows the error of the measures for a central slice. Note that most of the measures show a reduced error, similar to the MD, and smaller than that of the FA. Similar to what we have seen in the synthetic experiments, the GFA shows no robustness to the changes of *b*, with the worst results. On the other hand, the CVD, which also conveys anisotropy information, presents a much more robust behavior to the change of acquisition parameters than FA and GFA.

Finally, results for experiment *real dataset 3 - discrimination power* are presented in Fig. 11. The results corroborate the FA showing little robustness to changes in the number of gradient directions, since when the HC and SZ groups are acquired with different numbers of gradient directions, differences between groups are either greatly over or underestimated. On the other hand, MD shows great robustness, as also found with the pseudo-synthetic phantom. With regard to the new measures, DV, ASD and SMD$_2$, show increased robustness with respect to FA and even with respect to MD. At the same time, the new measures show a high discriminative power, as the number of voxels with significant differences remains very similar to that of FA or MD. The CVD parameter, which incorporates anisotropy information, is less sensitive to differences than FA, but more robust with respect to changes in the number of gradient directions.

5. Discussion and conclusions

New quantitative measures of diffusion based on DWIs have been proposed in this paper and their performance, robustness and discriminant power have been evaluated on a pseudo-synthetic phantom
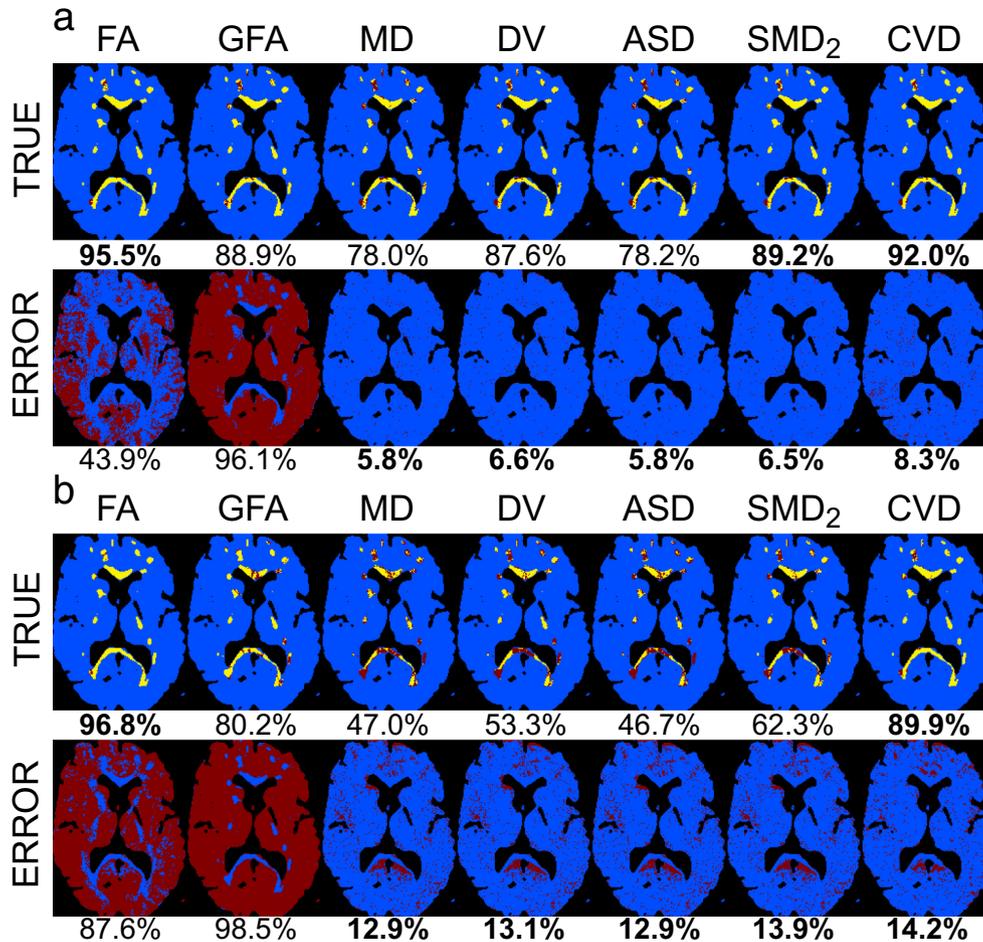
Fig. 7. Points of a single slice where the T-test (50 realizations) detects significant differences with $p < 0.01$ for various scalar measures and different configurations: (a) $N_g = \{41, 41\}$, $b = \{1200, 1200\}$ s/mm^2, $\sigma = \{3, 2\}$, (b) $N_g = \{41, 11\}$, $b = \{1200, 1000\}$ s/mm^2, $\sigma = \{3, 3\}$, (c) $N_g = \{41, 41\}$, $b = \{1000, 1200\}$ s/mm^2, $\sigma = \{3, 2\}$, and (d) $N_g = \{10, 41\}$, $b = \{800, 1200\}$ s/mm^2, $\sigma = \{4, 2\}$. Top figures (TRUE): yellow color denotes the area of modified diffusion that is detected by the T-test and red color denotes the area of modified FA that is not detected (and it should). Bottom figures (ERROR): red color denotes the area of the non-modified white matter FA that is detected as *different* by the T-test (and it should not). The percentage written below the row labeled as TRUE illustrates the ratio of correctly detected differences, while below the row labeled as ERROR indicates the ratio of points incorrectly detected as different. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. (*continued*)

and real data. The experiments have emphasized the two main features of these new indexes: **First**, they are more robust to the changes of the acquisition parameters than the FA, specifically to changes in the number of gradients, *b*-value and level of noise. **Second**, they have confirmed the ability to accurately detect changes in the characteristics of diffusion. The use of a phantom allowed us to numerically grade these positive features.

On the other hand, one of the initial requirements of these new measures was to be directly calculated over the diffusion signal. This requirement was grounded on the purpose to define measures independent of a particular model and also to allow a deep analysis of their behavior under different conditions. This idea was reflected on the proposal:

1. The new indexes are based on different sample moments of the DWIs, which avoids the middle step of a DT estimation. Thus, the variability due to the diffusion tensor estimation step is reduced and the selection of a specific diffusion model is avoided.
2. The definition of the measures allows a total characterization of their bias and variance assuming a Rician noise distribution. The noise analysis carried out in this paper has been eased due to the structure of the new measures. Accordingly, unbiased versions of the same measures have been proposed to make them robust to changes in the SNR.
3. Since they are not based on the DT model, the measures are not limited to a minimum of 6 gradient directions. Diffusion metrics can
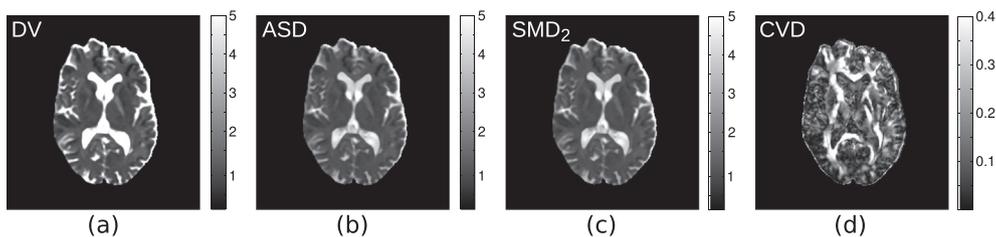
Fig. 8. Diffusion scalar measures obtained from axial slice of a realistic phantom with only 3 gradient directions: (a) $DV \times b^{3/2}$, (b) $ASD \times b$, (c) $(SMD_2)^{1/2} \times b$, and (d) CVD.
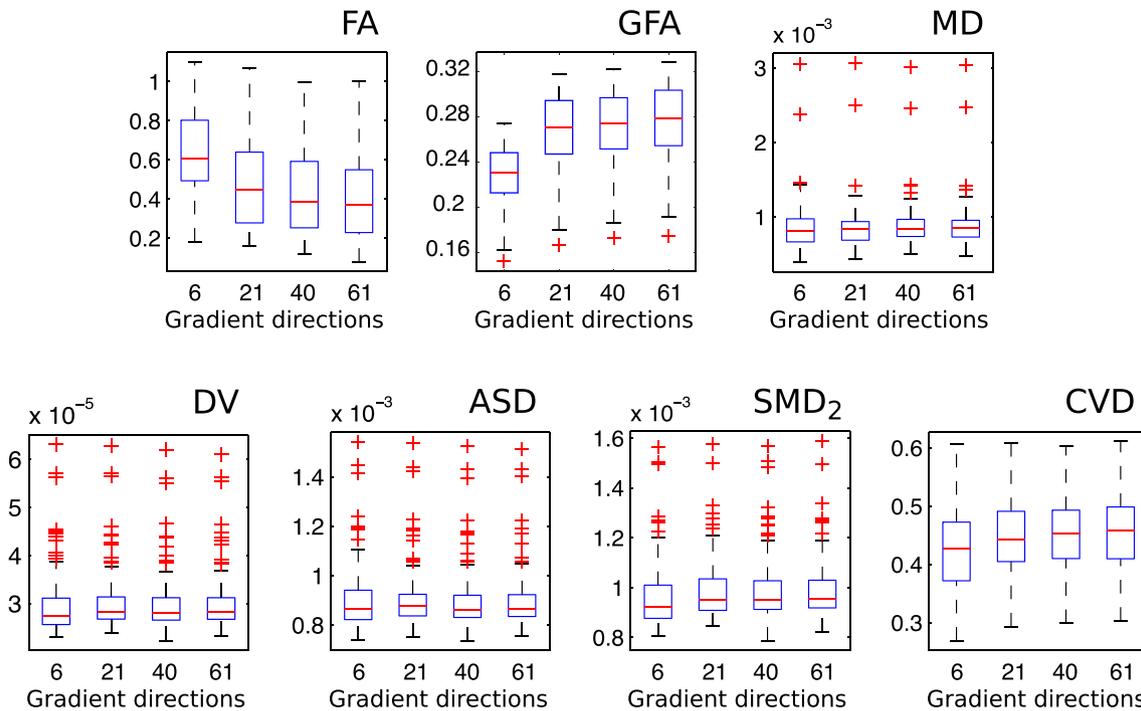
Fig. 9. Robustness test to variation of gradient directions with real data set 1: boxplot of the average values inside 5 regions (genu, body and splenium of corpus callosum; right and left Cingulum) for 13 patients for different number of gradient directions. (For the sake of visual comparison, $SMD_2^{1/2}$ and $CVD^{1/2}$ are shown, instead of the original measures).
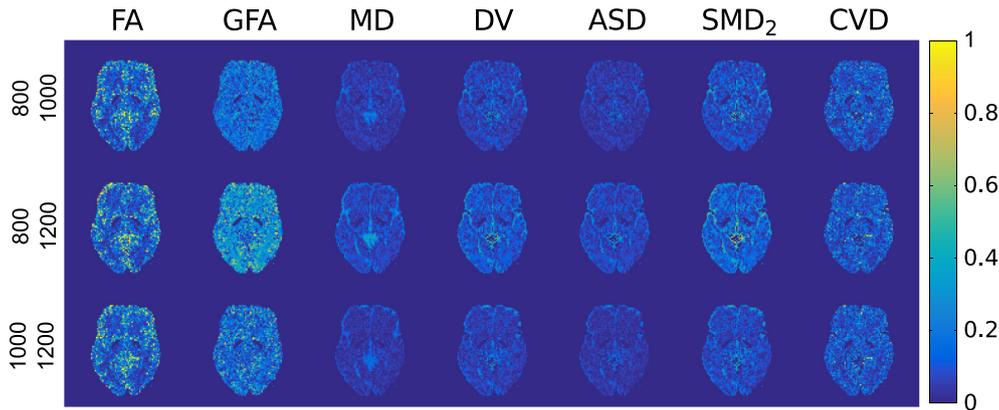
Fig. 10. Robustness test to variation of b-value with real data set 2: error between the same slice for three different b-values.

be obtained with only 3 gradient directions, although the cost to pay is to seriously increment the variance of the estimators.

The measures also present some **limitations** that must be considered:

- First, for the sake of simplicity, the measures rely on the assumption of an equally spaced sampling of one shell in the **q**-space. Although it is a reasonable assumption, if it is not fulfilled, the immediate solution would be the use of alternative calculation of the integrals via Spherical Harmonics. However, since one of the initial requirements was to keep the measures simple, we have opted for the discretization here suggested.
- The biological interpretation of the proposed measures is also an issue. These new indexes range from completely focused on the amount of diffusion (DV, ASD) to completely focused on anisotropy (CVD), while SMD_2 incorporates information from both features. On

the other hand, although many years of usage of FA and MD have provided the researchers with intuition about the nature of their findings using these measures, it is important to note that the relationship between changes in these measures and microstructural changes in the white matter is not straightforward at all. For instance, a decrease in FA can be caused by different factors, such as demyelination, lower packing density or different membrane permeability [60]. In any case, any new measure will initially suffer from this problem, and clinical studies over different pathologies will be needed to correctly interpret the measures in relation with physical processes. Previous experiments already carried out, point out that the behavior of these indexed is not far from traditional ones, and interpretation will be very similar, with the advantage of more robust results.

- It is important to note that experiment *real dataset 3 - discrimination power* was performed employing a specific tool, TBSS. This choice has some relevant implications. On the one hand, TBSS relies on a
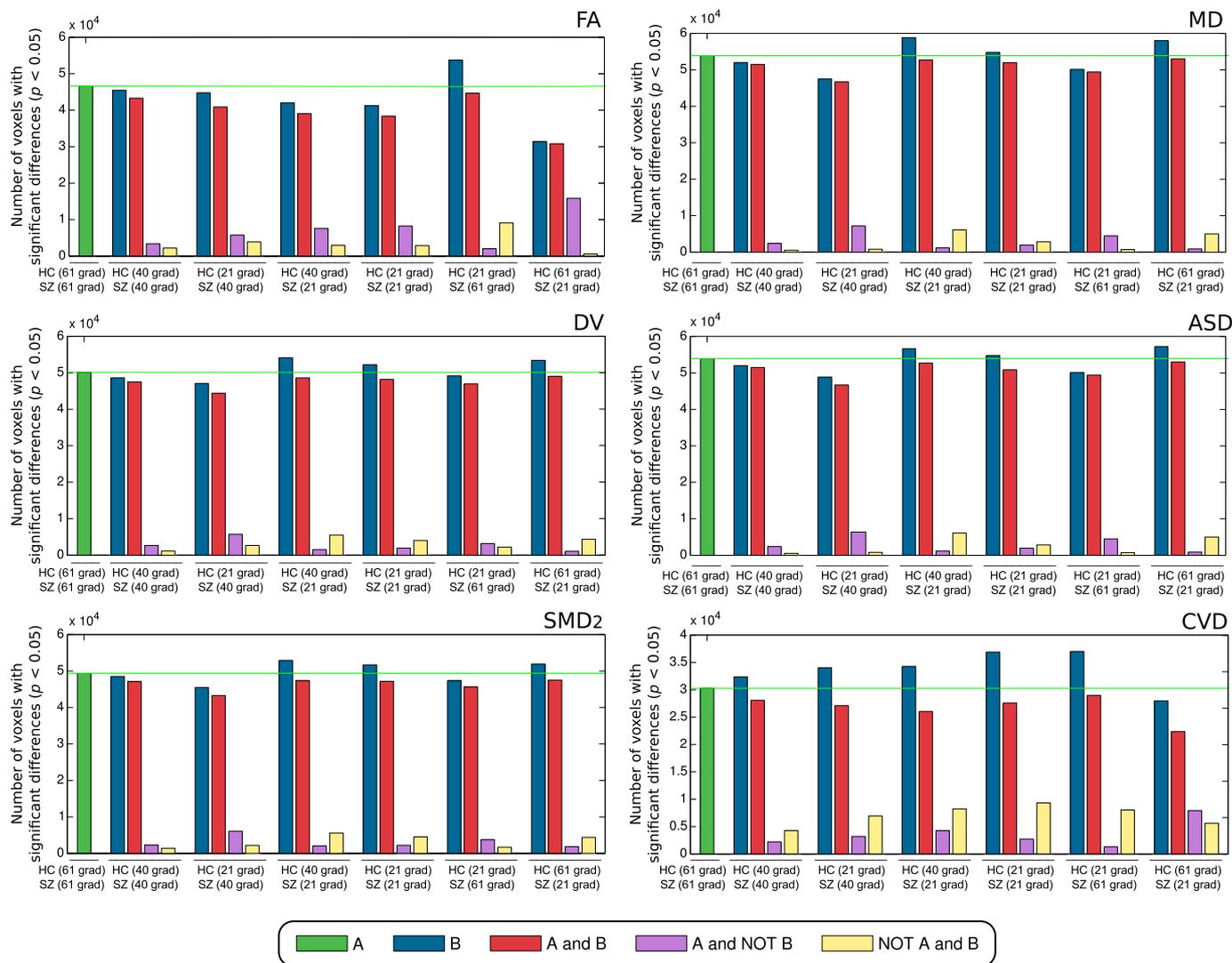
Fig. 11. Results from experiment *real dataset 3 - discrimination power and robustness to number of gradients*. Graphs show the number of voxels where significant differences were found in a TBSS group study between schizophrenia patients (SZ) and control subjects (HC), using different combinations of number directions for both groups (see description of the experiment in Section 3.2 for further details).

registration process followed by the extraction of a skeleton of the white matter from the FA volumes. If the dataset in a group study originates from acquisitions with different numbers of gradient directions, then the TBSS will need to register and obtain the skeleton from heterogeneous FA volumes. Although in principle this could negatively impact the results, we performed experiments (not shown for the sake of space) that mimic a realistic scenario where this registration process is performed on heterogeneous FA volumes. Results indicated TBSS to be robust, in this sense, to the use of FA volumes with different number of gradient directions. On the other hand, the conclusions and interpretations of the results obtained must also take into account the nature of the performed experiments. The behavior, for instance, of the proposed measures as metrics in tractography or connectomics-based analyses cannot be easily inferred, as these methods depend on a high number of factors before the application of any metric.

All in all, the new measures here proposed are a solid alternative to

standard scalar measures derived from DTI and they can alternatively be used in group studies that combine acquisitions with different acquisition parameters.

Acknowledgments

Appendix A. Derivation of the diffusion volume

Assuming a Gaussian diffusion model, we can infer that the value of $P(\mathbf{R}|\Delta)$ in the origin is related to the volume of signal $E(\mathbf{q})$ as seen in Eq. (11). To calculate the integral we use spherical coordinates and a generic diffusion $D(\mathbf{q})$ instead of the tensor model:

$$
\begin{aligned}
P(\mathbf{0}) &= \int_V E(\mathbf{q})d\mathbf{q} = \int_V \exp(-b \cdot D(\mathbf{q}))d\mathbf{q} \\
&= \int_0^{2\pi} d\theta \int_0^\pi d\phi \int_0^\infty dq_0 \exp(-c_0 \cdot q_0^2 \cdot D(\theta, \phi))q_0^2 \sin\theta
\end{aligned}
$$

with $q_0 = \|\mathbf{q}\|$ and $c_0 = 4\pi^2\tau$. We assume that the diffusion $D(\mathbf{q})$ does not depend on the radial direction q_0, and therefore $D(\mathbf{q}) = D(\theta,\phi)$:

$$
\begin{aligned}
\int_V E(\mathbf{q})d\mathbf{q} &= \int_0^{2\pi} d\theta \int_0^\pi d\phi \frac{\sqrt{\pi}}{4(c_0 \cdot D(\theta,\phi))^{3/2}} \sin\theta \\
&= \int_S \frac{\sqrt{\pi}}{4(c_0 \cdot D(\theta,\phi))^{3/2}} dS
\end{aligned}
$$

i.e., the integral in the surface of the sphere S. We approximate now the integral by the sum of the sampled values

$$
\int_V E(\mathbf{q})d\mathbf{q} \approx 2 \sum_i \frac{\sqrt{\pi}}{4c_0^{3/2}D_i^{3/2}}\Delta S, \tag{22}
$$

where ΔS is an increment of a sphere on unitary radius and D_i are the sampled diffusion values. The value 2 comes to the fact that the samples are taken only over half of the sphere. If we drop the constants, we can write

$$
\int_V E(\mathbf{q})d\mathbf{q} \propto \frac{1}{N_g} \sum_{i=1}^{N_g} \frac{b^{3/2}}{(-\log E(\mathbf{q}_i))^{3/2}}
$$

and from Eq. (11):

$$
\frac{1}{N_g} \sum_{i=1}^{N_g} \frac{b^{3/2}}{(-\log E(\mathbf{q}_i))^{3/2}} \propto \frac{1}{b^{3/2} \cdot |\mathscr{D}|^{1/2}}.
$$

Thus, after some algebra we finally have

$$
\sqrt{|\mathscr{D}|} \propto \left[b^3 \frac{1}{N_g} \sum_{i=1}^{N_g} \frac{1}{(-\log E(\mathbf{q}_i))^{3/2}} \right]^{(-1)}. \tag{23}
$$

We define now Diffusion Volume as

$$
DV = \left[b^3 \frac{1}{N_g} \sum_{i=1}^{N_g} \frac{1}{(-\log E(\mathbf{q}_i))^{3/2}} \right]^{(-1)}. \tag{24}
$$

Since the inverse of the logarithm can produce numerical instability, an approximation is needed, using the following expansion:

$$
\mathbb{E}\left\{\frac{1}{X}\right\} = \frac{1}{\mathbb{E}\{X\}} + \frac{1}{\mathbb{E}\{X\}^3}\mathrm{Var}\{X\} + \dots \tag{25}
$$

Considering the average operator as an estimator of the sample mean and using the first order approximation we can write

$$
\frac{1}{N_g} \sum_{i=1}^{N_g} \frac{1}{X_i} \approx \left(\frac{1}{N_g} \sum_{i=1}^{N_g} X_i \right)^{-1}.
$$

With this approximation, Eq. (24) becomes

$$
\begin{aligned}
DV(\mathbf{x}) &= \frac{1}{b^3} \frac{1}{N_g} \sum_{i=1}^{N_g} (-\log E(\mathbf{q}_i))^{3/2} \\
&= \frac{1}{b^3} \left\langle (-\log E(\mathbf{q}_i))^{3/2} \right\rangle. \tag{26}
\end{aligned}
$$

Appendix B. Noise analysis of the diffusion measures

B.1. Bias of ASD

We calculate the expectation of the ASD in Eq. (18):

$$
\begin{aligned}
\mathbb{E}\{ASD\} &= \mathbb{E}\left\{ -\frac{1}{b} \left\langle \log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right\rangle \right\} \\
&= -\frac{1}{b}(\langle \mathbb{E}\{\log S_i(\mathbf{x})\}\rangle - \mathbb{E}\{\log S_0(\mathbf{x})\})
\end{aligned}
$$

The term $\mathbb{E}\{\log S_i(\mathbf{x})\}$ is the mean of a logRician random variable, which is [8]

$$
\mathbb{E}\{\log S_i(\mathbf{x})\} = \log A_i(\mathbf{x}) + \frac{1}{2}\Gamma\left(0, \frac{A_i^2(\mathbf{x})}{2\sigma^2}\right)
$$

where $\Gamma(n,x)$ is the upper incomplete Gamma function that, for the particular case of $n = 0$ equals the exponential integral function $\Gamma(0, x) = E_1(x)$.

which fulfills the following property for $x > 0$:

$$E_1(x) < \exp\left(-x\right)\log\left(1 + \frac{1}{x}\right).$$

If we choose any positive $x_0 > 0$, the monotonicity of $\log\left(1 + \frac{1}{x}\right)$ assures that the bias of logRician signals is bounded by $\exp(-x)\log\left(1 + \frac{1}{x_0}\right)$ for $x > x_0$ i.e., it decreases exponentially, hence faster than any negative power of x. Thus, the bias of the measure

$$\text{bias(ASD)} = \frac{1}{2b}\left(\left\langle E_1\left(0, \frac{A_i^2(\mathbf{x})}{2\sigma^2}\right)\right\rangle - E_1\left(0, \frac{A_0^2(\mathbf{x})}{2\sigma^2}\right)\right)$$

decreases fast for high SNR. A clearer look can be obtained from a series expansion of $\log S_i(\mathbf{x})$. If we assume a high SNR we can write

$$
\begin{aligned}
\log S_i(\mathbf{x}) &= \tfrac{1}{2}\log\left(A_i^2 + 2N_r A_i + N_r^2 + N_i^2\right) \\
&= \log A_i + \tfrac{1}{2}\log\left(1 + \frac{2N_r}{A_i} + \frac{N_r^2 + N_i^2}{A_i^2}\right) \\
&= \log A_i(\mathbf{x}) + \tfrac{1}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n \cdot A_i^n}\left[(N_r - j \cdot N_i)^n + (N_r + j \cdot N_i)^n\right]
\end{aligned}
$$

From this, the expectation of $\log S_i(\mathbf{x})$ is calculated as

$$\mathbb{E}\{\log S_i(\mathbf{x})\} = \log A_i(\mathbf{x}) + \frac{1}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n \cdot A_i^n}\left[\mathbb{E}\{(N_r - j \cdot N_i)^n\} + \mathbb{E}\{(N_r + j \cdot N_i)^n\}\right].$$

Since $\mathbb{E}\{(N_r \pm j \cdot N_i)^n\} = 0$ (see Appendix C), for high SNR we can conclude that

$$\mathbb{E}\{\log S_i(\mathbf{x})\} \approx \log A_i(\mathbf{x}) \tag{27}$$

and therefore

$$\mathbb{E}\{\text{ASD}\} = -\frac{1}{b}\left\langle \log \frac{A_i(\mathbf{x})}{A_0(\mathbf{x})}\right\rangle, \tag{28}$$

i.e. the measure is unbiased.

B.2. Bias of SMD_2

We calculate the expectation of the SMD_2 in Eq. (19):

$$
\begin{aligned}
\mathbb{E}\{\text{SMD}_2\} &= \mathbb{E}\left\{\frac{1}{b^2}\left\langle\left[\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right]^2\right\rangle\right\} \\
&= -\frac{1}{b^2}\left\langle\mathbb{E}\{(\log S_i(\mathbf{x}))^2 + (\log S_0(\mathbf{x}))^2 - 2\log S_i(\mathbf{x})\log S_0(\mathbf{x})\}\right\rangle
\end{aligned}
$$

and using the approximation for high SNR in Eq. (27) and a series expansion

$$\mathbb{E}\{(\log S_i)^2\} = (\log A_i)^2 + \frac{\sigma^2}{A_i^2} - \frac{4}{3}\frac{\sigma^4}{A_i^4} + O\left(\frac{1}{A_i^6}\right).$$

we can write

$$\mathbb{E}\{\text{SMD}_2\} \approx \frac{1}{b^2}\left(\left\langle\left[\log\frac{A_i(\mathbf{x})}{A_0(\mathbf{x})}\right]^2\right\rangle + \left\langle\frac{\sigma^2}{A_i^2}\right\rangle + \frac{\sigma^2}{A_0^2}\right) \tag{29}$$

and therefore

$$\text{bias(SMD}_2) = \frac{1}{b^2}\left(\left\langle\frac{\sigma^2}{A_i^2}\right\rangle + \frac{\sigma^2}{A_0^2}\right). \tag{30}$$

B.3. Bias of DV

The expectation of the DV is

$$\mathbb{E}\{\text{DV}\} = \frac{1}{b^{3/2}}\mathbb{E}\{\langle(\log S_0(\mathbf{x}) - \log S_i(\mathbf{x}))^{3/2}\rangle\}.$$

In order to calculate the expectation, we calculate the Taylor series of $(\log S_i(\mathbf{x}) - \log S_0(\mathbf{x}))^{3/2}$ for high SNR and truncate at order two:

$$\mathbb{E}\{DV\} \approx \frac{1}{b^{3/2}} \langle (-(\log A_i(\mathbf{x}) - \log A_0(\mathbf{x})))^{3/2} \rangle + \frac{3}{8} \frac{\sigma^2}{b^{3/2}} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \left(-\log \frac{A_i}{A_0} \right)^{-1/2} \right\rangle$$

and therefore

$$\text{bias(DV)} = \frac{3}{8} \frac{\sigma^2}{b^{3/2}} \left\langle \left(\frac{1}{A_i^2} + \frac{1}{A_0^2} \right) \left(\log \frac{A_0}{A_i} \right)^{-1/2} \right\rangle. \tag{31}$$

B.4. Bias of CVD

For the sake of simplicity, we will calculate the bias of CVD^2 instead of CVD. Let us define two variables:

$$X = \mathcal{V}\left(\log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right)$$

$$Y = \left\langle \left[\log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right]^2 \right\rangle$$

so that $CVD^2 = X/Y$. The expectation of CVD^2 can be written as [61]

$$\mathbb{E}\{CVD^2\} = \mathbb{E}\left\{ \frac{X}{Y} \right\} \approx \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} - \frac{\text{Cov}(X, Y)}{\mathbb{E}\{Y\}^2} + \frac{\text{Var}\{Y\}\mathbb{E}\{X\}}{\mathbb{E}\{Y\}^3}$$

If we assume a high SNR we can make a first order simplification and then

$$\mathbb{E}\{CVD^2\} \approx \frac{\mathcal{V}\left(\log \frac{A_i(\mathbf{x})}{A_0(\mathbf{x})} \right) + \left\langle \frac{\sigma^2}{A_i^2} \right\rangle}{\left\langle \left[\log \frac{A_i(\mathbf{x})}{A_0(\mathbf{x})} \right]^2 \right\rangle + \left\langle \frac{\sigma^2}{A_i^2} \right\rangle + \frac{\sigma^2}{A_0^2}} \tag{32}$$

Note that the bias here is not additive, as in the previous cases.

B.5. Variance of ASD

For the variance of the sum of N random variables X_i the following relation holds

$$\text{Var}\left\{ \sum_{i=1}^{N} X_i \right\} = \sum_{i=1}^{N} \text{Var}\{X_i\} + \sum_{i \neq j} \text{Cov}(X_i, X_j).$$

Simplifying our considerations, we can define the variance of ASD as follows:

$$\text{Var}\{ASD\} = \text{Var}\left\{ -\frac{1}{b} \frac{1}{N_g} \sum_{i=1}^{N_g} \log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right\} \approx \frac{1}{b^2 N_g^2} \sum_{i=1}^{N_g} \text{Var}\left\{ \log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right\}$$

$$= \frac{1}{b^2} \frac{1}{N_g^2} \sum_{i=1}^{N_g} (\mathbb{E}\{(\log S_i(\mathbf{x}) - \log S_0(\mathbf{x}))^2\} - (\mathbb{E}\{\log S_i(\mathbf{x}) - \log S_0(\mathbf{x})\})^2).$$

Given the approximations of expectations for high SNRs, we conclude that

$$\text{Var}\{ASD\} \approx \frac{1}{b^2} \frac{\sigma^2}{N_g} \left(\left\langle \frac{1}{A_i^2} \right\rangle + \frac{1}{A_0^2} \right). \tag{33}$$

B.6. Variance of SMD_2

Following a similar reasoning we can write

$$\text{Var}\{SMD_2\} = \text{Var}\left\{ \frac{1}{b^2} \frac{1}{N_g} \sum_{i=1}^{N_g} \left(\log \frac{S_i(\mathbf{x})}{S_0(\mathbf{x})} \right)^2 \right\}$$

$$\approx \frac{1}{b^4} \frac{1}{N_g^2} \sum_{i=1}^{N_g} (\mathbb{E}\{(\log S_i(\mathbf{x}) - \log S_0(\mathbf{x}))^4\} - (\mathbb{E}\{(\log S_i(\mathbf{x}) - \log S_0(\mathbf{x}))^2\})^2).$$

To resolve this, let us define series expansion of the expectations $\mathbb{E}\{(\log S_i)^3\}$ and $\mathbb{E}\{(\log S_i)^4\}$:

$$\mathbb{E}\{(\log S_i)^3\} = (\log A_i)^3 + \frac{3\sigma^2}{A_i^2} \log A_i + \frac{3\sigma^3}{A_i^3} + O\left(\frac{1}{A_i^4} \right),$$

$$\mathbb{E}\{(\log S_i)^4\} = (\log A_i)^4 + \frac{6\sigma^2}{A_i^2} (\log A_i)^2 + \frac{4\sigma^3}{A_i^3} \log A_i + \frac{\sigma^4}{A_i^4} + O\left(\frac{1}{A_i^5} \right)$$

and from this, we obtain the variance of SMD_2:

$$\text{Var}\{SMD_2\} \approx \frac{1}{b^4}\frac{4\sigma^2}{N_g}\left\langle\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)\left(\log\frac{A_i}{A_0}\right)^2\right\rangle - \frac{1}{b^4}\frac{\sigma^4}{N_g}\left\langle\frac{1}{A_i^2 A_0^2}\right\rangle.$$

(34)

B.7. Variance of DV

Once again we will employ a similar strategy as previously presented:

$$\text{Var}\{DV\} = \text{Var}\left\{\frac{1}{b^{3/2}}\frac{1}{N_g}\sum_{i=1}^{N_g}\left(-\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)^{3/2}\right\}$$

$$\approx \frac{1}{b^3}\frac{1}{N_g^2}\sum_{i=1}^{N_g}\left(\mathbb{E}\{(\log S_0(\mathbf{x})-\log S_i(\mathbf{x}))^3\}-(\mathbb{E}\{(\log S_0(\mathbf{x})-\log S_i(\mathbf{x}))^{3/2}\})^2\right).$$

We define the expectation $\mathbb{E}\{(\log S_0-\log S_i)^3\}$ to be

$$\mathbb{E}\{(\log S_0-\log S_i)^3\} = (\log A_0-\log A_i)^3 + 3\sigma^2\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)\log\frac{A_0}{A_i} + O\left(\frac{1}{A_i^3}\right)$$

and then

$$\text{Var}\{DV\} \approx \frac{1}{b^3}\frac{9\sigma^2}{4N_g}\left\langle\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)\log\frac{A_0}{A_i}\right\rangle + \frac{1}{b^3}\frac{9\sigma^4}{64N_g}\left\langle\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)^2\left(\log\frac{A_0}{A_i}\right)^{-1}\right\rangle.$$

(35)

B.8. Variance of CVD

For the sake of simplicity, we will again follow the assumptions made for calculation the bias of CVD measure. To calculate the variance of the ratio X/Y the following approximation can be used:

$$\text{Var}\left\{\frac{X}{Y}\right\} \approx \frac{\text{Var}\{X\}}{\mathbb{E}\{Y\}^2} - \frac{2\,\mathbb{E}\{X\}}{\mathbb{E}\{Y\}^3}\text{Cov}(X, Y) + \frac{\mathbb{E}\{X\}^2}{\mathbb{E}\{Y\}^4}\text{Var}\{Y\}.$$

Assuming a high SNR, we make a first order simplification leading to

$$\text{Var}\{CVD^2\} \approx \frac{\text{Var}\left\{\mathcal{V}\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)\right\}}{\mathbb{E}\left\{\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)^2\right\}^2} = \frac{\text{Var}\left\{\mathcal{V}\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)\right\}}{\frac{1}{b^4}\left(\left\langle\left(\log\frac{A_i}{A_0}\right)^2\right\rangle + \left\langle\frac{\sigma^2}{A_i^2}\right\rangle + \frac{\sigma^2}{A_0^2}\right)^2}.$$

We need to calculate the variance in the numerator:

$$\text{Var}\left\{\mathcal{V}\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)\right\} = \text{Var}\left\{\frac{1}{b^2}\frac{N_g}{N_g-1}\left[\frac{1}{N_g}\sum_{i=1}^{N_g}\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)^2 - \left(\frac{1}{N_g}\sum_{i=1}^{N_g}\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)^2\right]\right\}$$

$$\approx \frac{1}{b^4}\frac{1}{(N_g-1)^2}\text{Var}\left\{\sum_{i=1}^{N_g}(\log S_i(\mathbf{x})-\log S_0(\mathbf{x}))^2\right\}$$

$$+ \frac{1}{b^4}\frac{1}{(N_g-1)^2 N_g^2}\text{Var}\left\{\left(\sum_{i=1}^{N_g}(\log S_i(\mathbf{x})-\log S_0(\mathbf{x}))\right)^2\right\}.$$

The first part of the equation has been already derived as a part of the variance for SMD_2 measure. In order to calculate the variance of the second part, we will use the following approximation:

$$\text{Var}\{X^2\} \approx 4\,\mathbb{E}\{X\}^2\text{Var}\{X\}.$$

From this, we obtain the variance of the measure

$$\text{Var}\left\{\mathcal{V}\left(\log\frac{S_i(\mathbf{x})}{S_0(\mathbf{x})}\right)\right\} \approx \frac{1}{b^4}\frac{4N_g\sigma^2}{(N_g-1)^2}\left\langle\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)\left(\log\frac{A_i}{A_0}\right)^2\right\rangle$$

$$- \frac{1}{b^4}\frac{N_g\sigma^4}{(N_g-1)^2}\left\langle\frac{1}{A_i^2 A_0^2}\right\rangle + \frac{1}{b^4}\frac{4N_g\sigma^2}{(N_g-1)^2}\left\langle\log\frac{A_i}{A_0}\right\rangle^2\left(\left\langle\frac{1}{A_i^2}\right\rangle + \frac{1}{A_0^2}\right).$$

and therefore

$$\text{Var}\{CVD^2\} \approx \frac{4\sigma^2}{N_g}\frac{\left\langle\left(\frac{1}{A_i^2}+\frac{1}{A_0^2}\right)\left(\log\frac{A_i}{A_0}\right)^2\right\rangle - \frac{\sigma^2}{4}\left\langle\frac{1}{A_i^2 A_0^2}\right\rangle + 4\left\langle\log\frac{A_i}{A_0}\right\rangle^2\left(\left\langle\frac{1}{A_i^2}\right\rangle + \frac{1}{A_0^2}\right)}{\left(\left\langle\left(\log\frac{A_i}{A_0}\right)^2\right\rangle + \left\langle\frac{\sigma^2}{A_i^2}\right\rangle + \frac{\sigma^2}{A_0^2}\right)^2}.$$

(36)

Appendix C. Calculation of expectations

Let us assume that $N_1(\mathbf{x};0,\sigma^2)$ and $N_2(\mathbf{x};0,\sigma^2)$ are two IID Gaussian random processes with zero mean and variance σ^2. Thus

$$\mathbb{E}\{(N_1 \pm j \cdot N_2)^n\} = 0. \tag{1}$$

Let us see the demonstration for the " $+$ " case, being the one for " $-$ " similar.

$$
\begin{aligned}
\mathbb{E}\{(N_1 + j \cdot N_2)^n\} &= \mathbb{E}\left\{\sum_{k=0}^{n}\binom{n}{k}N_1^{n-k}(j \cdot N_2)^k\right\} \\
&= \sum_{k=0}^{n}\binom{n}{k}\mathbb{E}\{N_1^{n-k}\}j^k \cdot \mathbb{E}\{N_2^k\}
\end{aligned}
$$

The moments of a Gaussian RV are known to be

$$
\mathbb{E}\{N_i^n\} = \begin{cases} 0 & n \quad \text{odd} \\ 2^{m/2}\dfrac{m!}{(m/2)!} & n \quad \text{even} \end{cases}
$$

so we can rewrite

$$
\begin{aligned}
\mathbb{E}\{(N_1 + j \cdot N_2)^n\} &= \sum_{k=0}^{n/2}\binom{n}{2k}\mathbb{E}\{N_1^{n-2k}\}j^{2k} \cdot \mathbb{E}\{N_2^{2k}\} \\
&= \frac{2^{-n/2}n!}{(n/2)!}\sum_{k=0}^{n/2}\binom{n/2}{k}(-1)^k.
\end{aligned}
$$

Since the summation

$$\sum_{k=0}^{M}\binom{M}{k}(-1)^k = \delta[M]$$

with $\delta[M]$ the Kronecker delta, i.e. the solution is only different from zero only if $M = 0$. Then, we can conclude that

$$\mathbb{E}\{(N_1 + j \cdot N_2)^n\} = 0 \quad \forall\, n > 0.$$

References

[1] Basser P, Pierpaoli C. Microstructural features measured using diffusion tensor imaging. J Magn Reson B 1996;111(3):209–19.

[2] Kubicki M, Park H, Westin C-F, et al. DTI and MTR abnormalities in schizophrenia: analysis of white matter integrity. Neuroimage 2005;26(4):1109–18.

[3] Sundgren P, Dong Q, Gomez-Hassan D, Mukherji S, Maly P, Welsh R. Diffusion tensor imaging of the brain: review of clinical applications. Neuroradiology 2004;46(5):339–50.

[4] Rovaris M, Filippi M. Diffusion tensor MRI in multiple sclerosis. J Neuroimaging 2007;17(s1):27S–30S.

[5] Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. Radiology 2011;258(3):906–14.

[6] Smith JJ, Sorensen AG, Thrall JH. Biomarkers in imaging: realizing radiology's future. Radiology 2003;227(3):633–8.

[7] Sorensen AG. Magnetic resonance as a cancer imaging biomarker. J Clin Oncol 2006;24(20):3274–81.

[8] Tristán-Vega A, Aja-Fernández S, Westin C-F. Least squares for diffusion tensor estimation revisited: propagation of uncertainty with Rician and non-Rician signals. Neuroimage 2012;59(4):4032–43.

[9] Aja-Fernández S, Niethammer M, Kubicki M, Shenton ME, Westin C-F. Restoration of DWI data using a Rician LMMSE estimator. IEEE Trans Med Imag 2008;27(10):1389–403.

[10] Tensaouti F, Lahlou I, Clarisse P, Lotterie JA, Berry I. Quantitative and reproducibility study of four tractography algorithms used in clinical routine. J Magn Reson Imaging 2011;34(1):165–72.

[11] Vaessen M, Hofman P, Tijssen H, Aldenkamp A, Jansen JF, Backes WH. The effect and reproducibility of different clinical DTI gradient sets on small world brain connectivity measures. Neuroimage 2010;51(3):1106–16.

[12] Danielian LE, Iwata NK, Thomasson DM, Floeter MK. Reliability of fiber tracking measurements in diffusion tensor imaging for longitudinal study. Neuroimage 2010;49(2):1572–80.

[13] Wakana S, Caprihan A, Panzenboeck MM, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 2007;36(3):630–44.

[14] Farrell J, Landman B, Jones C, Smith S, Prince J, van Zijl P, et al. Effects of SNR on the accuracy and reproducibility of DTI-derived fractional anisotropy, mean diffusivity and principal eigenvector measurements at 1.5 T. J Magn Reson Imag 2007;26:756–67.

[15] Landman B, Farrell J, Jones C. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. Neuroimage 2007;36:1123–38.

[16] Wang J, Abdi H, Bakhadirov K, Diaz-Arrastia R, Devous M. A comprehensive reliability assessment of quantitative diffusion tensor tractography. Neuroimage 2012;60(2):1127–38.

[17] Hui E, Cheung M, Chan K, Wu E. B-value dependence of DTI quantitation and sensitivity in detecting neural tissue changes. Neuroimage 2010;49:2366–74.

[18] Metzler-Baddeley C, O'Sullivan M, Bells S, Pasternak O, Jones D. How and how not to correct for CSF-contamination in diffusion MRI. Neuroimage 2012;59:1394–403.

[19] Jones D, Basser P. Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. Magn Reson Med 2004;52:979–93.

[20] Barrio-Arranz G, de Luis-García R, Tristán-Vega A, Martín-Fernández M, Aja-Fernández S. Impact of MR acquisition parameters on DTI scalar indexes: a tractography based approach. PloS one 2015;10(10):e0137905.

[21] Bisdas S, Bohning D, Besenski N, Nicholas J, Rumboldt Z. Reproducibility, interrater agreement, and age-related changes of fractional anisotropy measures at 3T in healthy subjects: effect of the applied b-value. Am J Neuroradiol 2008;29(6):1128–33.

[22] DeSantis S. Water diffusion in soft matter by means of NMR: from conventional to anomalous behaviour [Ph.D. thesis]. Universitá di Roma; 2010.

[23] Fujiwara S, Sasaki M, Kanbara Y, Inoue T, Hirooka R, Ogawa A. Feasibility of 1.6-mm isotropic voxel diffusion tensor tractography in depicting limbic fibers. Neuroradiology 2008;50:131–6.

[24] Santarelli X, Garbin G, Ukmar M, Longo R. Dependence of the fractional anisotropy in cervical spine from the number of diffusion gradients, repeated acquisition and voxel size. J Magn Reson Imag 2010;28:70–6.

[25] Jones D. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. Neuroimage 2004;51:807–15.

[26] Gao W, Zhu H, Lin W. A unified optimization approach for diffusion tensor imaging technique. Neuroimage 2009;44(3):729–41.

[27] Rogers BP, Blaber J, Welch EB, Ding Z, Anderson AW, Landman BA. Stability of gradient field corrections for quantitative diffusion MRI. SPIE medical imaging. 2017. p. 81–90.

[28] Hutchinson EB, Avram AV, Irfanoglu MO, Koay CG, Barnett AS, Komlosh ME, et al. Analysis of the effects of noise, DWI sampling, and value of assumed parameters in diffusion MRI models. Magn Reson Med 2017;78(5):1767–80.

[29] Vollmar C, O'Muircheartaigh J, Barker GJ, Symms MR, Thompson P, Kumari V, et al. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. Neuroimage 2010;51(4):1384–94.

[30] Zhu T, Hu R, Qiu X, Taylor M, Tso Y, Yiannoutsos C, et al. Quantification of accuracy and precision of multi-center DTI measurements: a diffusion phantom and

human brain study. Neuroimage 2011;56(3):1398–411.

[31] Zhu T, Liu X, Connelly PR, Zhong J. An optimized wild bootstrap method for evaluation of measurement uncertainties of DTI-derived parameters in human brain. Neuroimage 2008;40(3):1144–56.

[32] Venkatraman VK, Gonzalez CE, Landman B, Goh J, Reiter DA, An Y, et al. Region of interest correction factors improve reliability of diffusion imaging measures within and across scanners and field strengths. Neuroimage 2015;119:406–16.

[33] Pagani E, Hirsch JG, Pouwels PJ, et al. Intercenter differences in diffusion tensor MRI acquisition. J Magn Reson Imaging 2010;31(6):1458–68.

[34] Pohl KM, Sullivan EV, Rohlfing T, et al. Harmonizing DTI measurements across scanners to examine the development of white matter microstructure in 803 adolescents of the NCANDA study. Neuroimage 2016;130:194–213.

[35] Mirzaalian H, Ning L, Savadjiev P, et al. Inter-site and inter-scanner diffusion MRI data harmonization. Neuroimage 2016;135:311–23.

[36] Mirzaalian H, Ning L, Savadjiev P, et al. Multi-site harmonization of diffusion MRI data in a registration framework. Brain Imaging Behav 2017;12(1):284–95.

[37] Tuch DS, Reese TG, Wiegell MR, Wedeen VJ. Diffusion MRI of complex neural architecture. Neuron 2003;40:885–95.

[38] Kreher B, Schneider J, Mader I, Martin E, Henning J, Il'yasov K. Multitensor approach for analysis and tracking of complex fiber configurations. Magn Reson Med 2005;54(5):1216–25.

[39] Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Apparent Diffusion Profile estimation from high angular resolution diffusion images: estimation and applications. Magn Reson Med 2006;56(2):395–410.

[40] Tristán-Vega A, Westin C-F, Aja-Fernández S. Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. Neuroimage 2009;47(2):638–50.

[41] Özarslan E, Vemuri BC, Mareci TH. Generalized scalar measures for diffusion MRI using trace, variance, and entropy. Magn Reson Med 2005;53(4):866–76.

[42] Tuch DS. Q-ball imaging. Magn Reson Med 2004;52:1358–72.

[43] Özarslan E, Koay CG, Shepherd TM, Komlosh ME, İrfanoğlu MO, Pierpaoli C, et al. Mean apparent propagator (MAP) MRI: a novel diffusion imaging method for mapping tissue microstructure. Neuroimage 2013;78:16–32.

[44] Wu Y-C, Field AS, Alexander AL. Computation of diffusion function measures in q-space using magnetic resonance hybrid diffusion imaging. IEEE Trans Med Imag 2008;27(6):858–65.

[45] Stejskal E-O, Tanner J-E. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J Chem Phys 1965;42:288–92.

[46] Özarslan E, Sepherd TM, Vemuri BC, Blackband SJ, Mareci TH. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). Neuroimage 2006;31:1086–103.

[47] Assaf Y, Mayk A, Cohen Y. Displacement imaging of spinal cord using q-space diffusion-weighted MRI. Magn Reson Med 2000;44(5):713–22.

[48] Wu Y-C, Alexander AL. Hybrid diffusion imaging. Neuroimage 2007;36(3):617–29.

[49] Ning L, Westin C-F, Rathi Y. Estimating diffusion propagator and its moments using directional radial basis functions. IEEE Trans Med Imag 2015;34(10):2058–78.

[50] Hosseinbor AP, Chung MK, Wu YC, Fleming JO, Field AS, Alexander AL. Extracting quantitative measures from EAP: a small clinical study using BFOR. Med Image Comput Comput Assist Interv. Vol. 7511. Springer; 2012. p. 280–7.

[51] Tristán-Vega A. A novel framework for the study of neural architectures in the human brain with diffusion MRI [Ph.D. thesis]. Valladolid, Spain: Universidad de Valladolid; 2009 https://www.lpi.tel.uva.es/thesis.

[52] Aja-Fernández S, Vegas-Sánchez-Ferrero G. Statistical analysis of noise in MRI. Springer; 2016.

[53] Tristán-Vega A, Aja-Fernández S. Design and construction of a realistic DWI phantom for filtering performance assessment. Med Image Comput Comput Assist Interv. vol. 5761. 2009. p. 951–8.

[54] Tristán-Vega A, Aja-Fernández S. DWI filtering using joint information for DTI and HARDI. Med Imag Anal 2010;14(2):205–18.

[55] Hansen B, Jespersen SN. Data for evaluation of fast kurtosis strategies, *b*-value optimization and exploration of diffusion MRI contrast. Sci Data 2016;3:160072.

[56] Molina V, Lubeiro A, Soto O, Rodriguez M, Álvarez A, Hernández R, et al. Alterations in prefrontal connectivity in schizophrenia assessed using diffusion magnetic resonance imaging. Prog Neuropsychopharmacol Biol Psychiatry 2017;76:107–15.

[57] Lee S-H, Kubicki M, Asami T, Seidman LJ, Goldstein JM, Mesholam-Gately RI, et al. Extensive white matter abnormalities in patients with first-episode schizophrenia: a diffusion tensor imaging (DTI) study. Schizophr Res 2013;143(2):231–8.

[58] Smith S, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 2006;31:1487–505.

[59] Smith S, Jenkinson M, Woolrich M, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23(S1):208–19.

[60] Kochunov P, Thompson P, Lancaster J, Bartzokis G, Smith S, Coyle T, et al. Relationship between white matter fractional anisotropy and other indices of cerebral health in normal aging: tract-based spatial statistics study of aging. Neuroimage 2007;35(2):478–87.

[61] Papoulis A. Probability, random variables, and stochastic processes. 3rd ed. McGraw-Hill; 1991.