Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 164 (2003) 1–9
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

A measure of curve fitting error for noise filtering diffusion tensor MRI data

Nikos G. Papadakis,[a,*] Kay M. Martin,[b] Iain D. Wilkinson,[c] and Chris L.-H. Huang[d]

[a] *Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK*
[b] *New Hall, University of Cambridge, Cambridge, UK*
[c] *Unit of Academic Radiology, University of Sheffield, Sheffield, UK*
[d] *Department of Physiology, University of Cambridge, Cambridge, UK*

Received 3 September 2002; revised 24 February 2003

Abstract

A parameter, χ^2_p, based on the fitting error was introduced as a measure of reliability of DT-MRI data, and its properties were investigated in simulations and human brain data. Its comparison with the classic χ^2 revealed its sensitivity to both the goodness of fit and the pixel signal-to-noise-ratio (SNR), unlike the classic χ^2, which is sensitive only to the goodness of fit. The new parameter was thus able to separate effectively pixels with coherent signals (having small fitting error and/or high SNR) from those with random signals (having inconsistent fitting and/or low SNR). A practical advantage of χ^2_p over the classic χ^2 was that χ^2_p is quantified directly from the data of each pixel, without requiring accurate estimation of data-dependent parameters (such as noise variance), which often makes application of the classic χ^2 problematic. Analytical approximations of χ^2_p enabled an objective (data-independent) and automated calculation of a threshold value, used for internal scaling of the χ^2_p map. Apart from assessing data reliability on a pixel-by-pixel basis, χ^2_p was used to develop an objective and generic methodology for the exclusion of pixels with unreliable DT information by discarding pixels with χ^2_p values exceeding the threshold. Pixels corresponding to very low SNR, and poorly fitted cerebrospinal fluid and surrounding brain tissue, had increased χ^2_p values and were successfully excluded, providing DT anisotropy maps free from artifactual anisotropic appearance.

Keywords: Diffusion tensor imaging; Fitting error; Noise; Filtering; Anisotropy

1. Introduction

Diffusion tensor magnetic resonance imaging (DT-MRI) has emerged as an efficient neuroimaging modality for the description of water self-diffusion properties through its provision of a variety of diffusion tensor (DT) parameters, such as the trace, eigenvalues/eigenvectors of the DT, and anisotropy indices [1]. The pixel values of these parameters are sensitive to noise and their quantitative reliability is affected by several spatially varying factors [2,3]. For example, the signal-to-noise ratio (SNR) of the baseline (non-diffusion-weighted) DT-MRI data varies strongly across the brain being of high value in the cortex and of low value in deep gray matter. Furthermore, high SNR parenchymal regions close to cerebrospinal fluid (CSF) may be affected adversely by CSF partial volume averaging (PVA) and CSF flow and pulsation [4,5]. It is therefore often necessary to exclude pixels containing unreliable DT information. Most commonly, such pixels are excluded on the basis of the magnitude of their baseline data [5]. However, such data pixel exclusion criteria are biased towards a specific source of corrupted DT information (namely, low baseline magnitude), without providing a generic measure of the quality of the DT information in any particular pixel. Furthermore, the applied thresholds are empirical, often vary between different studies and rely on the calculation of data-dependent parameters; for example, magnitude

* Corresponding author. Fax: +44-114-2766515.
E-mail address: n.papadakis@shef.ac.uk (N.G. Papadakis).

1090-7807/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1090-7807(03)00202-7

thresholding requires data noise estimation, using the intensity of the image background [6].

Since DT-MRI data are first fitted to the DT equations, the fitting error itself may offer a sensitive indicator of the DT-MRI data quality on a pixel-by-pixel basis. Recently [7], a novel parameter based on the fitting error was briefly introduced and used as part of an empirical (data-dependent) pixel exclusion procedure. This study characterises that parameter more fully and thereby develops an objective (data-independent) methodology for assessing the quantitative reliability of the DT-MRI data.

2. Methods

2.1. Definition of fitting error measures

The N measured DT-MRI signals S_{mi} ($i = 1, 2, \ldots, N$) at a given pixel are fitted to model signal equations of multivariate monoexponential decay [1]; consequently, the baseline signal S_{f0} and the diffusion tensor $\mathbf{D_f}$ are estimated. As a result, for every S_{mi}, a fitted S_{fi} is calculated as

$$S_{fi} = S_{f0} \exp(-\mathbf{B}_i : \mathbf{D_f}), \tag{1}$$

where $\mathbf{B}_i$ is the b-matrix corresponding to S_{mi} and : denotes matrix scalar product [1]. Therefore, the squared fitting error ΔS^2 for N fitted DT-MRI signals at a given pixel is

$$\Delta S^2 = \sum_{i=1}^{N} (S_{mi} - S_{fi})^2. \tag{2}$$

The classic χ^2 measure, χ^2_c, is a normalised version of ΔS^2, where each term in ΔS^2 is divided by the noise variance σ^2 of the respective measured signal; σ is the standard deviation (SD) [8]. Since all measured signals in any image pixel are assumed to have the same σ, the same normalisation applies to ΔS^2 of all image pixels

$$\chi^2_c = \Delta S^2 / \sigma^2. \tag{3}$$

The proposed parameter χ^2_p introduced in [7] is based on ΔS^2, but uses a pixel-dependent normalisation factor; this is the total energy of the measured signals at each pixel:

$$\chi^2_p = \frac{\Delta S^2}{\sum_{i=1}^{N} S^2_{mi}}. \tag{4}$$

2.2. Simulations

Simulations of DT-MRI signals with added noise (in quadrature) for various anisotropy configurations and SNR values of the baseline signal were performed as described in [3]. Three DT sampling schemes were considered, all with $N = 78$. Scheme A consisted of two sets of isotropically arranged DW gradient directions: six directions at b-value $b_l = 100\ \mathrm{s/mm^2}$ and 72 directions at b-value $b_h = 1600\ \mathrm{s/mm^2}$ [3]. Scheme B used 13 isotropically arranged DW gradient directions at six equidistant b-values: 100, 400, 700, 1000, 1300, and 1600 s/mm^2. Scheme C was identical to A with the exception $b_h = 1200\ \mathrm{s/mm^2}$.

2.3. Imaging experiments

DT-MRI was performed on normal volunteers using a 1.5 T clinical MRI scanner (Eclipse, Philips Medical Systems, Cleveland, OH) equipped with an actively shielded whole body gradient set (maximum strength per axis of 27 mT/m, and slew rate of 72 mT/m/s). Standard (spin-echo) and CSF-suppressed (FLAIR-prepared) DT-MRI was performed on 13 normal volunteers (8 males and 5 females, age range 25–42 years), using Scheme A. At the same imaging session, standard DT-MRI was performed on three of the above volunteers (males, age range 28–42 years), using Scheme B. Scheme C was implemented using standard DT-MRI on nine normal volunteers (5 males, 4 females, age range 24–46 years), different from those above. Details of the experimental protocols are given in [7].

3. Results

Fig. 1 shows simulated χ^2_c and χ^2_p distributions for isotropy ($\lambda = 0.7 \times 10^{-3}\ \mathrm{mm^2/s}$) and for various SNR values of the baseline signal SNR_0, using DT scheme A. For any $\mathrm{SNR}_0 \gg 0$, χ^2_c follows the same distribution, namely the χ^2 distribution with $N - 7$ ($= 71$) degrees of freedom (df). For $\mathrm{SNR}_0 = 0$, χ^2_c is shifted towards smaller values because, for this case, it is $\chi^2_c/0.655^2$ (rather than χ^2_c) which follows the χ^2 distribution (Eq. (A.1)(a), Appendix A). While large χ^2_c values reflect unreliable DT fitting, small χ^2_c values do not necessarily correspond to useful DT information, since they may arise from random signals ($\mathrm{SNR}_0 \to 0$). On the other hand, χ^2_p is sensitive to SNR_0, being a decreasing function of SNR_0. Importantly, it gives well separated distributions between coherent ($\mathrm{SNR}_0 \neq 0$) and random ($\mathrm{SNR}_0 = 0$) signals. This property holds independently of the number N of fitted signals, as shown in Fig. 2, which plots a lower limit $\lambda_{\min}$ of the ratio $\chi^2_p(\mathrm{SNR}_0 = 0)/\chi^2_p(\mathrm{SNR}_0 \gg 0)$ (Eq. (A.7), Appendix A) as a function of N, for $\mathrm{SNR}_0 = 15$.

Fig. 3 assesses the sensitivity of χ^2_p on fibre shape and DT scheme. The dependence of DW signal magnitude (and thus of the normalisation factor for ΔS^2 in Eq. (4)) on fibre shape and DT scheme shifts the χ^2_p distributions. For any DT scheme, the χ^2_p distribution of the isotropic case is shifted to the right of the χ^2_p distribution of the most anisotropic fibre ($\lambda_1 = 1.9, \lambda_2 = \lambda_3 = 0.1 \times 10^{-3}\ \mathrm{mm^2/s}$). However, the shift between the peaks of

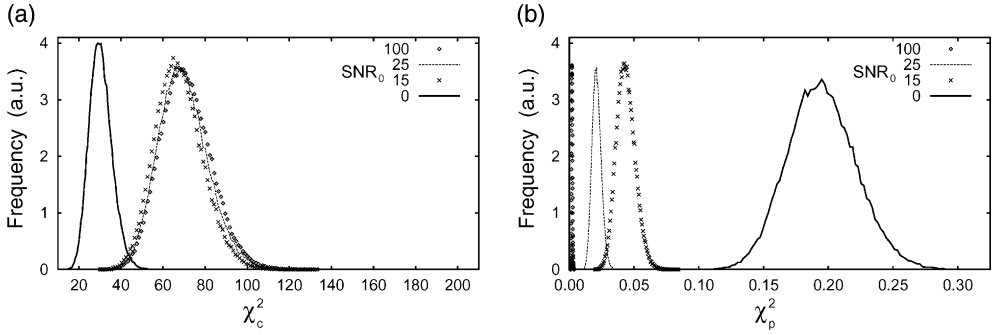

Fig. 1. Simulated χ^2_c (a) and χ^2_p (b) distributions for isotropy ($\lambda = 0.7 \times 10^{-3}$ mm^2/s) and for various SNR values of the baseline signal SNR_0. The scaling of the vertical axes is in arbitrary units (a.u.).

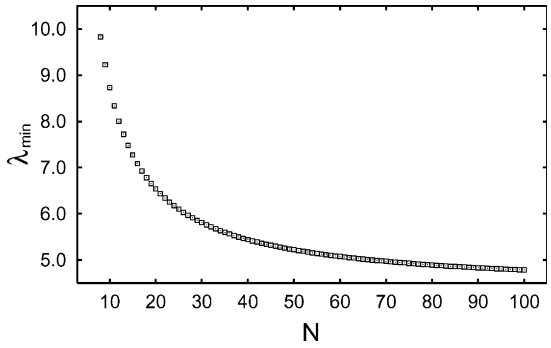

Fig. 2. Plot of λ_{min} (lower limit of the ratio $\chi^2_p(SNR_0 = 0)/\chi^2_p(SNR_0 \gg 0)$, given by Eq. (A.7)) as a function of the number N of fitted signals, for $SNR_0 = 15$.

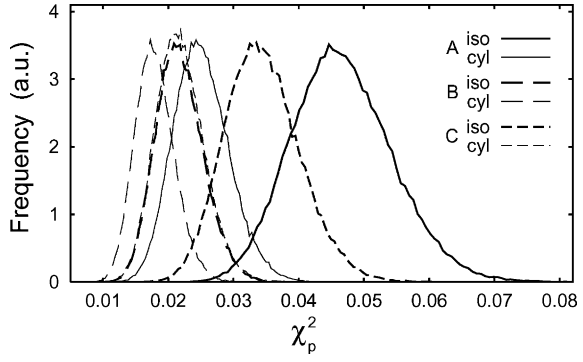

Fig. 3. Simulated χ^2_p distributions of isotropic (iso) and cylindrically symmetric strongly anisotropic (cyl) fibres for DT schemes A, B and C, using $SNR_0 = 15$.

the isotropic and anisotropic χ^2_p distributions varies with the DT scheme; it is largest for scheme A and smallest for scheme B. Since scheme A acquires the largest number of DW signals at the highest b-value, it produces the rightmost χ^2_p distributions for a given fibre type. Similarly, scheme B gives the leftmost χ^2_p distributions. Although the distributions of Fig. 3 correspond to $SNR_0 = 15$ (representative of the SNR_0 of our experimental data), the same relative shifts of the χ^2_p distributions occur for other SNR_0 values. It should be noted that, despite their dependence on experimental (DT scheme) and physiological (fibre shape) parameters, all the χ^2_p distributions of Fig. 3 remain well-separated from the χ^2_p distribution for $SNR_0 = 0$ (Fig. 1b).

Figs. 4a and b plot χ^2_p and χ^2_c pixel distributions over all the 13 normal subjects for standard and CSF-suppressed (FLAIR) DT-MRI, using DT scheme A. The distributions of both χ^2_p and χ^2_c are bimodal. For χ^2_p (Fig. 4b), pixels with low SNR_0 and/or inconsistent fitting contribute to the right-hand side lobe (referred thereafter as "noise" lobe) while those with consistent fitting (brain tissue) contribute to the left-hand side lobe (referred thereafter as "signal" lobe). For χ^2_c (Fig. 4a) the relative position of the two lobes is reversed. The χ^2_p sensitivity to SNR_0 and the FLAIR-induced decrease in SNR_0 account for the shift of the χ^2_p signal lobe in FLAIR compared with that in standard DTI. In agreement with the simulations, these distributions demonstrate that while the χ^2_p signal lobes are upper bound by the respective noise lobe, such a bound is absent for the case of χ^2_c. Therefore, χ^2_p as opposed to χ^2_c intrinsically provides a reliability measure of the DT-MRI data, since it ensures that "random" pixels will necessarily have large χ^2_p, and can be discarded.

Fig. 4c plots χ^2_p pixel distributions for the three DT schemes, using standard DT-MRI. In agreement with the simulations (Fig. 3), the position of the signal lobes of χ^2_p depends on the DT scheme, with the lobes of schemes C and B being shifted to the left of the signal lobe for scheme A. Specifically, the peaks of the signal lobes for schemes B and C are shifted relative to the peak of the signal lobe for scheme A, by about 40% and 20%, respectively (the shifts are reported as percentages of the χ^2_p value, corresponding to the peak position of the signal lobe for scheme A). This is comparable with the relative shift between the peaks of the signal lobes in standard and FLAIR DT-MRI for scheme A (Fig. 4b), which is 30% of the χ^2_p value, corresponding to the peak position of the signal lobe for FLAIR DT-MRI. Furthermore, the signal lobes of the three schemes have different widths; this is because, as shown in Fig. 3, the dependence of χ^2_p on fibre anisotropy varies with the DT

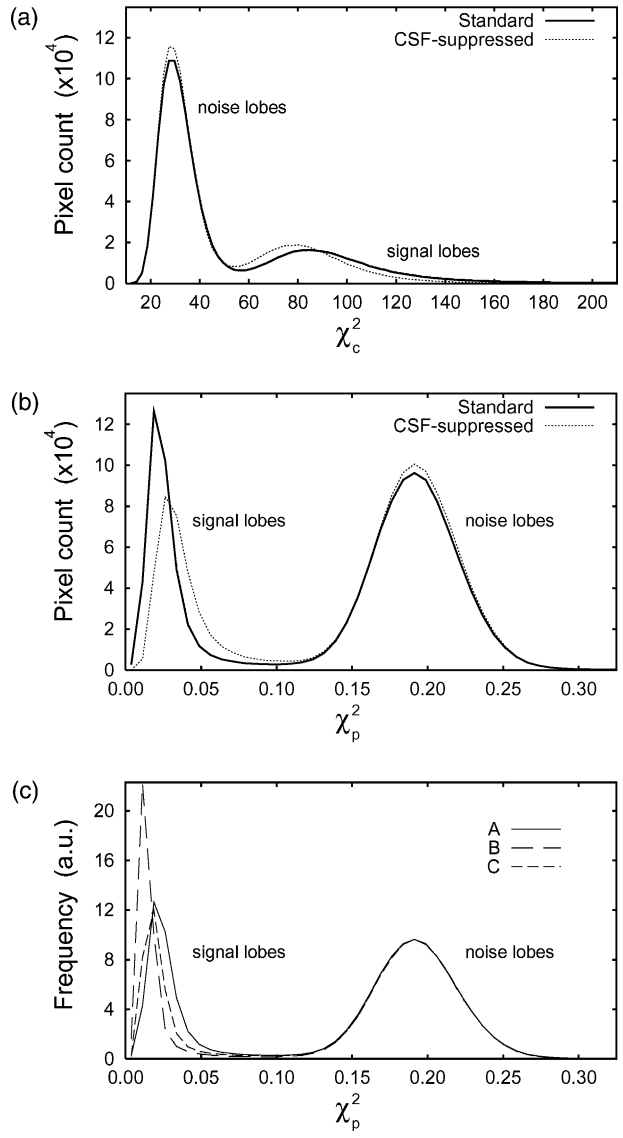

Fig. 4. Experimental χ^2_p and χ^2_c pixel distributions over all 13 normal subjects for standard (a) and CSF-suppressed (b) DT-MRI using DT scheme A. (c) Experimental χ^2_p pixel distributions for DT schemes A, B and C (the vertical axes have been scaled, so that the noise lobes have equal height).

scheme. Since this dependence was larger for scheme A, this scheme gave the widest signal lobe. Similarly, scheme B gave the narrowest signal lobe. Importantly, there is excellent agreement between the χ^2_p noise lobes of the three DT schemes, and these lobes are clearly separated from the signal lobes.

The experimental χ^2_p noise lobes (Fig. 4b) and the simulated χ^2_p distribution for $SNR_0 = 0$ (Fig. 1b) are superposed in Fig. 5a. These distributions peak at the same χ^2_p, and have very similar shape; the slight mismatch at their lower half most likely arises from pixels with $SNR_0 \gg 0$ and inconsistent fitting, such as those corresponding to Nyquist ghosting and the brain skull. As a result, the χ^2_p noise lobe can be characterised purely numerically using the simulated χ^2_p distribution for $SNR_0 = 0$. It also provides a data-independent reference relative to which the χ^2_p value of a specific pixel is assessed. A threshold χ^2_{p0} is therefore determined, so that only pixels with $\chi^2_p < \chi^2_{p0}$ are acceptable. Specifically, the DT-MRI data are tested against the null hypothesis $\mathscr{H}$, that fitting is entirely due to random magnitude variations in the data. The probability density function (pdf) p of $\mathscr{H}$ is the χ^2_p distribution for $SNR_0 = 0$. A small value P_0 of the cumulative probability P of $\mathscr{H}$ is then chosen (for example, $P_0 = 0.001$), such that when $P(\chi^2_p) < P_0$, the null hypothesis is rejected and the pixel is considered to contain reliable DT information. Thus, $P(\chi^2_{p0}) = P_0$. Since P is not analytically known, the χ^2_p distribution for $SNR_0 = 0$ needs to be generated empirically from simulations. Alternatively, χ^2_p for $SNR_0 = 0$ can be approximated by a scaled χ^2 statistic with $N - 7$ df (Eq. (A.4)), for which p and P are available analytically [8]. Fig. 5a (label "Analytical") plots p for the scaled χ^2 corresponding to the parameters used in this work ($N = 78, \mathrm{df} = 71, m = 1$). The analytical χ^2_p distribution peaks at the same χ^2_p as the empirical χ^2_p distributions (experimental or simulated) and is slightly wider than the latter. Therefore, the analytical threshold χ^2_{p0} will be smaller than that derived from the simulations, and this observation holds consistently over

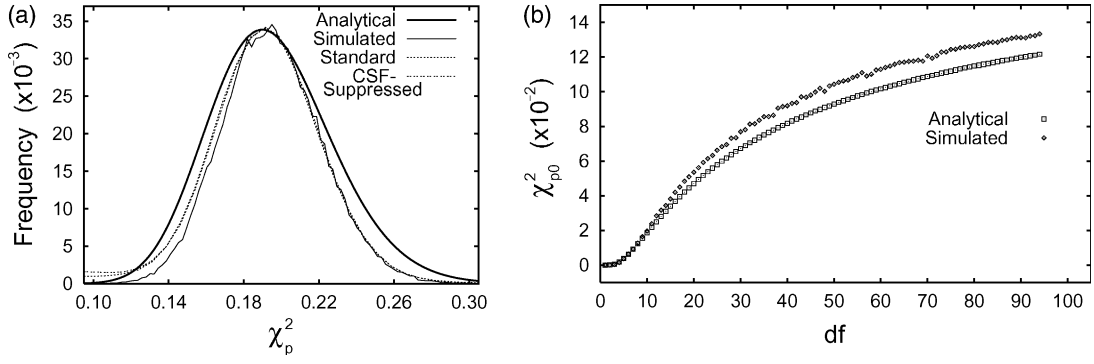

Fig. 5. (a) Superposition of empirical (simulated χ^2_p for $SNR_0 = 0$, Fig. 1b and experimental noise lobes, Fig. 4b) and analytical (scaled χ^2 with 71 df, using Eq. (A.4)) χ^2_p distributions for $SNR_0 = 0$. (b) Plot of χ^2_{p0} derived from simulations and the analytical approximation as a function of the degrees of freedom (df); $\mathrm{df} = N - 7$, N = number of fitted signals.

a wide range of df (Fig. 5b). Apart from its computational efficiency, the analytical approximation will thus result in reduced, more conservative thresholds χ^2_{p0}, increasing the confidence that the surviving pixels contain reliable DT information.

Fig. 6 shows χ^2 (χ^2_c and χ^2_p) maps of the same slice of a volunteer for standard and FLAIR DT-MRI. The χ^2_p maps (Figs. 6b and g for standard and FLAIR DT-MRI, respectively) have been scaled linearly from 0 to the analytical χ^2_{p0} ($=0.109$). The χ^2_c maps (Figs. 6c and h for standard and FLAIR DT-MRI, respectively) have been scaled linearly between 0 and 175, which corresponds to 10^{-10} probability that χ^2_c will exceed this value by chance (using the notation in [8], $Q(\chi^2_c = 175) = 1 - P(\chi^2_c = 175) \simeq 10^{-10}$, for df $= 71$). Baseline images are also shown for anatomical reference (Figs. 6a and f for standard and FLAIR DT-MRI, respectively). The χ^2 maps confirm the different sensitivity of χ^2_c and χ^2_p on SNR_0 and goodness of fit. The contrast in the χ^2_p maps is modulated jointly by SNR_0 and the fitting quality. Thus pixels with χ^2_p exceeding χ^2_{p0} correspond to: (i) very low SNR_0 (image background in all maps, and CSF areas in the FLAIR χ^2_p maps), and (ii) poorly fitted CSF areas (in standard DTI) and bordering brain tissue (in all maps). Within brain tissue, χ^2_p is smaller than χ^2_{p0} and the contrast is modulated locally by the goodness of fit (for example, hyperintense rim running diagonally between the left side of the sagittal sinus and the posterior tip of the right ventricle, most likely arising from systematic hardware-induced instability, causing, for example, incomplete fat suppression), SNR_0 (increased intensity in the left frontal lobe where SNR_0 is low) and degree of anisotropy (decreased intensity in strongly anisotropic regions, such as the internal capsule and the splenium of the corpus callosum). Contrast in the χ^2_c maps is modulated mainly by the goodness of fit. As a result, for standard DT-MRI, within brain volume the hyperintense areas in the χ^2_c maps agree well with those in the χ^2_p maps. At the same time, pixels with low SNR_0 appear hypointense in the χ^2_c maps and thus well fitted. Since FLAIR DT-MRI reduces PVA between tissue and CSF, hyperintense pixels (in the χ^2_c maps of standard DT-MRI, Fig. 6c) bordering cerebral tissue and CSF spaces appear isointense in the χ^2_c maps of FLAIR DT-MRI (Fig. 6h). However, these areas may still contain unreliable DT information due to their low SNR_0 and are clearly demarcated in the χ^2_p maps of FLAIR DT-MRI (Fig. 6g), causing the marked difference between χ^2_c and χ^2_p maps of FLAIR DT-MRI.

Figs. 6d, e, i, and j show relative anisotropy (RA) maps before (Figs. 6d and i) and after (Figs. 6e and j) χ^2_p thresholding (using the analytical χ^2_{p0}). In accordance with the respective χ^2_p maps, the RA maps confirm that thresholding had a stronger effect on the FLAIR than on the standard DT-MRI RA maps. Importantly, pixels within CSF and neighbouring tissue appearing as pseudo-anisotropic structures in standard RA maps (Fig. 6d, anterior part of the left ventricle) were also eliminated in Fig. 6e, as a result of the χ^2_p filtering.

Fig. 7 compares χ^2_p maps of the three DT schemes for standard DT-MRI. The maps for schemes A and B correspond to the same slice of one subject while the

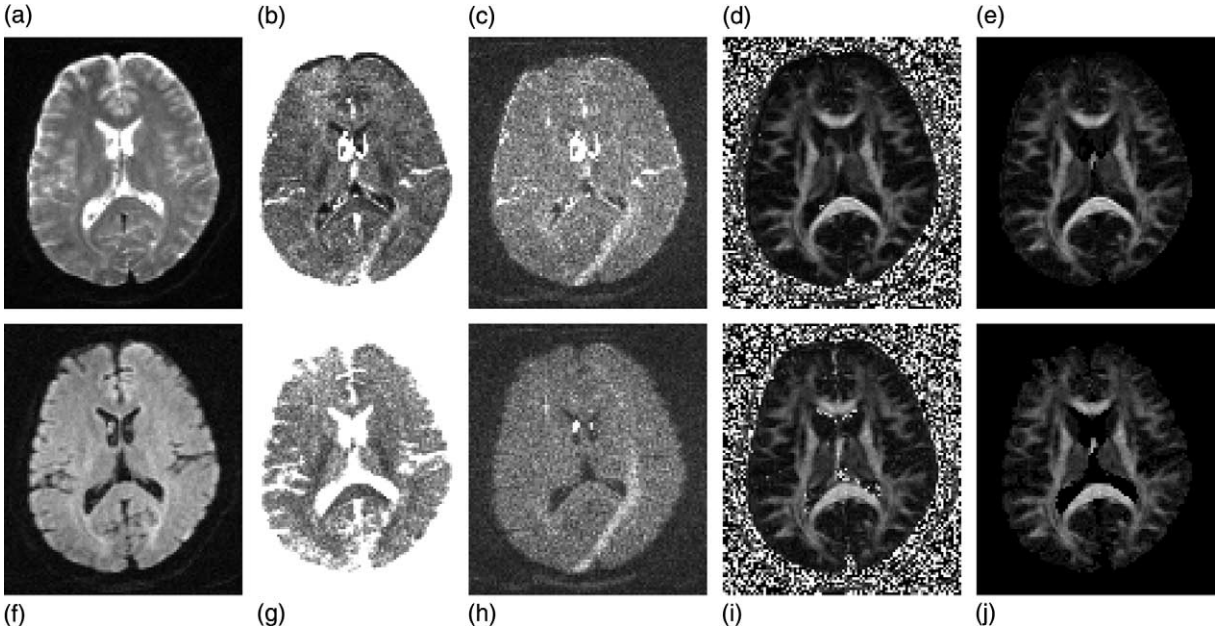

Fig. 6. Standard (a–e) and CSF-suppressed (f–j) DT-MRI data of a single slice from one subject: (a,f) baseline ($b=0$) images, (b,g) χ^2_p maps, scaled 0–0.109 ($=$ analytical χ^2_{p0}), (c,h) χ^2_c maps, scaled 0–175 ($Q(\chi^2_c = 175) = 1 - P(\chi^2_c = 175) \simeq 10^{-10}$, for df $= 71$), relative anisotropy (RA) maps, scaled 0–1, before (d,i) and after (e,j) χ^2_p thresholding.

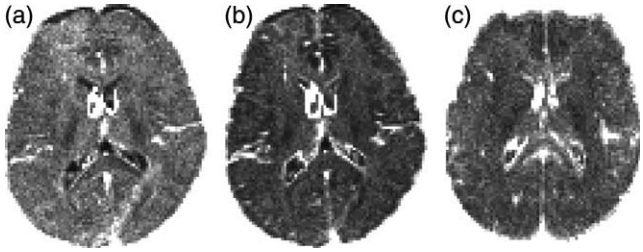

Fig. 7. χ^2_p maps of the three DT schemes using standard DT-MRI, scaled as in Figs. 6b,g. The same slice of one subject is shown for schemes A (a) and B (b); a homologous slice of a different subject is shown for scheme C (c). Phase-encoding direction: left–right (A and B), anterior–posterior (C).

map for scheme C corresponds to a homologous slice of a different subject. Phase-encoding direction is left–right for schemes A and B, while it is anterior–posterior for scheme C. Parenchymal tissue appears with highest χ^2_p values (though still smaller than χ^2_{p0}) in the map of scheme A and with lowest in that of scheme B. Furthermore, contrast between low- and high-anisotropy parenchymal areas appears stronger in the map of scheme A and smallest in that of scheme B. Thus, the maps of Fig. 7 confirm the results from the simulations (Fig. 3) and the data histograms (Fig. 4), concerning the dependence of χ^2_p for $SNR_0 \gg 0$ (or equivalently, the χ^2_p signal lobes) on DT scheme and fibre anisotropy. However, the most important finding is that the maps reveal similar hyperintense areas of unreliable DT information independent of the DT scheme, while using the same intensity scaling ($0 - \chi^2_{p0}$). It should be noted that, due to its phase-encoding direction, scheme C was the most sensitive to ventricular CSF flow and pulsation and, thus Fig. 7c showed more extensive hyperintense areas in the vicinity of the ventricles.

4. Discussion

This paper characterises and explores the properties of a modified fitting error parameter, χ^2_p, for DT-MRI data. The key feature of this parameter is that it provides clear separation between noise and signal lobes with the former lobe being an upper bound of the latter. This property is due to the normalisation of the squared fitting error ΔS^2 (indicative of the goodness of fit) by a quantity representative of the pixel SNR_0 (the total energy of the measured signals). Since, for given N and $SNR_0 \gg 0$, χ^2_p (or equivalently the χ^2_p signal lobes) depends on SNR_0 by definition (otherwise it would not have been able to differentiate between signal and noise lobes), it is not uniquely distributed. However, the distribution of χ^2_p when $SNR_0 = 0$ was unique for given N and in excellent agreement with experimental χ^2_p noise lobes. Thus, the rationale for using χ^2_p is essentially based on the invariant, and objectively defined (not affected by specific experimental parameters, apart from N) distribution of χ^2_p for $SNR_0 = 0$ (or equivalently the χ^2_p noise lobes). The exact position of the χ^2_p signal lobes is not critical as long as they are clearly positioned below the noise lobes; quantitatively, this is ensured by imposing the condition that the signal lobe is below the χ^2_p value (χ^2_{p0}) corresponding to the 0.001 percentile of the noise lobe. The additional dependence of χ^2_p (when $SNR_0 \gg 0$) on fibre shape and DT scheme, did not affect the validity of this condition. This was demonstrated by χ^2_p distributions (simulated and experimental) and maps using different DT schemes; all the χ^2_p signal lobes were clearly separated from the noise lobe while the maps depicted areas of unreliable DT information which were independent of DT scheme and fibre shape.

Since these results were obtained using diverse DT schemes, the DT scheme will not confound the separation between signal and noise χ^2_p lobes for $N \neq 78$. This ensures the applicability of χ^2_p beyond the specific DT schemes used in this work. However, it should be noted that, similar to χ^2_c, χ^2_p requires $N > 7$ so that it is not zero by definition. In practice this condition is often met, since the need for improved estimation of DT, causes most studies to use $N > 7$, applying DTI schemes with multiple b-values and/or DW directions [9]. Furthermore, if parameters such as SNR_0 and N are so low that the χ^2_p signal and noise lobes overlap, then the use of χ^2_p would not be informative because DT information will then be unreliable globally (for every pixel) not regionally; for example, calculated DT will have negative eigenvalues.

An implication of the dependence of χ^2_p on SNR_0 is that it cannot provide a goodness-of-fit measure for the DT model. On the contrary, similar to the majority of DT-MRI applications [9], it implicitly accepts this model, and, provides an ad hoc and heuristic measure of reliability of DT information. Low reliability (large χ^2_p) arises from large fitting error (for example, due to PVA between tissue and CSF, CSF flow/pulsation) and/or low energy of the measured signals (for example, due to low SNR_0), and may increase the apparent anisotropy of the affected areas. Highlighting such problematic areas is the main utility of the χ^2_p measure. Thus, χ^2_p should not be considered as an alternative for measures which test the validity of the DT model (classic χ^2_c) or evaluate and implement more complex models for analysis of DT-MRI data [10,11]. For example, taking into account the areas with $\chi^2_p > \chi^2_{p0}$ in the χ^2_p maps, it is unlikely that, brain regions, which do not obey the DT model just because of the presence of multiple intravoxel fibres, will have $\chi^2_p > \chi^2_{p0}$.

Alternative quantities may be used for normalisation of ΔS^2, aiming to reflect the pixel SNR_0. An example is the square of the baseline signal, either acquired or S_{f0} (Eq. (1)). Although the resultant alternative error measure will give signal lobes independent of DT scheme

and fibre shape, it will still depend on SNR_0. Most importantly, it will suffer several constraints. For example, using only the baseline signal for normalisation, this measure will underestimate areas of unreliable DT information, which have strong baseline signal; for example CSF spaces and surrounding tissue affected by PVA and CSF flow/pulsation. In contrast, since these areas have large apparent diffusivity, their DW signals will have small magnitude, leading to increased values of χ^2_p. Further constraints concern the noise lobes. For example, using a few averages for the measured baseline signal is not sufficient to produce well-defined noise lobes; the robustness of the χ^2_p noise lobes was due to the effective averaging of N noise signals. In fact, for given N, the noise lobes and the threshold of the alternative measure will not be uniquely determined, because they will depend on either the number of low b-value signals (which is specific to DT scheme) or the number of averages of the baseline ($b = 0$) signals (which may be acquired independently of the DT scheme); thus such alternative measure may be inappropriate for our descibed methodology.

A second feature of χ^2_p is that it is quantified directly from the data of the respective pixel, without relying on accurate calculation of data-dependent parameters, which may often be problematic. An example of such a parameter is the data SD (σ), required for calculation of χ^2_c. The presence of non-quantifiable noise sources (such as physiological noise, PVA and eddy currents), often causes σ to be underestimated leading to increased χ^2_c pixel values and exceedingly low χ^2_c probabilities of model acceptance [10]. This problematic quantification of χ^2_c is illustrated in: (i) Fig. 4a, where a large part of the signal χ^2_c lobes has $\chi^2_c > 114$, corresponding to $Q \ll 0.001$, and (ii) in Figs. 6c and h, where the upper intensity scaling corresponds to $Q \simeq 10^{-10}$, and brain tissue appears with uniformly high grayscale level, especially in Fig. 6c. As a result, calculation of χ^2_c is often limited to either spatially averaged values [12] (not allowing assessment on a pixel-by-pixel basis) or relative values with respect to the χ^2_c of a reference region (not allowing χ^2_c quantification through the χ^2 probability) [13]. In contrast, absolute pixelwise quantification of χ^2_p is possible without these complications. Furthermore, the analytic approximation of χ^2_p for $SNR_0 = 0$ (Eq. (A.4)) enables an objective (data-independent) and automated calculation of the threshold χ^2_{p0}, which is used for internal scaling of the χ^2_p maps and for data filtering. In contrast, the absence of the present analysis in [7] led to non-robust (data-dependent) determination of χ^2_{p0}, giving markedly different χ^2_{p0} for standard and FLAIR DT-MRI.

Compared with commonly used filtering methods, based on the magnitude of the baseline signal, χ^2_p provides more generic filtering since it is not biased towards an individual source of erroneous DT information. For example, magnitude thresholding could not have discarded the high SNR_0 pixels at the interface between CSF and tissue and within ventricular areas in the standard DT-MRI RA map of Fig. 6d. These areas appear often pseudo-anisotropic (cf, similar regions of the anisotropy maps in [14]), and thus, filtering using χ^2_p will eliminate such artifacts. It should be noted that since χ^2_p and χ^2_c give very similar hyperintense areas within brain for standard DT-MRI (Figs. 6b and c), a potential alternative to χ^2_p filtering, would have been the combination of magnitude thresholding with χ^2_c thresholding; the χ^2_c threshold should correspond to exceedingly small Q values (eg, 10^{-10}, which was used as the upper grayscale level in Figs. 6c and h), in order to account for underestimation of noise, as explained above. Compared with this alternative, χ^2_p thresholding has the advantages that it is both objective (it does not require estimation of data-dependent noise SD) and quantitative (it represents a confidence interval for the thersholding; 0.001 in this work).

Various extensions of the use and applications of χ^2_p are possible, but their detailed description exceeds the scope of this study which focuses on the characterisation of χ^2_p; therefore they are only briefly discussed. First, in addition to discarding pixels with $\chi^2_p > \chi^2_{p0}$, χ^2_p may be used to generate a continuous weighting function for data with $\chi^2_p < \chi^2_{p0}$, in which the smaller the χ^2_p value the higher the weight. Such weighting function may be useful in assessing mean values of DT parameters over selected regions or among different subjects. Since χ^2_p tends to take smaller values as anisotropy increases, the weighting will increase differences between low and high anisotropy areas. Secondly, given that the χ^2_p maps (Figs. 6b and g) show stronger contrast between brain structures than the χ^2_c maps (Figs. 6c and h), χ^2_p (as opposed to χ^2_c) may also be used for the assessment of tissue homogeneity within a region and for delineation of boundaries between tissue and CSF, complimentary to other scalars, such as the trace of the DT [15]. The sensitivity of χ^2_p to fitting error, SNR_0 and fibre shape will be particularly advantageous for this purpose. Conversely, the χ^2_p maps may be useful in assessing the regional extent of tissue with altered diffusion properties due to pathologies. For example, cysts and oedematous areas (due to stroke or trauma) will appear hypointense in the χ^2_p maps (because of their high SNR_0 and consistent fitting due to absence of flow), while the border of these areas will have increased χ^2_p (due to PVA with normal tissue). Importantly, although χ^2_p was defined and characterised using DT-MRI data, no assumption specific to DT-MRI was made. Thus, similar to χ^2_c, χ^2_p is generally applicable for the assessment of data fitting in the presence of noise; for example in the cases of apparent diffusion coefficient (ADC) and T_2 mapping.

In conclusion, the fitting error parameter χ^2_p separates effectively pixels with coherent signals from those with

random signals and is weighted by both the pixel SNR and the goodness of fit. It is quantified directly from the data of each pixel without requiring accurate estimation of data-dependent parameters, and its analytical approximations enable an objective and automated calculation of a threshold value, used for internal scaling of the χ^2_p maps. It is proposed as a measure for assessing reliability of DT-MRI data on a pixel-by-pixel basis.

Acknowledgments

The authors thank Dr. Alex Donev (Department of Probability and Statistics, University of Sheffield) for his advice on statistical aspects of this work and the anonymous reviewers for insightful suggestions.

Appendix A

Assuming m averages are performed in magnitude DT-MRI data, the following hold [6]:

$$\sigma_n = \frac{0.655}{\sqrt{m}}\sigma \quad \text{(a)} \qquad \langle S_n^2\rangle = (1.253\sigma)^2 + \sigma_n^2 \quad \text{(b)}, \tag{A.1}$$

where σ is the signal SD of a single average when $\mathrm{SNR}_0 \gg 0, S_n, \sigma_n$ are the magnitude and SD of averaged data containing only noise ($\mathrm{SNR}_0 = 0$) and $\langle\rangle$ denote mean value (mv). Following from Eqs. (2)–(4), it is, for $\mathrm{SNR}_0 = 0$:

$$\chi^2 = \Delta S^2/\sigma_n^2, \tag{A.2}$$

$$\sum_{i=1}^{N} S_{mi}^2 \approx N\langle S_n^2\rangle. \tag{A.3}$$

In Eq. (A.2), χ^2 follows the χ^2 distribution with $N-7$ degrees of freedom (df). Combining Eqs. (4), (A.1)–(A.3), we get:

$$\chi^2_p \approx \alpha\chi^2 \equiv \chi^2_s \qquad \alpha \equiv N^{-1}\left(1 + m1.913^2\right)^{-1}, \tag{A.4}$$

where χ^2_s is the scaled version of the χ^2 statistic and α is the scaling factor. For the parameters used in this work ($m = 1, N = 78$), $\alpha = 2.75 \times 10^{-3}$.

For $\mathrm{SNR}_0 \gg 0$, Eq. (4) can be written:

$$\chi^2_p = \frac{\sigma^2}{\sum_{i=1}^{N} S_{mi}^2}\chi^2. \tag{A.5}$$

For given N, σ, SNR_0 and χ^2, χ^2_p becomes maximum when the denominator in Eq. (A.5) is minimum. This happens when the DTI dataset consists of one baseline signal and $N-1$ maximally attenuated DW signals. Therefore:

$$\chi^2_{pm} = \frac{\sigma^2}{S_0^2\left(1 + (N-1)\beta^2\right)}\chi^2,$$

where χ^2_{pm} is the maximum χ^2_p for $\mathrm{SNR}_0 \gg 0$, S_0 is the magnitude of the baseline signal and β is the attenuation factor of the DW signal magnitude. Given that usually $bD \leqslant 1.1$ [9], a value of β which safely maximises χ^2_{pm} is $e^{-1.1} \simeq 0.3$. Defining $\mathrm{SNR}_0 \equiv S_0/\sigma$, such that SNR_0 describes the SNR of the baseline signal after any signal averaging, we have:

$$\chi^2_{pm} = \mathrm{SNR}_0^{-2}\left(1 + (N-1)\beta^2\right)^{-1}\chi^2. \tag{A.6}$$

Combining, Eqs. (A.4) and (A.6), an estimation of the lower limit $\lambda_{\min}$ of the ratio between χ^2_p for $\mathrm{SNR}_0 = 0$ and χ^2_p for $\mathrm{SNR}_0 \gg 0$ is:

$$\lambda_{\min} = \frac{\chi^2_s}{\chi^2_{pm}} = \alpha\mathrm{SNR}_0^2\left(1 + (N-1)\beta^2\right). \tag{A.7}$$

References

[1] P.J. Basser, Inferring microstructural features and the physiological state of tissues from diffusion-weighted images, NMR Biomed. 8 (1995) 333–344.
[2] C. Pierpaoli, P.J. Basser, Toward a quantitative assessment of diffusion anisotropy, Magn. Reson. Med. 36 (1996) 893–906.
[3] N.G. Papadakis, C.D. Murrills, L.D. Hall, C.L.-H. Huang, T.A. Carpenter, Minimal gradient encoding for robust estimation of diffusion anisotropy, Magn. Reson. Imaging 18 (2000) 671–679.
[4] K.K. Kwong, R.C. McKinstry, D. Chien, A.P. Crawley, CSF-suppressed quantitative single-shot diffusion imaging, Magn. Reson. Med. 21 (1991) 157–163.
[5] N.G. Zacharopoulos, P.A. Narayana, Selective measurement of white matter and gray matter diffusion trace values in normal brain, Med. Phys. 11 (1998) 2237–2241.
[6] R.M. Henkelman, Measurement of signal intensities in the presence of noise in MR images, Med. Phys. 12 (1985) 232–233.
[7] N.G. Papadakis, K.M. Martin, M.H. Mustafa, I.D. Wilkinson, P.D. Griffiths, C.L.-H. Huang, P.W.R. Woodruff, Study of the effect of csf suppression on white matter diffusion anisotropy mapping of healthy human brain, Magn. Reson. Med. 48 (2002) 394–398.
[8] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C. The Art of Scientific Computing, second ed., Cambridge University Press, Cambridge, 1992.
[9] D. LeBihan, J.-F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, H. Chabriat, Diffusion tensor imaging: concepts and applications, J. Magn. Reson. Imaging 13 (2001) 534–546.
[10] R. Shrager, D.K. Jones, S. Pajevic, P. Munson, P.J. Basser, When is a Gaussian displacement distribution adequate to describe water diffusion in tissues?, ISMRM Workshop on Diffusion MRI: Biophysical issues (what can we measure?) (2002) 21–25.
[11] D.C. Alexander, G.J. Barker, S.R. Arridge, Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data, Magn. Reson. Med. 48 (2002) 331–340.
[12] C.A. Clark, D. LeBihan, Water diffusion compartmentalization and anisotropy at high b values in the human brain, Magn. Reson. Med. 44 (2000) 852–859.
[13] F.A. Mamata, H. Jolesz, R.V. Mulkern, S.E. Maier, Tissue characterization with χ^2 maps of monoexponential diffusion signal fits, in: ISMRM 9th Scientific Meeting, vol. 2, 2001, p. 1512.

[14] K.O. Lim, S.J. Choi, N. Pomara, A. Wolkin, J.P. Rotrosen, Reduced frontal white matter integrity in cocaine dependence: a controlled diffusion tensor imaging study, Biol. Psychiatry 51 (2002) 890–895.

[15] A. Virta, A. Barnett, C. Pierpaoli, Visualizing and characterizing white matter fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI, Magn. Reson. Imaging 17 (8) (1999) 1121–1133.