The Non-Local Bootstrap – Estimation of Uncertainty in Diffusion MRI

Pew-Thian Yap*, Hongyu An, Yasheng Chen, and Dinggang Shen

Department of Radiology and Biomedical Research Imaging Center (BRIC)
The University of North Carolina at Chapel Hill, U.S.A.
{ptyap,hongyu_an,yasheng_chen,dgshen}@med.unc.edu

Abstract. Diffusion MRI is a noninvasive imaging modality that allows for the estimation and visualization of white matter connectivity patterns in the human brain. However, due to the low signal-to-noise ratio (SNR) nature of diffusion data, deriving useful statistics from the data is adversely affected by different sources of measurement noise. This is aggravated by the fact that the sampling distribution of the statistic of interest is often complex and unknown. In situations as such, the bootstrap, due to its distribution-independent nature, is an appealing tool for the estimation of the variability of almost any statistic, without relying on complicated theoretical calculations, but purely on computer simulation. In this work, we present new bootstrap strategies for variability estimation of diffusion statistics in association with noise. In contrast to the residual bootstrap, which relies on a predetermined data model, or the repetition bootstrap, which requires repeated signal measurements, our approach, called the *non-local bootstrap* (NLB), is non-parametric and obviates the need for time-consuming multiple acquisitions. The key assumption of NLB is that local image structures recur in the image. We exploit this self-similarity via a multivariate non-parametric kernel regression framework for bootstrap estimation of uncertainty. Evaluation of NLB using a set of high-resolution diffusion-weighted images, with lower than usual SNR due to the small voxel size, indicates that NLB is markedly more robust to noise and results in more accurate inferences.

1 Introduction

Diffusion magnetic resonance imaging (MRI) [16] reveals spectacular details of brain tissue micro-structures through observation of water diffusion patterns. It therefore captures vital information that is of paramount importance for *in vivo* investigation of white matter and connectivity alterations that are associated with brain diseases, development, and aging [24, 27–30]. However, the noisy nature of diffusion MRI data adversely affects the estimation precision of quantities such as local fiber orientations, which will eventually introduce uncertainty in important applications such as white matter fiber tractography. The impact of noise can be large, especially in high angular resolution diffusion imaging (HARDI), where relatively high diffusion weightings (i.e, b-values) are employed for increasing angular contrast.

* Corresponding author.

J.C. Gee et al. (Eds.): IPMI 2013, LNCS 7917, pp. 390–401, 2013.

Considerable efforts have been directed to modeling the variability caused by noise [8, 13]. However, the models used often assume normality and are yet to be verified in complex situations where various noise sources, such as physiologic variation, scanner instability, and imaging noise, might be simultaneously involved. The distributions of these types of noise are non-normal and cannot be adequately modeled using simple models that rely on the normality assumption. To avoid unrealistic assumptions, an appealing alternative is to use the bootstrap method.

The bootstrap method is a non-parametric procedure for estimating the statistical properties of a population from a limited number of measurement samples, without prior assumptions about the population distribution [6]. It was designed to replace complex and often inaccurate approximations to uncertainty measures with computer simulation based on real data [2]. For instance, the bootstrap has been shown to be capable of accurately estimating the true uncertainty in fiber orientations [14]. The repetition bootstrap and the residual bootstrap are two commonly used bootstrap techniques in medical image analysis. The repetition bootstrap [17, 18] depends on repeated measurements of signal for each diffusion-sensitizing gradient, a requirement which might be difficult to fulfill with limited acquisition time. On the other hand, the residual bootstrap [2, 15] is a model-based approach that resamples the residuals of a linear regression model that is fitted to the data. Since the residual bootstrap does not require repeated measurements, it can be applied to data that are acquired under clinically realistic scan times. However, since it relies on a model, the model needs to describe the signal measurements adequately so that the error terms across all diffusion gradients will have a common mean of zero. Extra care should also be taken so that the model does not overfit the signal, especially in noisy conditions, causing misleading reduction in variability.

In this work, we introduce novel non-parametric bootstrap strategies that will allow bootstrap samples to be generated from a single image without requiring repeated signal measurements as well as predetermined data models. Our approach hinges on the observation that local imaging information recurs in an image. This self-similarity implies that imaging information coming from spatially distant (non-local) regions can be exploited for more effective estimation of statistics of interest. In what follows, we will first show that estimation using non-local information can be seen as a non-parametric regression problem with a multivariate predictor variable that captures local neighborhood information and a univariate or multivariate response variable that is related to the statistic of interest. Based on this regression-based formulation, we then show that the uncertainty of the statistic can be estimated with bootstrap samples generated using case resampling or residual resampling. We will demonstrate with empirical evidence that the proposed bootstrap strategies are significantly more robust to noise and yields inferences that are markedly more accurate.

2 Approach

We will first discuss how the non-local estimation problem can be recast as a non-parametric kernel regression problem. We will then show how this regression-based formulation can be used to compute bootstrap estimates of the variability of statistics of concern.

2.1 Non-Local Estimation as Non-parametric Kernel Regression

Let $\mathbf{Z} = [(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_n, \mathbf{y}_n)]$ be a sample of n independent observations of bivariate random variable $(\mathbf{X}, \mathbf{Y})$. $\mathbf{X}$ is a $\mathbb{R}^{d_X}$-valued predictor random variable and $\mathbf{Y}$ is a $\mathbb{R}^{d_Y}$-valued response random variable. The regression function of $\mathbf{Y}$ on $\mathbf{X}$ is

$$m(\mathbf{x}) = \mathrm{E}(\mathbf{Y}|\mathbf{X} = \mathbf{x}). \tag{1}$$

The problem is to obtain an estimate of $m(\mathbf{x})$, $\hat{m}(\mathbf{x})$, using the n observations, such that $\hat{m}(\mathbf{x})$ tends to $m(\mathbf{x})$ as $n \to \infty$.

Nadaraya [21] and Watson [26] proposed to estimate $m(\mathbf{x})$ as a locally weighted average, using a kernel as a weighting function. The Nadaraya-Watson estimator is

$$\hat{m}_{\mathbf{H}}(\mathbf{x}) = \frac{\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)\mathbf{y}_i}{\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)}, \tag{2}$$

where $K_{\mathbf{H}}(\cdot) = |\mathbf{H}|^{-1} K(\mathbf{H}^{-1}\cdot)$ is a multivariate kernel function with symmetric positive-definite bandwidth matrix $\mathbf{H}$ [11]. $K(\mathbf{u})$ satisfies

$$\begin{aligned} &① \; K(\mathbf{u}) \geq 0, \; \forall \mathbf{u}; \quad ② \int K(\mathbf{u}) d\mathbf{u} = 1; \quad ③ \int \mathbf{u} K(\mathbf{u}) d\mathbf{u} = 0; \\ &④ \int \mathbf{u}\mathbf{u}^{\mathrm{T}} K(\mathbf{u}) d\mathbf{u} = \mu_2(K)\mathbf{I}, \; \mu_2(K) < \infty. \end{aligned} \tag{3}$$

The first and second requirements ensure that, when used for density estimation, the kernel results in a density estimate that is indeed a probability density function (i.e., non-negative with unit integral). The third requirement ensures that the expected value of the random variable computed from the estimated distribution is equal to the average of the observations. The last requirement ensures that the estimation bias is bounded. The estimation of $m(\mathbf{x})$ can be improved by employing locally weighted least squares regression, as shown in [23].

The estimator can be derived by noting that estimates of the unknown joint density function $f(\mathbf{x}, \mathbf{y})$ of $\mathbf{X}$ and $\mathbf{Y}$ and density function $f(\mathbf{x})$ of $\mathbf{X}$ can be obtained via kernel density estimation as

$$\hat{f}_{\mathbf{H},\mathbf{H}'}(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) K_{\mathbf{H}'}(\mathbf{y} - \mathbf{y}_i), \quad \hat{f}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i). \tag{4}$$

These, together with the fact that

$$\mathrm{E}(\mathbf{Y}|\mathbf{X}) = \int \mathbf{y} f(\mathbf{y}|\mathbf{x}) d\mathbf{y} = \int \mathbf{y} \frac{f(\mathbf{x}, \mathbf{y})}{f(\mathbf{x})} d\mathbf{y}, \tag{5}$$

lead to (2). From the theory of kernel density estimation, it is known that $\hat{f}(\mathbf{x}, \mathbf{y})$ and $\hat{f}(\mathbf{x})$ converge asymptotically ($n \to \infty$, $\mathbf{H} \to \mathbf{0}$, $n\mathbf{H} \to \boldsymbol{\infty}$) to the true densities of the underlying distribution.

Determining the non-local means [1] can be recast as a regression problem with voxel neighborhoods as the predictor and the corresponding central voxels as the response. The mean estimates are non-local as in principle they are computed from voxels

throughout the image, not limited by physical distance. For instance, if we let $\mathbf{x}_1 \ldots, \mathbf{x}_n$ be the intensity values of voxel blocks throughout the image (d_X = # voxels in each block, assuming each voxel is scalar-valued) and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be the intensity values of the corresponding central voxels ($d_Y = 1$), then, if $\mathbf{x}$ is the neighborhood intensity values of a voxel located at $\mathbf{p}$, $\hat{m}_{\mathbf{H}}(\mathbf{x})$ in (2) becomes the non-local mean at voxel location $\mathbf{p}$. This important observation, in addition to allowing us to improve non-local estimation using different kernel regression estimators that have been vastly studied [9], also provides the underpinning of the bootstrap strategies that will be discussed next. Before proceeding, we would like to highlight the fact that there is nothing in the regression framework that limits $\mathbf{X}$ and $\mathbf{Y}$ to random variables representing image intensity values; they can in fact be any features that are derived from the image. This regression framework also provides an explanation as to why non-local means denoising produces good white 'method noise' [1].

2.2 Non-Local Bootstrap Strategies

Bootstrap methods depend on the notion of *bootstrap samples*. Having observed a random sample of size n from a distribution with cumulative density function F (or probability density function f),

$$F \rightarrow (\mathbf{z}_1, \ldots, \mathbf{z}_n)\,, \tag{6}$$

the *empirical distribution function* $\hat{F}$ is defined to be the discrete distribution that puts probability $1/n$ on each $\mathbf{z}_i$. The arrow notation ($\rightarrow$) indicates that the sample values are outcomes of independent and identically distributed random variables, each with distribution function F, i.e., $\mathbf{z}_i \overset{\text{i.i.d.}}{\sim} F$. A bootstrap sample is defined as a random sample of size n, $\mathbf{Z}^* = (\mathbf{z}_1^*, \mathbf{z}_2^*, \ldots, \mathbf{z}_n^*)$, that is drawn from $\hat{F}$, i.e.,

$$\hat{F} \rightarrow (\mathbf{z}_1^*, \mathbf{z}_2^*, \ldots, \mathbf{z}_n^*)\,. \tag{7}$$

The star (*) notation indicates that $\mathbf{Z}^*$ is not the actual dataset $\mathbf{Z}$, but rather a randomized version of $\mathbf{Z}$ obtained via resampling with replacement. Generation of a significant amount of bootstrap samples allows us to estimate the sampling distribution of a statistic T, which can be used to make inferences about a population parameter θ. If we denote the estimate $\hat{\theta} = t(\mathbf{Z})$, for each bootstrap sample we can compute a *bootstrap replication* of $\hat{\theta}$,

$$\hat{\theta}^* = t(\mathbf{Z}^*), \tag{8}$$

a collection of which gives us an estimate of the sampling distribution of $\hat{\theta}$.

Recasting the non-local estimation problem in the form of kernel regression allows us to devise a number of bootstrap strategies for the estimation of the variability of statistics computed from the image. Out of the many possibilities, we will introduce here two non-local bootstrap strategies.

Case-Resampling Non-Local Bootstrap (CR-NLB): The first bootstrap scheme that we propose is called the *case-resampling non-local bootstrap* (CR-NLB). Recall from the discussion in the previous section that $\hat{m}_{\mathbf{H}}(\mathbf{x})$ is the regression function

for $\mathbf{Y}$ with respect to $\mathbf{X}$. If $\mathbf{p}_1, \ldots, \mathbf{p}_n$ denote the spatial locations of all voxels in an image, with corresponding neighborhood $\mathbf{x}_1, \ldots, \mathbf{x}_n$, then $\hat{m}_h(\mathbf{x}_1), \ldots, \hat{m}_{\mathbf{H}}(\mathbf{x}_n)$ are the non-local means at these locations. Determining the variability of the non-local means can hence be interpreted as determining the variability of the regression function. One way to achieve this is by resampling with replacement from the sample $\mathbf{Z} = [(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_n, \mathbf{y}_n)]$ to generate bootstrap sample $\mathbf{Z}^* = [(\mathbf{x}_{i_1}, \mathbf{y}_{i_1}), \ldots, (\mathbf{x}_{i_n}, \mathbf{y}_{i_n})]$, where $(i_1, \ldots, i_n)$ is a random sample of integers from 1 to n. For each bootstrap sample $\mathbf{Z}^*$, we generate a bootstrap replication of $\hat{m}_{\mathbf{H}}(\mathbf{x})$, denoted as $\hat{m}^*_{\mathbf{H}}(\mathbf{x})$. To estimate the standard error of $\hat{m}_{\mathbf{H}}(\mathbf{x})$, the bootstrap algorithm works by drawing a large number (B) of independent bootstrap samples $\mathbf{Z}^*_1, \ldots, \mathbf{Z}^*_B$, evaluating the corresponding bootstrap replications, and computing the empirical standard deviation of the replications. Note that bootstrap statistics of quantities other than the mean can be determined similarly.

CR-NLB is similar to the conventional repetition bootstrap as applied in [18]; but the 'repeated measurements' are now coming from non-local regions with voxel neighborhoods that are similar to the neighborhood of a voxel of interest. Since n is typically very large (possibly a few hundreds of thousands), the small sample size problem is now less of an issue. This approach does not rely on any assumptions regarding residuals of the sample with respect to the regression function and is hence less susceptible to difficulties caused by residual heteroscedasticity. But if the residuals are indeed homoscedastic, a residual resampling approach, discussed next, is a viable alternative.

Residual-Resampling Non-Local Bootstrap (RR-NLB): The second bootstrap scheme is called the *residual-resampling non-local bootstrap* (RR-NLB). RR-NLB resamples the residuals of the observations $\mathbf{Z} = [(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_n, \mathbf{y}_n)]$ with respect to the regression curve $\hat{m}_{\mathbf{H}}(\mathbf{x})$, i.e.,

$$\mathbf{r}_i = \mathbf{y}_i - \hat{m}_{\mathbf{H}}(\mathbf{x}_i). \tag{9}$$

The residuals will not necessarily have zero mean; so to let the resampled residuals reflect the behavior of the true observation errors, they should first be recentered as $\tilde{\mathbf{r}}_i = \mathbf{r}_i - \bar{\mathbf{r}}$, where $\bar{\mathbf{r}} = \frac{1}{n}\sum_i^n \mathbf{r}_i$. The residuals can be further corrected for leverage [5] by dividing each recentered residual by $\sqrt{1-g_i}$, where $g_i = K_{\mathbf{H}}(\mathbf{0}) / \sum_{j=1}^n K_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x}_j)$. Each bootstrap sample is generated via

$$\mathbf{Z}^* = \left[(\mathbf{x}_1, \hat{m}_{\mathbf{H}}(\mathbf{x}_1) + \tilde{\mathbf{r}}_{i_1}), \ldots, (\mathbf{x}_n, \hat{m}_{\mathbf{H}}(\mathbf{x}_n) + \tilde{\mathbf{r}}_{i_n})\right], \tag{10}$$

where $(i_1, \ldots, i_n)$, as before, is a random sample of integers from 1 to n. This bootstrap strategy allows residuals to be used from throughout the whole image, giving us a great deal more information at our disposal.

If $\mathbf{Y}$ is multivariate, as in the case of diffusion MRI where each voxel can be seen as containing a vector of signal measurements corresponding to different diffusion-sensitizing directions, a more confined sampling approach can be employed by limiting the sampling within each voxel to capture more localized and subtle variation. That is, assuming homoscedasticity across elements of the residual vector, we can generate a bootstrap sample via

$$\mathbf{Z}^* = \left[(\mathbf{x}_1, \hat{m}_{\mathbf{H}}(\mathbf{x}_1) + \tilde{\mathbf{r}}^*_1), \ldots, (\mathbf{x}_n, \hat{m}_{\mathbf{H}}(\mathbf{x}_n) + \tilde{\mathbf{r}}^*_n)\right], \tag{11}$$

where in this case, if we use $[\cdot]^{(k)}$ to denote the k-th element of a vector and if $\tilde{\mathbf{r}}_i = [\tilde{\mathbf{r}}_i^{(1)}, \ldots, \tilde{\mathbf{r}}_i^{(d_Y)}]^{\mathrm{T}}$, then $\tilde{\mathbf{r}}_i^* = [\tilde{\mathbf{r}}_i^{(j_1)}, \ldots, \tilde{\mathbf{r}}_i^{(j_{d_Y})}]^{\mathrm{T}}$, with $(j_1, \ldots, j_{d_Y})$ being a random sample of integers from 1 to d_Y. The residual is recentered differently from above as $\tilde{\mathbf{r}}_i^{(j)} = \mathbf{r}_i^{(j)} - \bar{r}_i$, where $\bar{r}_i = \frac{1}{d_Y} \sum_{j=1}^{d_Y} \mathbf{r}_i^{(j)}$. This approach is similar to the conventional residual bootstrap [2] but does not require a predetermined model. We will use this formulation in the experiments so that comparison can be made with respect to the conventional residual bootstrap [2].

2.3 Kernel and Bandwidth

A variety of kernel functions are possible in general [9]. Consistent with non-local means [1], we use a Gaussian kernel, i.e.,

$$K(\mathbf{u}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{u}\right), \tag{12}$$

and hence

$$K_{\mathbf{H}}(\mathbf{u}) = |\mathbf{H}|^{-1} K(\mathbf{H}^{-1}\mathbf{u}) = \frac{1}{\sqrt{2\pi}|\mathbf{H}|} \exp\left(-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{H}^{-2}\mathbf{u}\right). \tag{13}$$

The choice of $\mathbf{H}$ is dependent on the application. For simplicity, we require equal bandwidth h in all dimensions, corresponding to $\mathbf{H} = h\mathbf{I}$. If different bandwidths are needed, we can set $\mathbf{H} = \mathrm{diag}(h_1, \ldots, h_d)$. We can also set $\mathbf{H} = \boldsymbol{\Sigma}^{-\frac{1}{2}}$, the covariance matrix of the data. Using such a bandwidth matrix corresponds to a transformation of the data to obtain an identity covariance matrix.

The choice of h depends upon a trade-off between the bias and variance of the estimate of the regression function: a small h gives small bias and large variance, whereas a large h results in the opposite. We determine h based on the fact that, in non-local estimation, we are mainly interested in voxels with neighborhoods that ideally differ only by noise. That is, if the standard deviation of the noise is σ_{noise}, following [3] we set $h = \sigma_{\text{noise}}\sqrt{d_X}$. The noise level σ_{noise} can be estimated globally as shown in [19] or spatial-adaptively as shown in [20]; we implemented the former for simplicity.

2.4 Application to Uncertainty Estimation in Diffusion MRI

The non-local bootstrap strategies described above can be applied to uncertainty estimation in diffusion MRI data. To do this, we have to first define the predictor vector $\mathbf{x}$ and the response vector $\mathbf{y}$. Since, as noted in [4], noise will interfere with block matching in non-local estimation, we average the signal measurements across gradient directions for increased signal-to-noise ratio (SNR). More specifically, for each voxel $\mathbf{p}$, we define a predictor vector $\mathbf{x}$ based on the corresponding local neighborhood: $[\langle S(\mathbf{g}, \mathbf{p}(1))\rangle_{\mathbf{g}}, \ldots, \langle S(\mathbf{g}, \mathbf{p}(k))\rangle_{\mathbf{g}}, \ldots]^{\mathrm{T}}$, where $\langle S(\mathbf{g}, \mathbf{p})\rangle_{\mathbf{g}}$ denotes the averaging of diffusion-weighted signal $S(\mathbf{g}, \mathbf{p})$ over gradient direction $\mathbf{g} \in \mathcal{G}$. Here, $\mathbf{p}(k) \in \mathcal{N}(\mathbf{p})$, with $\mathcal{N}(\mathbf{p})$ denoting the neighborhood of $\mathbf{p}$. If the noise level for the measured signal is σ_{noise}, the noise level for the average signal is $\sigma_{\text{noise}}/\sqrt{|\mathcal{G}|}$, where $|\mathcal{G}|$ denotes the total number of gradient directions. The response vector $\mathbf{y}$ is simply defined as the signal vector at location $\mathbf{p}$.

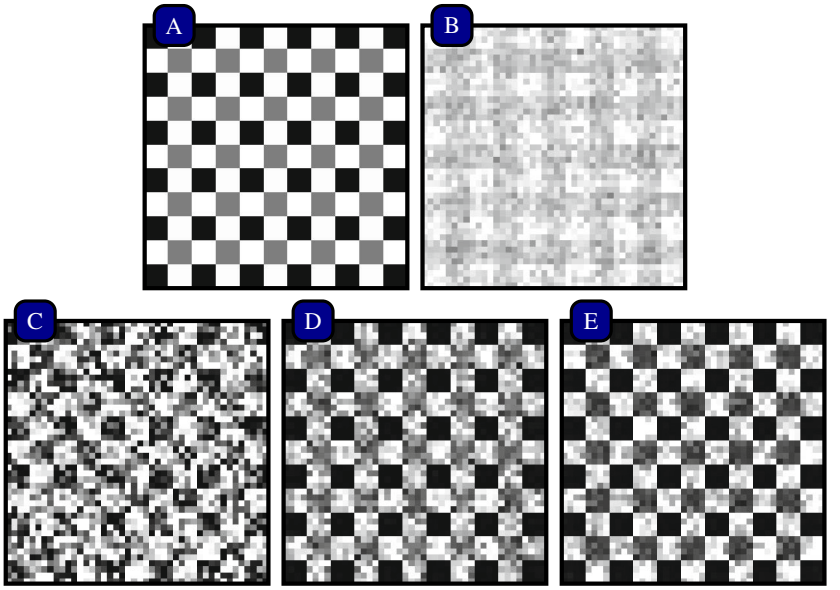

Fig. 1. Synthetic Data. Exemplar results for $\sigma_{\text{noise}} = 100$. (A) The anisotropy image of the noise-free DWI phantom. (B) The anisotropy image of a noisy realization of the phantom. Average ϕ-images over 10 noisy realizations of the phantom for (C) RR-B, (D) CR-NLB, and (E) RR-NLB.

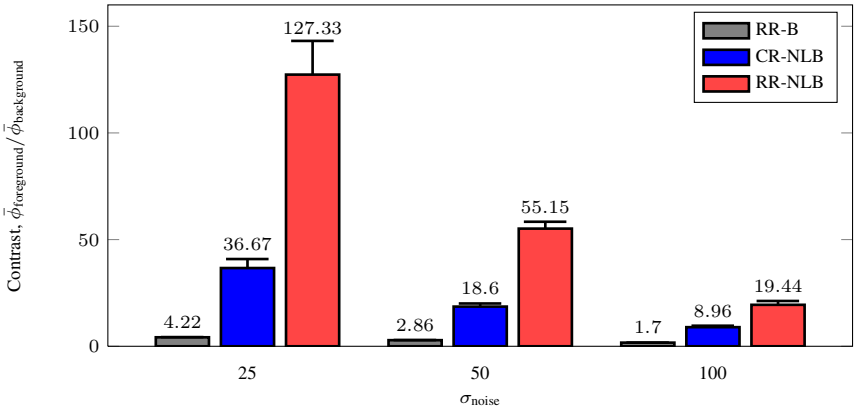

Fig. 2. Contrast. The average foreground-background contrast based on the ϕ-values. The error bars (some barely visible due to insignificant variability) indicate the standard deviations.

3 Experimental Results

We evaluated the effectiveness of the proposed bootstrap strategies (CR-NLB and RR-NLB) using *in silico* and *in vivo* diffusion MRI data. The conventional residual-resampling bootstrap (RR-B) [2], with spherical harmonics as the model [15], was used as comparison baseline. Unless otherwise stated, the neighborhood block radius was set to 2, giving a block of size $2(2) + 1 = 5$ in each dimension and a predictor vector $\mathbf{x}$ of size $d_X = 5^3|\mathcal{G}| = 125|\mathcal{G}|$. $B = 1{,}000$ bootstrap samples were generated for all cases. The bandwidth $\mathbf{H}$ was determined based on the data noise level, as described previously.

3.1 In Silico Data

We used diffusion tensors and their mixtures to generate a phantom of size 56×56 for evaluation [see Fig. 1(a)]. One group of tensors were oriented in the horizontal direction and another group in the vertical direction. At locations where these two groups cross, a mixture of two tensors with equal volume fraction was used to model the crossings. The tensor parameters were $\lambda_1 = 8 \times 10^{-4}\,\mathrm{mm}^2/\mathrm{s}$, $\lambda_2 = \lambda_3 = 1.5 \times 10^{-4}\,\mathrm{mm}^2/\mathrm{s}$, and $b = 1{,}000\,\mathrm{s/mm}^2$. The baseline signal was $S_0 = 1{,}000$. The background isotropic signal was generated using an isotropic tensor with $\lambda = 1.5 \times 10^{-3}\,\mathrm{mm}^2/\mathrm{s}$ and $S_0 = 500$. The (42) gradient directions were taken from the *in vivo* dataset (see next section). Note that these diffusion parameters were carefully chosen to mimic the *in vivo* data.

To evaluate the effectiveness of the proposed bootstrap strategies, we evaluated whether they result in correct inference of the anisotropy of the diffusion signal. To achieve this, we evaluated whether the high-order spherical harmonic coefficients are significant in the presence of noise. That is, we were interested whether the statistic

$$T = \sum_{l=2,4,6} \sum_{m=-l,-l+1,\ldots,l} c_{l,m}^2 \tag{14}$$

was significant. Coefficient $c_{l,m}$ is real-valued and is associated with the spherical harmonic function of order l and degree m. Only even orders were used because of the antipodal symmetric assumption associated with the diffusion signal. The more significant the coefficients associated with the higher order spherical harmonic functions ($l > 0$), i.e., the anisotropic energy, the greater the possibility that the diffusion signal is anisotropic. The coefficient for $l = 0$ is associated with isotropic diffusion. To gauge the significance, we determined the variability of T with respect to noise, i.e., the standard error of T, using CR-NLB, RR-NLB, and RR-B. An indicator of the significance of the observed anisotropic energy T_{observed} is given by measure

$$\phi = \frac{T_{\text{observed}}}{\hat{\text{se}}_B}, \tag{15}$$

where $\hat{\text{se}}_B$ is the bootstrap estimate of the standard error using B bootstrap samples. A large value for ϕ indicates high significance; a small value indicates otherwise.

For RR-B, the spherical harmonic coefficients estimated by fitting the spherical harmonic functions to the diffusion signal were used to compute T_{observed} as given by (14).

The standard error of the statistic was then estimated using the bootstrap replicates computed from the bootstrap samples that were generated via resampling of the residuals, as done in standard practice [2,12]. For CR-NLB and RR-NLB, T_{observed} was computed by spherical harmonics fitted to $\hat{m}_{\mathbf{H}}(\mathbf{x})$ and the respective standard errors were estimated based on the resampling schemes described in Section 2.2. A goal common to these methods is the attempt to determine the variability of some form of regression function.

For evaluation, various levels ($\sigma_{\text{noise}} = 25, 50, 100$) of Rician noise was added to the synthesized diffusion signal. Noise level $\sigma_{\text{noise}} = 100$ corresponds to that of the *in vivo* data. We computed the Rician noise corrupted signal $\tilde{S}$ as

$$\tilde{S} = \sqrt{(S + n_1)^2 + (n_2)^2} \tag{16}$$

where n_1 and n_2 are sampled from normal distributions with zero mean and variance σ^2_{noise}. Each value for $\tilde{S}$ is a sample from a Rician distribution with parameters S and σ_{noise}.

We computed ϕ with 10 noisy realizations of the phantom and the average ϕ-images for the case of $\sigma_{\text{noise}} = 100$ are shown in Fig. 1. Note that ideally the ϕ-value should be high for the (anisotropic) foreground and low for the (isotropic) background. RR-B [Fig. 1(C)] is apparently less capable in differentiating between isotropic and anisotropic voxels in severely noisy conditions. CR-NLB [Fig. 1(D)] and RR-NLB [Fig. 1(E)] yield much better contrast between the foreground and the background. Fig. 1(B) shows that, due to noise, the contrast of the (generalized) anisotropy between the anisotropic and isotropic regions is low. Despite the heavy amount of noise, the results indicate that the anisotropy of the original data [Fig. 1(A)] can still be inferred quite accurately using CR-NLB and RR-NLB. For quantitative evaluation, we computed the contrast between the foreground and the background, i.e., the ratio of the ϕ-value averages: $\bar{\phi}_{\text{foreground}}/\bar{\phi}_{\text{background}}$. The results for all tested noise levels, shown in Fig. 2, again validate that CR-NLB and RR-NLB give much improved contrast over RR-B.

3.2 In Vivo Data

A set of high-resolution $(1\text{mm})^3$ diffusion-weighted images were acquired using a Siemens 3T TIM Trio MR scanner with the acquisition technique reported in [22]. Diffusion gradients were applied in 42 non-collinear directions with diffusion weighting $b = 1{,}000\,\text{s/mm}^2$. The imaging matrix was 192×192 with a field of view of $192 \times 192\,\text{mm}^2$. The slice thickness was 1 mm. The low SNR of this dataset, due to the small voxel volume, makes it ideal for testing the various bootstrap strategies.

It is a well-accepted fact that the anisotropy of the white matter (WM) is high owing to diffusion anisotropy resulting from restricted diffusion. On the other hand, the diffusion signal captured from the gray matter (GM) and the cerebrospinal fluid (CSF) should be very close to isotropic. We expect to see similar patterns for results given by the different bootstrap methods.

The same evaluation, as performed on the synthetic data, was performed on this *in vivo* data. The results, shown in Fig. 3, indicate that CR-NLB and RR-NLB correctly yield high ϕ-values for WM regions and low ϕ-values for GM and CSF regions. RR-B

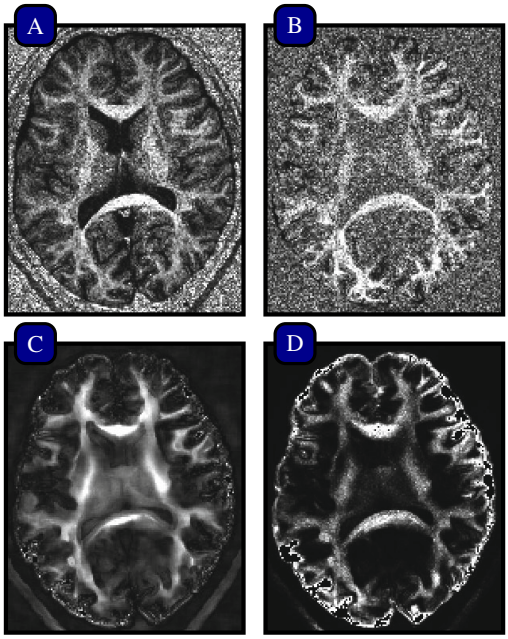

Fig. 3. Real Data. The ϕ-images as indicator of anisotropy. (A) The anisotropy image of the DWI data. The results given by (B) RR-B, (C) CR-NLB, and (D) RR-NLB.

shows a much reduced contrast between these regions, indicating low differentiability between them. As was done for the synthetic data, we computed the contrast between the anisotropic WM regions and the isotropic GM and CSF regions. We obtained results similar to those of the synthetic data: 1.68 for RR-B, 2.51 for CR-NLB, and 2.26 for RR-NLB. These values indicate that CR-NLB and RR-NLB are significantly more likely than RR-B to result in correct inferences of the anisotropy in the presence of noise.

4 Conclusion

We have presented two novel bootstrap strategies that show that bootstrapping can be performed within a regression framework that pulls together non-local information coming from the whole image. To the best of our knowledge, this is the first work that marries non-local estimation, non-parametric kernel regression, and the bootstrap in a single unified framework. Recasting non-local means as a regression problem allows us to further refine non-local means by studying different kernel estimators, by performing bias analysis [10], and by incorporating other more advanced techniques such as adaptive kernel regression [25] and local polynomial regression [7]. Our results indicate that both case-resampling and residual-resampling non-local bootstrap approaches yield results that are markedly better than the commonly used residual bootstrap. Future directions entail applying these bootstrap strategies to evaluating the variability of local fiber orientations, fiber trajectories, and connectivity between brain regions.

Acknowledgment. This work was supported in part by a UNC start-up fund and NIH grants (EB006733, EB008374, EB009634, MH088520, AG041721, and MH100217).

References

1. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. Multiscale Modeling and Simulation 4(2), 490–530 (2005)
2. Chung, S., Lu, Y., Henry, R.G.: Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. NeuroImage 33(2), 531–541 (2006)
3. Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Transaction on Medical Imaging 27, 425–441 (2008)
4. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3D transform-domain collaborative filtering. IEEE Transactions on Image Processing 16(8), 2080–2095 (2007)
5. Davison, A., Hinkley, D.: Bootstrap Methods and their Application. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (1997)
6. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. Monographs on Statistics and Applied Probablilty. Chapman and Hall (1994)
7. Fan, J., Gijbels, I.: Local Polynomial Modelling and Its Applications. Monographs on Statistics and Applied Probablilty. Chapman and Hall (1996)
8. Friman, O., Farnebäck, G., Westin, C.F.: A Bayesian approach for stochastic white matter tractography. IEEE Transactions on Medical Imaging 25, 965–977 (2006)
9. Härdle, W.: Applied Nonparametric Regression. Cambridge University Press (1992)
10. Härdle, W., Bowman, A.W.: Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands. Journal of the American Statistical Association 83(401), 102–110 (1988)
11. Härdle, W., Müller, M.: Multivariate and semiparametric kernel regression. In: Schimek, M.G. (ed.) Smoothing and Regression: Approaches, Computation, and Application. Wiley & Sons, Inc., Hoboken (2000)
12. Haroon, H.A., Morris, D.M., Embleton, K.V., Alexander, D.C., Parker, G.J.M.: Using the model-based residual bootstrap to quantify uncertainty in fiber orientations from Q-ball analysis. IEEE Transaction on Medical Imaging 28(4), 535–550 (2009)
13. Jbabdi, S., Woolrich, M., Andersson, J., Behrens, T.: A Bayesian framework for global tractography. NeuroImage 37(1), 116–129 (2007)
14. Jeurissen, B., Leemans, A., Tournier, J.D., Sijbers, J.: Can residual bootstrap reliably estimate uncertainty in fiber orientation obtained by spherical deconvolution from diffusion-weighted MRI? In: Proceedings 14th Annual Meeting of the Organization of Human Brain Mapping (2008)
15. Jeurissen, B., Leemans, A., Jones, D.K., Tournier, J.D., Sijbers, J.: Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. Human Brain Mapping 32(3), 461–479 (2011)
16. Johansen-Berg, H., Behrens, T.E. (eds.): Diffusion MRI — From Quantitative Measurement to In-Vivo Neuroanatomy. Elsevier (2009)
17. Jones, D.: Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. Magnetic Resonance in Medicine 49(1), 7–12 (2003)
18. Lazar, M., Alexander, A.L.: Bootstrap white matter tractography (BOOT-TRAC). NeuroImage 24(2), 524–532 (2005)

19. Manjón, J., Carbonell-Caballero, J., Lull, J., García-Martí, G., Martí-Bonmatí, L., Robles, M.: MRI denoising using non-local means. Medical Image Analysis 12(4), 514–523 (2008)
20. Manjón, J., Coupé, P., Martí-Bonmatí, L., Collins, D., Robles, M.: Adaptive non-local means denoising of MR images with spatially varying noise levels. Journal of Magnetic Resonance Imaging 31(1), 192–203 (2010)
21. Nadaraya, E.: On estimating regression. Theory of Probability and its Applications 9(1), 141–142 (1964)
22. Porter, D.A., Heidemann, R.M.: High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition. Magnetic Resonance in Medicine 62(2), 468–475 (2009)
23. Ruppert, D., Wand, M.: Multivariate locally weighted least squares regression. The Annals of Statistics 22(3), 1346–1370 (1994)
24. Shi, F., Yap, P.T., Gao, W., Lin, W., Gilmore, J., Shen, D.: Altered structural connectivity in neonates at genetic risk for schizophrenia: A combined study using morphological and white matter networks. NeuroImage 62(3), 1622–1633 (2012)
25. Silverman, B.: Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probablilty. Chapman and Hall (1998)
26. Watson, G.: Smooth regression analysis. Sankhyā: The Indian Journal of Statistics Series A 26(4), 359–372 (1964)
27. Wee, C.Y., Yap, P.T., Li, W., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D.: Enriched white matter connectivity networks for accurate identification of MCI patients. NeuroImage 54(3), 1812–1822 (2010)
28. Wee, C.Y., Yap, P.T., Zhang, D., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D.: Identification of MCI individuals using structural and functional connectivity networks. NeuroImage 59(3), 2045–2056 (2012)
29. Yap, P.T., Fan, Y., Chen, Y., Gilmore, J., Lin, W., Shen, D.: Development trends of white matter connectivity in the first years of life. PLoS ONE 6(9), e24678 (2011)
30. Yap, P.T., Wu, G., Shen, D.: Human brain connectomics: Networks, techniques, and applications. IEEE Signal Processing Magazine 27(4), 131–134 (2010)