Post mortem validation of DSI compressed sensing with optical imaging

Robert Jones, Chiara Maffei, Jean Augustinack, Bruce Fischl, Hui Wang, Berkin Bilgic, Anastasia Yendiki

Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital & Harvard Medical School, Charlestown, MA, USA

Abstract

Compressed sensing (CS) has been used to enhance the feasibility of diffusion spectrum imaging (DSI) by reducing the required acquisition time. CS applied to DSI (CS-DSI) attempts to reconstruct diffusion probability density functions (PDFs) from significantly undersampled q-space data. Dictionary-based CS-DSI using L2-regularized algorithms is an intriguing approach that has demonstrated high fidelity reconstructions, fast computation times and inter-subject generalizability when tested on *in vivo* data. CS-DSI reconstruction fidelity is typically evaluated using the fully sampled data as ground truth. However, it is difficult to gauge how great an error with respect to the fully sampled PDF we can tolerate, without knowing whether that error also translates to substantial loss of accuracy with respect to the true fiber orientations. Here, we obtain direct measurements of axonal orientations in *ex vivo* human brain tissue at microscopic resolution with polarization-sensitive optical coherence tomography (PSOCT). We employ dictionary-based CS reconstruction methods to DSI data from the same samples, acquired at high max b-value (40000 s/mm^2) and with high spatial resolution. We compare the diffusion orientation estimates from both CS and fully sampled DSI to the ground-truth orientations from PSOCT. This allows us to investigate the conditions under which CS reconstruction preserves the accuracy of diffusion orientation estimates with respect to PSOCT. We find that, for a CS acceleration factor of R=3, CS-DSI preserves the accuracy of the fully sampled DSI data. That acceleration is sufficient to make the acquisition time of DSI comparable to that of state-of-the-art single- or multi-shell acquisitions. We also show that, as the acceleration factor increases further, different CS reconstruction methods degrade in different ways. Finally, we find that the signal-to-noise (SNR) of the training data used to construct the dictionary can have an impact on the accuracy of the CS-DSI, but that there is substantial robustness to loss of SNR in the test data.

1. Introduction

Diffusion magnetic resonance imaging (dMRI) has played an integral role in the study of human brain circuitry *in vivo* by enabling non-invasive investigation of tissue architecture (Le Bihan et al., 1986). The molecular displacements resulting from water diffusion can be estimated from dMRI measurements acquired with a pulsed gradient spin-echo (PGSE) sequence (Stejskal & Tanner, 1965). Diffusion tensor imaging (DTI), the seminal approach for quantitative reconstruction of 3D water molecule displacement (Basser et al., 1994a, 1994b), assumes a 3D Gaussian distribution of water molecule displacements and thus can only model a single fiber population in each voxel. Diffusion spectrum imaging (DSI), combined with an acquisition that samples the entire q-space on a Cartesian grid (Wedeen et al., 2005), is capable of delineating multiple intravoxel fiber populations. The diffusion ensemble average propagator (EAP), or the 3D probability density function (PDF) of spin displacements in a voxel, can be recovered directly from a Fourier transform on the normalized Cartesian q-samples, and yields a plethora of information describing the angular and radial features of diffusion (Hagmann et al., 2008; Wedeen et al., 2008). DSI has been shown to provide more accurate fiber orientation estimates than other reconstruction methods in simulations (Daducci et al., 2013) and comparisons to optical imaging measurements (Jones et al., 2020), as well as more accurate tractography in comparisons to anatomical tracing (Maffei et al., 2020). Furthermore, dense q-space sampling provides the flexibility of resampling the data onto q-shells at various b-values, as needed for analyses that require shell acquisition schemes (Jones et al., 2020). However, the applicability of DSI is drastically limited by the lengthy acquisition times required to encode the full q-space and the long echo times (TEs) required to achieve high b-values, which lead to time-consuming and inefficient k-space acquisitions (Reese et al., 2009).

Much recent work has been devoted to accelerating MR acquisitions. Some approaches modify imaging sequences to allow multiple image slices to be acquired simultaneously. Multislice parallel imaging techniques, such as simultaneous multi-slice (SMS) (Setsompop, Cohen-Adad, et al., 2012; Setsompop et al., 2018; Setsompop, Gagoski, et al., 2012), simultaneous image refocusing (SIR) (Reese et al., 2009), and multiplexed (SMS+SIR) echo planar imaging (EPI), have played a crucial role in reducing dMRI scan times down to reasonable lengths. However, they do not address the large number of q-space samples required by DSI. To tackle this problem, compressed sensing (CS) reconstruction has been applied to DSI (CS-DSI). CS theory exploits transform sparsity to recover signals from sub-Nyquist acquisitions (Donoho, 2006; Lustig et al., 2007; Lustig et al., 2008). CS-DSI undersamples in q-space and reconstructs the missing samples with CS, allowing for a reduction in acquisition time directly proportional to the CS acceleration factor. Combining CS-DSI with SMS or multiplexed EPI can provide even higher accelerations (Setsompop et al., 2013), and render CS-DSI a practical diffusion protocol for long-term population studies (Tobisch et al., 2018).

A number of previous works have investigated CS-DSI reconstruction. Menzel et al. (2011) used wavelet and total variation (TV) penalties on PDFs combined with random Gaussian undersampling patterns, and concluded that angular and radial diffusion properties were preserved at R=4 acceleration. Paquette et al. (2015) performed a joint comparison of different wavelet-based sparsifying transforms and q-space sampling strategies, and found the best results when using a "uniform-angular, random-radial" undersampling mask combined with discrete wavelet transform (DWT). Both Menzel et al. and Paquette et al. used fixed transforms to generate sparse signal representations. Conversely, Bilgic et al. (2012) proposed CS-DSI using adaptive PDF dictionaries, combining the K-SVD algorithm (Aharon et al., 2006) for dictionary

training with the FOcal Underdetermined System Solver (FOCUSS) algorithm (Gorodnitsky & Rao, 1997) to solve the CS problem. While this approach yielded reduced reconstruction errors compared to fixed transforms, the iterative FOCUSS reconstruction resulted in full brain computation times on the order of days. This bottleneck was addressed in subsequent work (Bilgic et al., 2013), where two dictionary-based, L2-regularized methods were introduced that reduce computation times down to seconds per slice. They provide fast, simple formulations while preserving reconstruction quality compared to Dictionary-FOCUSS. One of the key findings from (Bilgic et al., 2013) was that forcing the PDFs to remain in the range of a dictionary was more important than the sparsity constraints imposed on the transform coefficients. In other words, the key to good reconstructions lies in the prior information encoded in a dictionary, and not the regularization norm that is applied on the dictionary transform coefficients.

Dictionary-based CS-DSI is a promising approach, but nonetheless has aspects that require further investigation. One of these areas is the effect of dMRI signal-to-noise ratio (SNR) on the CS algorithms and dictionary learning, which is yet to be well characterized (Bilgic et al., 2012). Determining how the SNR of the training data influences reconstructions and the minimum SNR level required for high-quality CS reconstruction would provide critical insights for the development of CS-DSI protocols. Traditionally, CS-DSI reconstruction fidelity is assessed by using the fully sampled DSI data as the gold standard to calculate error metrics (Bilgic et al., 2013; Bilgic et al., 2012; Menzel et al., 2011). A drawback of this approach is that the fully sampled DSI data are inherently corrupted with noise, particularly at high b-values. Bilgic et al. (2013) ameliorated this issue by sampling a few q-space locations 10 times, and using these low-noise references to evaluate reconstructions. Interestingly, they found that CS reconstructions exhibited lower errors than the fully sampled (1-average) data when compared to the 10-average data. These results exemplify potential denoising benefits of CS-DSI with respect to fully sampled data collected at the same SNR, but also highlight inadequacies of taking the fully sampled data as ground truth. Even if one could acquire noiseless fully sampled DSI data, it would be difficult to gauge how great an error with respect to those data we can tolerate in our CS reconstructions, without knowing how this translates to errors with respect to the true fiber orientations. Thus, it is critical to obtain independent measurements of fiber orientations from a modality that does not rely on water diffusion.

Here, we address this with a validation study of CS-DSI reconstructions in *ex vivo* human brain. Microscopic-resolution fiber orientations measured with polarization sensitive optical coherence tomography (PSOCT) serve as a reference for the mesoscopic-resolution diffusion orientations obtained from dMRI of the same tissue. PSOCT uses polarized light to probe tissue birefringence and obtain undistorted, direct measurements of in-plane fiber orientations at microscopic resolutions (Wang et al., 2014), and has been established in previous studies as a viable reference modality for the cross-validation of dMRI-derived orientation estimates (Jones et al., 2020; Wang et al., 2014). We use the dictionary-based techniques from Bilgic et al. (2013) to perform CS reconstructions, with dictionaries trained on DSI data from three different *ex vivo* human brain blocks to determine generalizability. We investigate if fiber orientations estimated from CS reconstructed data can achieve the same accuracy as those estimated from the fully sampled data. We investigate the influence of the SNR of the training or test data on the accuracy of the diffusion orientation estimates, and particularly how errors in the PDFs of CS-DSI with respect to the PDFs of fully sampled DSI translate to errors with respect to the ground-truth axonal orientations from PSOCT. Our findings elucidate the effect of the CS acceleration factor,

CS reconstruction method, and data SNR on the accuracy of dictionary-based CS-DSI reconstructions. We show that, for an acceleration factor of R=3, the accuracy of CS-DSI is very similar to that of fully sampled DSI, while its acquisition time becomes comparable to that of currently deployed multi-shell acquisitions. This could have implications for the adoption of CS-DSI in future dMRI studies.

2. Methods

2.1. Sample identification

The samples used in this study were extracted from two human brain hemispheres that were obtained from the Massachusetts General Hospital Autopsy Suite and fixed in 10% formalin for at least two months. Demographic information about the hemispheres is given in Table 1. Three samples were extracted from different anatomical locations of the hemispheres. Sample 1A (Figure 1A) was cut from brain 1 and was approximately 3x2x2 cm. It contained an area of deep white matter (WM) that included the corpus callosum (CC), the internal and external capsules (IC and EC, respectively), the caudate nucleus and the putamen. Sample 1B (Figure 1B) was taken from a different region of brain 1. The block was approximately 3x2x2 cm and contained an area of deep WM including the posterior internal capsule, putamen and thalamus. Sample 2 (Figure 1C) was cut from brain 2 and was sized approximately 2x2x3 cm. The superior part of the block contained the anterior superior frontal gyrus, the medial side included the cingulate sulcus, and the lateral side contained parts of the corticospinal tract and dorsal superior longitudinal fasciculus (SLF-I).

Brain	Age	Gender	Laterality	Cause of death	Diagnosis	PMI
1 (samples A and B)	43	F	Right	Cardiac arrest	Cognitive control	<24 h
2	70	M	Left	Coronary artery disease and multiorgan failure	Severe coronary artery disease, myocardial infarction, hypertension, hyperlipidemia, mild hypoxic/ischemic changes	24 h

Table 1. Demographic information on post mortem human hemispheres.

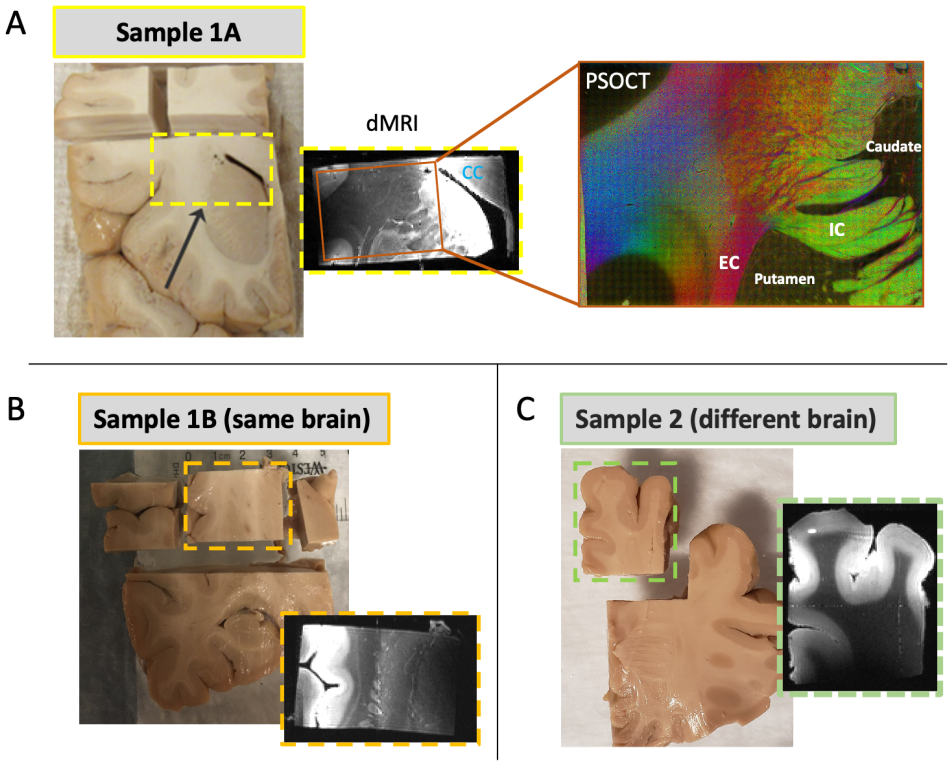

Figure 1. Overview of sample identification and data acquisition. *Ex vivo* human brain samples were extracted from coronal slabs (dashed rectangles) and dMRI data were acquired at 9.4T. Single slices from b=0 scans are shown for each sample. Samples 1A and 1B were cut from different anatomical locations of the same brain, and sample 2 was extracted from a different brain. Following dMRI, a piece of sample 1A was cut and imaged with PSOCT (A, right). Sample 1A, which contained the corpus callosum (CC), internal capsule (IC), external capsule (EC), caudate and putamen, was used as the test dataset. Each of samples 1A, 1B, and 2 were used as training datasets.

2.2. Data acquisition and processing

2.2.1. Diffusion MRI

All three *ex vivo* samples were scanned in a small-bore 9.4T Bruker Biospec system with gradients capable of $|G_{max}|$=480 mT/m. Prior to scanning, each block was placed in a plastic syringe filled with Fomblin and all air was removed. The dMRI data were acquired using a spin echo 3D single-shot EPI sequence with a maximum gradient amplitude of 393 mT/m, TR=750 ms, TE=43 ms, GRAPPA factor 2, matrix size 136x136x176, and 250 µm isotropic resolution. We used a DSI sampling scheme consisting of one b=0 image and 514 gradient directions arranged on a Cartesian lattice in q-space and zero padded to an 11x11x11 grid (Wedeen et al., 2005). Diffusion encoding was applied with b_{max}=40000 s/mm^2, δ=15 ms, and Δ=21 ms, corresponding to q_{max}=250 mm^{-1}. The total acquisition time was approximately 48 hours. We will refer to the datasets containing all 515 diffusion volumes as the "fully sampled" or "FS" data.

A 4-channel phased array surface receive (Rx) coil was used in dMRI acquisitions (diagramed in Figure 2A, bottom), leading to a decrease in coil sensitivity and SNR as the distance between the sample and Rx coil increased. The coronal planes of the dMRI data were approximately parallel to the surface coil, so coronal slices closer to the coil had higher signal than the slices farther away. Figure 2A (top) shows three example coronal slices from the b=0

volume of sample 1A, each with varying distances from the surface Rx coil. To quantify the dMRI SNR, we used the b=0 signal intensities from a region of interest (ROI) in the deep WM. The SNR of each coronal slice was calculated as the mean signal (Figure 2B, solid black line) divided by the standard deviation of the signal intensities from the ROI (Figure 2B, solid gray line). In order to relate the SNR across slices to the sensitivity profile of the coil, we fit a linear model to the calculated SNR values to obtain an expression for SNR as a function of slice (Figure 2B, dotted green line). For the dMRI slice numbering of each sample, we will refer to slice 1 as the slice closest to the surface coil, with increasing slice numbers indicating increased distance from the coil.

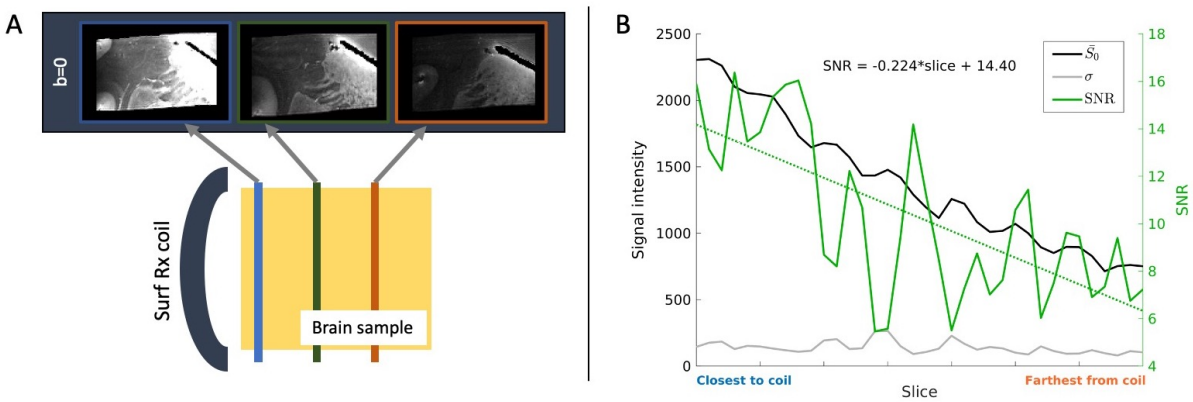

Figure 2. dMRI acquisition and SNR. (A) Bottom: A surface receive coil was used in dMRI acquisitions. Top: This led to a decrease in sensitivity with increasing distance from the coil, as shown in coronal slices from the sample 1A b=0 volume. The same intensity scaling was used for all three slices. **(B)** The b=0 SNR in each coronal slice of sample 1A (solid green line) was calculated from the mean (black line) and standard deviation (gray line) of signal intensities from a deep WM ROI. A linear regression was performed to find the slope of SNR as a function of slice (dotted green line). The left y-axis shows the signal and noise intensities, and the right y-axis shows SNR values. The x-axis in indexed by slice number.

2.2.2. PSOCT

Following dMRI acquisition of sample 1A, a piece of the tissue block was extracted for imaging with PSOCT (Figure 1A, right). Details on the setup, acquisition, processing and analysis of the PSOCT dataset used in this paper were previously described in (Jones et al., 2020). Briefly, the sample was imaged with a polarization maintaining fiber (PMF) based, spectral domain PSOCT system developed in-house (Wang et al., 2016). The light source consisted of two broadband super-luminescent diodes (Thorlabs Inc., LSC2000C), with a center wavelength of 1300 nm and a bandwidth of 170 nm. The axial resolution was 2.9 µm in tissue and the lateral resolution was estimated at 3.5 µm. PSOCT produces measurements of optic axis orientation, which represent the in-plane (2D) orientation of the fiber axis. We downsampled the PSOCT data to an in-plane resolution of 10 µm to facilitate processing and analysis. To cover the entire tissue surface, 1120 tiles (FOV = 1 mm^2) were acquired using a 50% overlap and a snaked configuration scheme. The tiles were stitched using the Fiji software (Preibisch et al., 2009; Schindelin et al., 2012). A vibratome cut off a 75 µm slice after the superficial region of the tissue block was imaged, which consequently allowed deeper regions to be exposed by PSOCT. There were 63 total slices acquired for the sample block. One critical advantage of the technique is that PSOCT images the blockface of the tissue before slicing (Wang et al., 2018). This avoids

the nonlinear inter-slice distortions that are present in traditional histological techniques, where slices are imaged after they are cut. As a result, we can simply stack the slices into a volume.

Affine registration was used to align the dMRI fractional anisotropy (FA) and PSOCT retardance volumes. Retardance represents the phase delay between orthogonal polarization channels that is induced by birefringence, a property of anisotropic structures. The myelinated axons that compose WM bundles possess birefringent properties and are highlighted in the retardance, providing a similar contrast to FA. We used a robust, inverse consistent registration method that detects and down-weighs outlier regions in the images (Reuter et al., 2010). The dMRI volumes were transformed to PSOCT space using nearest-neighbor interpolation, and the dMRI orientation vectors (see section 2.3.2 below) were also rotated accordingly using the rotational component of the affine transformation. The transformed dMRI vectors were projected onto the PSOCT imaging plane for the purposes of comparison to optic axis measurements, which represent the in-plane fiber orientations.

2.3. dMRI reconstructions

2.3.1. Compressed sensing

CS reconstructions were performed using two dictionary-based CS-DSI methods previously introduced by Bilgic et al. (2013). One is PCA-based reconstruction (PCA) and the other is Tikhonov-regularized pseudoinverse reconstruction using the training set of PDFs as the dictionary (PINV). CS undersampling masks were generated for acceleration factors R=3, 5 and 9 using a variable-density power-law function (Bilgic et al., 2013; Lustig et al., 2007). Nine CS masks with different sampling patterns were created for each acceleration factor. Retrospective undersampling was applied to fully sampled q-space data, followed by CS reconstruction using either PCA or PINV. Each reconstruction method had one free parameter: the number of principal components T and the Tikhonov regularization parameter λ, for PCA and PINV respectively. The optimal parameter for each dictionary was determined using the parameter sweeping approach from Bilgic et al. (2013), in which the training data were undersampled and reconstructed using a range of parameters, and the one yielding the lowest root-mean-square error (RMSE) in PDFs when compared to the fully sampled data was selected. This procedure was performed separately for each combination of dictionary and acceleration factor.

Reconstructions and analyses were performed in MATLAB (R2019a, 9.6) based on publicly available code (https://www.martinos.org/~berkin/software.html). Both PCA and PINV methods used a single matrix multiplication to reconstruct an entire slice, with computation times of ~10-15 seconds per slice (~6000 voxels per slice) on a workstation with a 3.4GHz Intel i7 processor, 8 cores, 32GB RAM. For full volume reconstructions (~50 slices), we compiled the Matlab function into a standalone application and processed individual slices in parallel on a high-performance compute cluster. Simultaneously running each slice as a separate process with 8GB of memory resulted in full volume reconstruction times between 2-10 minutes (depending on the number of available CPUs).

2.3.2. Fiber orientations

At each dMRI voxel, we computed the orientation distribution function (ODF), *i.e.,* the marginal PDF of the diffusion orientation angle. Diffusion tractography algorithms use the ODFs, rather than the full PDFs, to reconstruct WM bundles. In DSI, ODFs are typically obtained by interpolating the PDFs onto uniform radial vertices and summing them along each radial projection. The truncation of q-space causes ringing in the PDFs, which introduces

artifacts into the ODFs. One approach to mitigating these artifacts is to apply a windowing function on the q-space data, like a Hanning filter (Wedeen et al., 2005). This smooths the signal decay at the edges of q-space but diminishes the contributions of high-frequency diffusion terms, potentially oversmoothing the PDFs and ODFs and reducing angular resolution. An alternative approach is to use unfiltered q-space data and carefully define the starting and ending displacement distances for integration of the PDF (Lacerda et al., 2016; Paquette et al., 2016; Tian et al., 2016). This way, one can restrict the integration range so that PDF ringing is omitted from ODF computations without having to taper the high frequency q-space data. We used the latter approach for ODF reconstructions.

DSI ODF reconstructions were performed in Python with the *Dipy* (version 1.3.0.) library (Garyfallidis et al., 2014). For each voxel, the q-space data were zero padded to a 75x75x75 grid and the 3D FFT was applied to obtain the PDF, with negative PDF values clipped to zero. For ODF reconstruction, we used a PDF integration lower bound of $r_{start} = 10$ and upper bound of $r_{end} = 16$ (which were roughly 0.27x and 0.43xFOV), a radial step size of $r_{step} = 0.05$, and a radial weighting factor of 2. ODF peaks were extracted using a maximum of 3 peaks per voxel, minimum peak separation angle of 25° and minimum peak magnitude of 0.05. The same parameters were used for all DSI ODF reconstructions.

We also fit the DTI model to the fully sampled data using the FSL command *dtifit* and extracted the orientations of the primary eigenvectors of the tensors. These served as a baseline for "worst-case" dMRI orientation accuracy, *i.e.,* where only a single fiber population can be resolved in each voxel. We will refer to the fully sampled DTI and DSI orientations as FS-DTI and FS-DSI, respectively.

2.4. Error metrics

We performed a voxel-wise comparison of dMRI and PSOCT orientations using absolute angular error as the accuracy metric, as in (Jones et al., 2020). In each voxel, we selected the peak of the diffusion ODF with the in-plane orientation that matched the corresponding PSOCT orientation most closely and computed the angular error between that peak and the PSOCT measurement. Thus, for a voxel with *N* dMRI ODF peaks, the absolute in-plane angular error *AE* was calculated as:

$$AE = \min_{n}\{\arccos\left(|\boldsymbol{e}_n^T \boldsymbol{v}|\right)\}, \tag{1}$$

where $\boldsymbol{e}_n$ is the unit vector along the nth dMRI peak orientation, and $\boldsymbol{v}$ is the unit vector along the measured PSOCT orientation.

PSOCT optic axis measurements rely on birefringence of anisotropic processes, such as axon bundles in WM, but are not necessarily accurate in gray matter regions where there is a lack of birefringence and low retardance. Thus, a WM mask was created to exclude all voxels where the retardance intensity was below 50% of the maximum retardance. Only voxels in this WM mask were considered when computing angular error metrics. The PSOCT imaging plane was nearly parallel to the dMRI coronal plane, which itself was nearly parallel to the surface Rx coil. We exploited this arrangement, and the decreasing SNR of dMRI slices at increasing distance from the coil, to assess the accuracy of CS-DSI as a function of SNR.

As a more conventional CS-DSI error metric, we also computed the normalized RMSE of the PDFs obtained from CS reconstruction with respect to those obtained from the fully sampled DSI data:

$$RMSE = \frac{\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2}{\|\boldsymbol{x}\|_2}, \tag{2}$$

where $\boldsymbol{x}$ is the PDF from the fully sampled data, $\hat{\boldsymbol{x}}$ is the PDF from the CS reconstructed data, and $\|\cdot\|_2$ is the L2-norm.

2.5. CS-DSI validation experiments

We constructed a total of six PDF dictionaries, each trained on a single slice of fully sampled DSI data from sample 1A, sample 1B, or sample 2. For each sample we created two dictionaries, one from a high-SNR slice (slice 3) and one from a low-SNR slice (slice 13). We applied CS reconstruction to DSI data from sample 1A, undersampled by a factor of R=3, 5, or 9, using one of these dictionaries. We then computed DSI ODFs and extracted the orientations of the ODF peaks. We transformed the orientations to PSOCT space and projected them onto the PSOCT plane. We calculated the absolute angular error with respect to PSOCT at each WM voxel.

Angular non-uniformities in the CS undersampling patterns may introduce directional biases into the ODFs, and thus affect our angular error computations. We accounted for this potential source of variability in our error metrics by repeating the CS reconstructions with 9 different CS undersampling masks, for each combination of dictionary and acceleration factor.

2.5.1. Effect of CS acceleration factor and training sample on angular error

We assessed the efficacy of the CS algorithms at different acceleration factors, by comparing CS reconstructions of data from sample 1A that had been undersampled by a factor of R=3, 5, and 9. In this comparison, we used the PCA and PINV methods with dictionaries trained on a high-SNR slice from each sample. The accuracy of dMRI orientations was quantified by the mean angular error across each PSOCT slice, as well as across all WM voxels in the PSOCT volume. This error was averaged over the reconstructions that were obtained with the 9 different CS undersampling masks. The mean angular error of FS-DSI was used as a reference for evaluating the quality of CS reconstructions.

2.5.2. Effect of SNR on reconstruction error metrics

One goal of this study was to determine the influence of SNR on metrics of CS reconstruction quality. To this end, we calculated the b=0 SNR (as described in section 2.2.1.) for each dMRI slice and computed the average RMSE in PDFs (with respect to the fully sampled data) and the average angular error (with respect to PSOCT) across all WM voxels in each slice. Then, for each CS reconstruction, we performed a linear regression of RMSE or angular error against SNR. This was done for FS-DTI, FS-DSI, and CS-DSI. For CS-DSI, error metrics from each combination of acceleration, training sample, and method were calculated, each time averaging the mean errors over CS reconstructions from the 9 different undersampling masks.

2.5.3. Effect of SNR on dictionary training

To evaluate the effect that the SNR of the training data has on CS reconstructions, voxels in sample 1A were reconstructed at acceleration R=3 with dictionaries trained on low-SNR data

(slice 13) from each sample using PCA and PINV methods. These results were then compared to the corresponding results from the experiment described in section 2.5.1. that used high-SNR training data. The accuracy of dMRI orientations was quantified by the mean angular error across all WM voxels in the PSOCT volume. The CS reconstructions included here were performed with one randomly selected CS undersampling mask.

3. Results

3.1. Visual inspection

Figure 3 provides visualizations of dMRI and PSOCT orientations from a representative slice of sample 1A. Figure 3A shows fiber orientation maps as color-coded RGB images. The color wheel shows the correspondence between pixel color and in-plane orientation. For fully sampled DSI and for CS-DSI at acceleration factors R=3, 5, 9, the color maps show the orientations of the ODF peaks that most closely matched the PSOCT orientations in the same voxel. All CS-DSI results are shown for the same CS undersampling mask. For fully sampled DTI, the color maps show the orientations of the primary eigenvector of the diffusion tensor. Figure 3B shows heat maps of the absolute angular error between dMRI and PSOCT orientations in each voxel.

Despite the large disparity in voxel size, there was good overall agreement between dMRI and PSOCT fiber orientation maps (Figure 3A). The dMRI maps showed the closest resemblance to PSOCT in the medial half of the slice, with greater differences in the lateral half of the slice. Examination of the angular error maps (Figure 3B) confirms that the greatest angular errors occurred in the middle and lateral regions of the slice. This distribution of errors was most obvious in the FS-DTI error map (Figure 3B, top right). The fully sampled (Figure 3V, top middle) and CS-DSI maps (Figure 3B, bottom) show similarly good agreement with PSOCT.

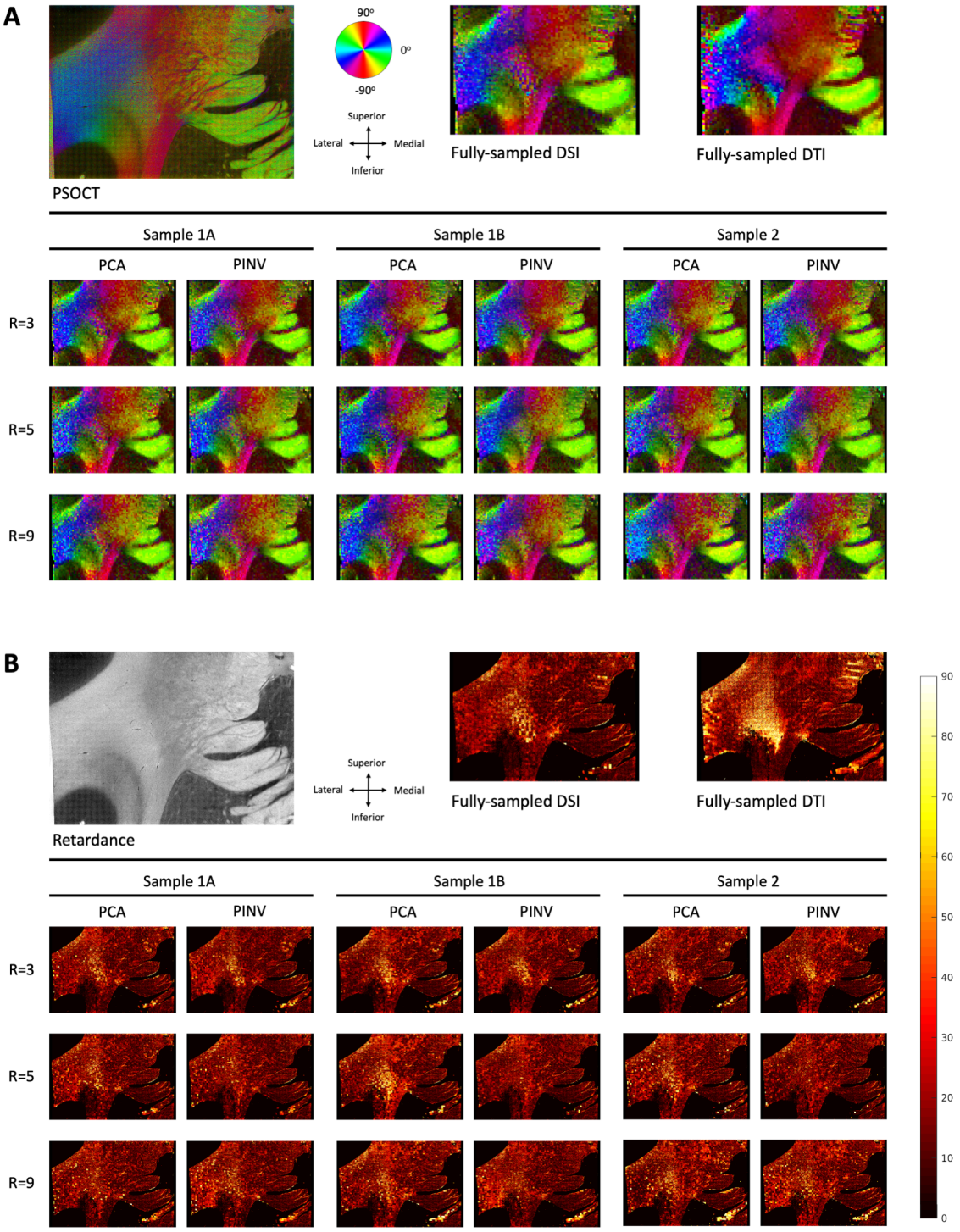

Figure 3. Fiber orientations estimated from dMRI vs. PSOCT. (A) Color-coded maps of PSOCT (top left) and dMRI orientations from fully sampled data (top right) and CS-DSI data (bottom). **(B)** Absolute angular error of dMRI orientations with respect to PSOCT. A WM mask was created by thresholding the PSOCT retardance (top left). The heat maps were masked to include only voxels classified as WM. CS-DSI reconstructions are shown for high-SNR training data and one of the CS undersampling masks.

3.2. Effect of CS acceleration factor and training sample on angular error

Figure 4 shows bar plots of the mean angular errors of dMRI with respect to PSOCT, averaged over all WM voxels analyzed, from CS reconstructions of sample 1A with high-SNR training data at different acceleration factors. Error bars show the standard error over different CS undersampling masks. The corresponding statistics are given in Table 2. The mean angular error of FS-DSI (green bar and dotted line) is shown as the benchmark for assessing CS-DSI results. The mean angular error of FS-DTI (red bar and dotted line) is shown as a worst-case scenario.

At an acceleration factor of R=3, CS-DSI achieved very similar angular error to FS-DSI (within +/-1.27 º), for both PCA and PINV reconstruction methods, and regardless of whether the training data came from the same or a different sample than the test data. For PCA reconstruction, the angular error increased with the acceleration factor. This increase was most dramatic in the (more realistic) scenario where the training data came from a different brain than the test data. Conversely, higher acceleration factors imparted only minor changes on the accuracy of PINV reconstructions.

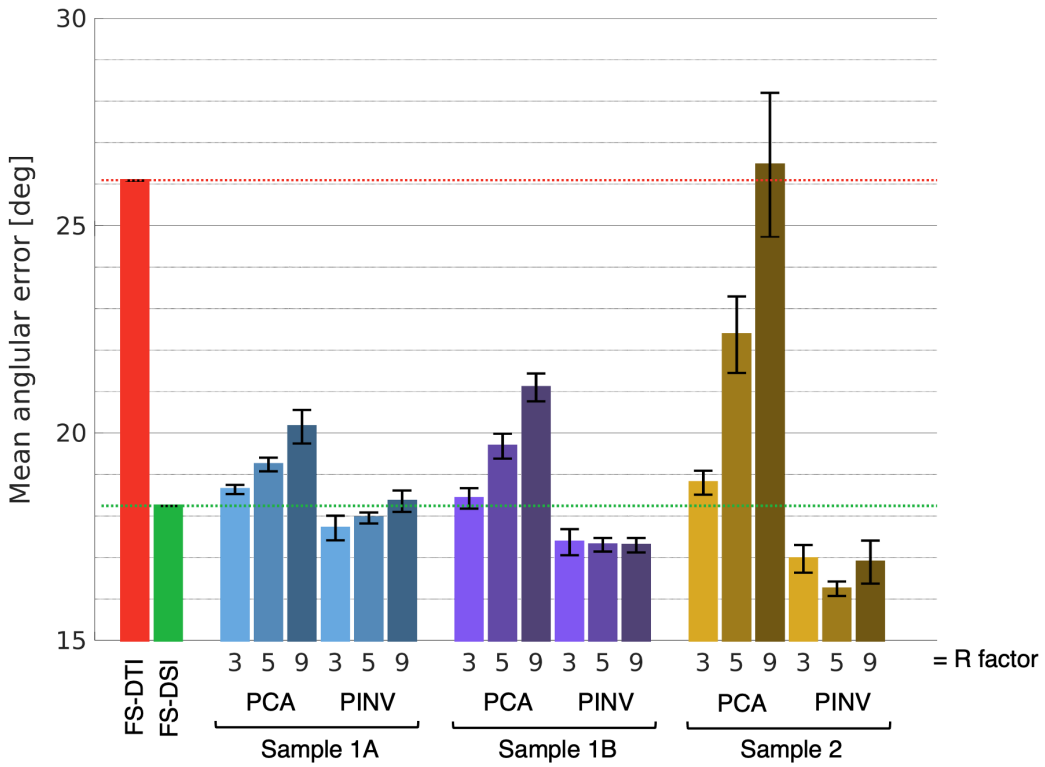

Figure 4. Mean angular error of dMRI with respect to PSOCT. The plots show average error over all WM voxels in sample 1A. For CS-DSI, the error was also averaged over 9 CS undersampling masks, with error bars representing standard error of the mean across these 9 masks. Results are grouped by training sample (sample 1A, blue; sample 1B, purple; sample 2, yellow) and CS method. Bar shades correspond to acceleration factor (3, 5, 9). All CS-DSI reconstructions used high-SNR training data from each sample. Mean angular errors from FS-DTI (red) and FS-DSI (green) are shown on the far left for comparison.

	Angular error (degrees)	
FS-DTI	26.10	
FS-DSI	18.25	
	PCA	**PINV**
Training sample 1A		
R=3	18.64 ± 0.11	17.71 ± 0.30
R=5	19.25 ± 0.16	17.95 ± 0.14
R=9	20.16 ± 0.40	18.36 ± 0.26
Training sample 1B		
R=3	18.43 ± 0.25	17.37 ± 0.32
R=5	19.69 ± 0.30	17.31 ± 0.16
R=9	21.10 ± 0.33	17.30 ± 0.17
Training sample 2		
R=3	18.81 ± 0.29	16.97 ± 0.33
R=5	22.37 ± 0.92	16.25 ± 0.18
R=9	26.48 ± 1.73	16.89 ± 0.52

Table 2. Angular error of dMRI orientations. Mean angular errors of dMRI with respect to PSOCT across all analyzed WM voxels. CS-DSI reconstructions used high-SNR training data from each sample. For CS-DSI, standard errors of the mean are also shown, computed over the 9 CS undersampling masks.

We delved deeper into this difference between the performance of PCA and PINV by comparing how many peaks were detected in the ODFs from the data reconstructed by each method. The bar plot in Figure 5A (left) shows the average number of reconstructed ODF peaks per WM voxel from each reconstruction (for a maximum of 3 peaks per voxel), with standard error bars. CS-DSI results were averaged over the 9 undersampling masks used for CS reconstructions. The number of ODF peaks reconstructed from FS-DSI (green bar) is shown for reference.

FS-DSI produced an average of approximately 2 peaks per voxel. For an acceleration factor of R=3, PCA reconstructions produced a similar number (less than 3.2% difference vs. FS-DSI). As the acceleration factor increased, PCA tended to return slightly fewer peaks. On the other hand, PINV reconstructions returned a greater number of peaks than FS-DSI (greater than 7.6% difference vs. FS-DSI). This number increased as the acceleration factor increased, and when the training data came from a different sample than the test data.

Figure 5B plots the number of ODF peaks against the mean angular error. The PCA and PINV reconstructions are denoted by triangle and circle markers, respectively, and colored as in Figure 5A. Error bars indicate the standard error for each metric. FS-DSI (green diamond) served as the benchmark in terms of angular error. As indicated by the red dashed line in Figure 5B, there was a clear separation between PINV (above the line) and PCA (below the line). FS-DSI was situated at the knee of the curve, and PCA reconstruction with an acceleration factor of R=3 was closest to FS-DSI, regardless of the sample that was used as the training data set. Thus, we conclude that CS-DSI with a combination of R=3 acceleration and PCA reconstruction preserved the accuracy of FS-DSI with respect to the ground truth orientations, but without having to introduce additional peaks.

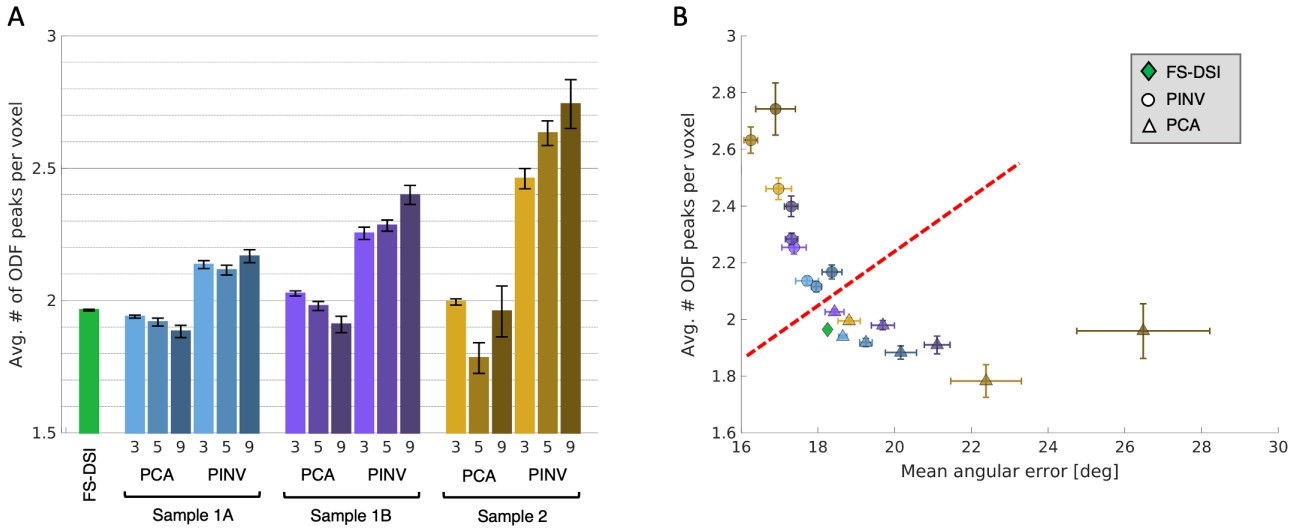

Figure 5. Number of reconstructed ODF peaks. (A) Average number of ODF peaks per WM voxel from fully sampled (green bar) and CS reconstructed (blue, purple, and brown bars) DSI data. CS results show the average from the 9 q-space volumes reconstructed with different CS undersampling masks. Error bars display the standard error. **(B)** Plot of the number of ODF peaks as a function of the mean angular error from fully sampled (green diamond) and CS reconstructed q-space data. Circle markers correspond to PINV, and triangle markers correspond to PCA. Marker colors are the same as the bar plots in (A). Error bars indicate the standard error. PINV reconstructions (above red dashed line) produced more peaks and lower angular errors than PCA (below red dashed line). The CS-DSI reconstruction that was closest to FS-DSI on both axes was PCA with acceleration R=3.

3.3. Effect of SNR on reconstruction error metrics

Figure 6A shows the slice-wise mean angular errors with respect to PSOCT, plotted against the SNR of the corresponding slice. Error bars indicate the standard error. The CS reconstructions included here used acceleration R=3 and dictionaries trained on high-SNR data from each sample. The curves for CS-DSI reconstructions closely resembled that of FS-DSI (green line) throughout the entire volume, with less than 2.52° variation in mean angular errors between all DSI reconstructions in each slice. FS-DTI (red line) had significantly greater errors than both FS-DSI and CS-DSI and showed a sharp increase in error as SNR decreased. The DSI angular errors were relatively robust to decreased SNR.

Table 3 shows statistics from the linear regressions of the mean angular error against SNR. The slopes were noticeably flatter for FS-DSI (-0.24° per unit SNR, $p=0.051$) and CS-DSI reconstructions (-0.28° to -0.46° per unit SNR, $p<0.003$) than FS-DTI (-1.20° per unit SNR, $p=0.00011$).

For comparison, Figure 6B shows the slice-wise mean RMSE in PDFs between each CS-DSI reconstruction and the fully sampled DSI data, as a function of SNR. The mean RMSE was calculated from the real part of the diffusion PDFs and averaged across all WM voxels in each slice. Error bars depict the standard error. Statistics from the linear regressions of the RMSE versus SNR are also given in Table 3. All reconstructions exhibited a strong negative correlation between RMSE and SNR ($r=0.99$, $p<0.001$) and showed a nearly linear increase in RMSE as SNR decreased (Figure 6B), with consistent linear regression slopes (-0.66 to -0.96 % RMSE per unit SNR). It should be noted that the linear fit of SNR likely smoothed the observed relationship, but that notwithstanding, the correlation between SNR and CS RMSE remained markedly apparent across all reconstructions.

These results suggest that, although the error in the PDFs obtained from CS-DSI increases noticeably as the SNR decreases, this translates to only a minor increase of the error in the peak orientations, which is what would have an impact on tractography. Note that these trends were observed both for reconstructions that tended to add more diffusion peaks (PINV) and for those that did not (PCA), hence this relationship between each error metric and SNR did not appear to be explained by the number of peaks.

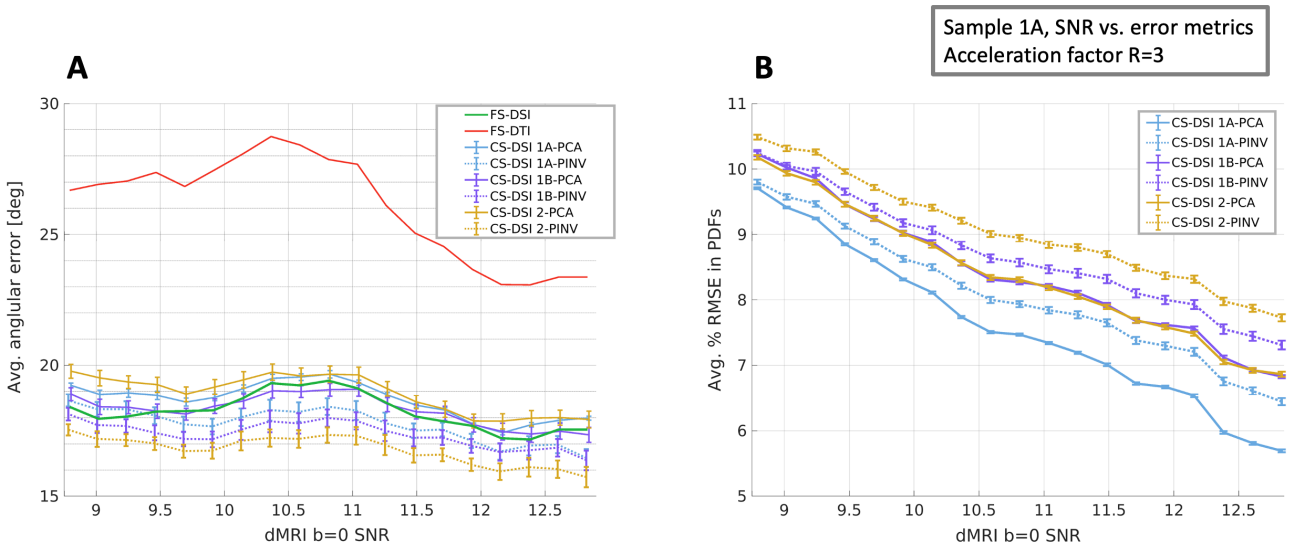

Figure 6. Reconstruction error as a function of SNR. (A) Average angular error of FS-DSI, FS-DTI, and CS-DSI with respect to ground-truth axonal orientations from PSOCT. **(B)** Average RMSE in PDFs between CS-DSI and FS-DSI. Each error metric is averaged across all WM voxels in each slice and plotted against the SNR of the corresponding b=0 dMRI slice. Line colors and styles correspond to different dMRI reconstructions. For CS reconstructions, each error metric was averaged across 9 CS undersampling masks, with error bars showing the standard error of the mean across the 9 masks. For CS-DSI, line colors denote different training samples and line styles denote different CS reconstruction methods. All CS reconstructions used an acceleration factor of R=3 and high-SNR training data.

		Angular error			RMSE in PDFs		
		r	Slope (° per unit SNR	*p*-value	r	Slope (% per unit SNR)	*p*-value
FS-DSI		-0.45	-0.25	0.051	-	-	-
FS-DTI		-0.77	-1.21	0.00011	-	-	-
CS training	CS method						
Sample 1A	PCA	-0.67	-0.36	0.0017	-0.99	-0.96	4.35×10^{-17}
	PINV	-0.84	-0.43	7.0×10^{-6}	-0.99	-0.79	9.65×10^{-17}
Sample 1B	PCA	-0.65	-0.30	0.0028	-0.99	-0.81	3.20×10^{-16}
	PINV	-0.73	-0.28	0.00038	-0.99	-0.70	1.82×10^{-16}
Sample 2	PCA	-0.81	-0.46	2.6×10^{-5}	-0.99	-0.80	1.11×10^{-17}
	PINV	-0.81	-0.35	2.4×10^{-5}	-0.99	-0.66	6.86×10^{-17}

Table 3. Linear regressions of reconstruction error metrics vs. SNR. Linear correlation coefficient (r), slope, and p-value from the linear regressions of the angular error with respect to PSOCT and the RMSE in PDFs with respect to fully sampled data as a function of SNR. CS reconstructions used an acceleration factor of R=3 and high-SNR training data.

3.4. Effect of SNR on dictionary training

After observing that CS reconstructions at acceleration R=3 using high-SNR training data performed nearly as well as FS-DSI in terms of angular error, we tested whether this was also true with low-SNR training data. The bar plot in Figure 7 compares the effect of training data SNR on the mean angular error of CS reconstructions at acceleration R=3 across all WM voxels analyzed. The CS reconstructions included here used one of the R=3 undersampling masks (out of the 9 undersampling masks used for Figures 4-6). All CS reconstructions exhibited greater mean angular error when using low-SNR than high-SNR training data, although the extent of differences varied depending on the training sample and reconstruction method. Specifically, PCA was more sensitive to the SNR level of the training data than PINV.

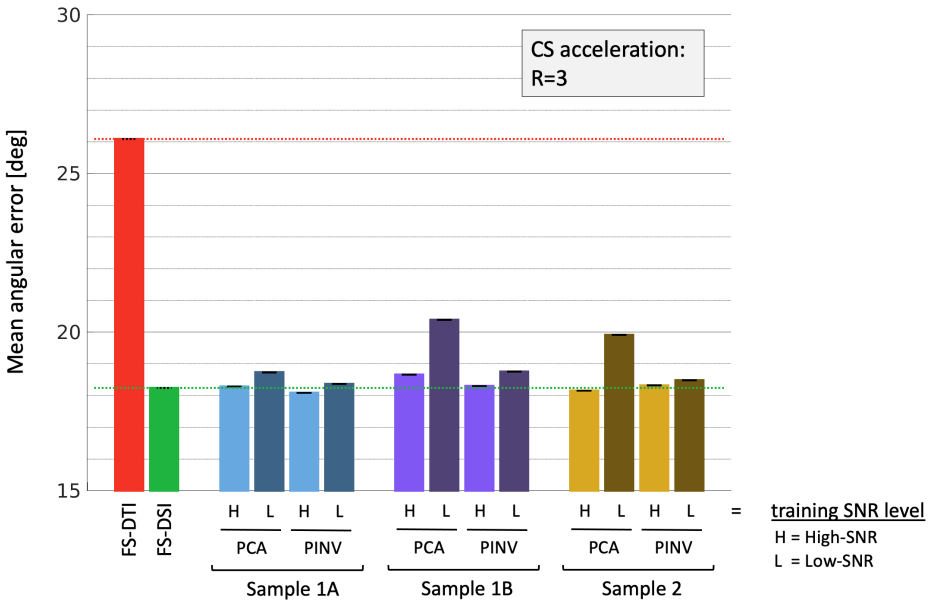

Figure 7. Effect of the SNR of the training data on angular error. Mean angular error across all analyzed WM voxels for CS reconstructions at acceleration R=3. Dictionaries were trained using either high-SNR ("H", light shade) or low-SNR ("L", dark shade) slices from sample 1A (blue), sample 1B (purple), or sample 2 (yellow). Results from FS-DTI (red) and FS-DSI (green) are shown on the far left. Error bars show standard error across voxels.

4. Discussion

In this study we evaluated the accuracy of diffusion orientation estimates obtained from CS-DSI by comparing them to ground-truth measurements of fiber orientations from optical imaging. This complements previous work that used simulated or fully sampled DSI data as the ground truth (Bilgic et al., 2013; Menzel et al., 2011; Paquette et al., 2015). While the error between PDFs obtained with fully sampled and CS-DSI is informative, it is not clear what error can be tolerated without incurring a noticeable loss of accuracy with respect to the underlying fiber orientations. The benefit of using PSOCT as the reference modality is that it provides direct measurements of axonal orientations that are independent of water diffusion.

We employed two dictionary-based algorithms for CS reconstruction and investigated the effects of training data and undersampling rate on the accuracy of estimated fiber orientations. We found that a PCA reconstruction at an acceleration factor of R=3 retained the fiber orientation accuracy of the fully sampled DSI data (Figure 4), without introducing additional peaks (Figure 5). Its performance was as good when the training data for the dictionary came from a different sample than the test data as it was when the training and test data came from the same sample. However, we also found that it was important to use high-SNR training data to achieve this performance (Figure 7).

We conclude that PCA reconstruction of CS-DSI data can be a reliable approach for achieving 3-fold acceleration of DSI without degrading its accuracy. This acceleration factor allows the equivalent of 514-direction DSI to be reconstructed from a 171-direction acquisition, thus bringing DSI acquisition times in line with those of state-of-the-art high angular resolution (*e.g.,* multi-shell) acquisitions. As has been shown in previous validation studies, the DSI acquisition scheme achieves high accuracy (Daducci et al., 2013, Jones et al., 2020, Maffei et al., 2020), and DSI data can be readily resampled onto q-shells to facilitate any analysis that requires a shell-based acquisition (Jones et al., 2020).

We exploited the sensitivity profile of the surface Rx coil used in our *ex vivo* dMRI acquisitions to quantify the accuracy of CS-DSI as a function of SNR in the test data. We found that, as SNR decreased, the RMSE between the PDFs obtained from CS-DSI and those obtained from fully sampled DSI data increased faster than the angular error of the peak orientations with respect to PSOCT (Figure 6). This suggests that there is a range of errors in the PDF that we can tolerate without impacting the accuracy of the peak orientations. Specifically, as seen in Figure 7, a roughly 40% increase in the PDF-based RMSE due to loss of SNR leads to at most a 1° increase in the angular error with respect to the ground-truth axonal orientation measurements. This increase is very small compared to the overall errors of dMRI orientations with respect to the ground-truth measurements.

In regard to the dictionary-based CS algorithms that we investigated, our findings indicate that dictionary generalizability differs between the PINV and PCA methods, and that the extent of these differences depends on both the CS acceleration factor (Figure 4) and the training data SNR (Figure 7). One potential factor contributing to these differences may be the ways in which PDFs are represented within the algorithms. PCA performs reconstructions in a reduced-dimensionality space consisting of only the T principle components that describe the greatest variance in the training PDFs, and the optimal number of components T decreases at higher acceleration factors in order to improve the conditioning of the pseudoinverse in the least-squares reconstruction (Bilgic et al., 2013). In our experiments, T was around 20 at acceleration R=3 and less than that at higher accelerations, which are moderately lower than the optimal *in vivo* parameters. Intuitively, reducing the number of principal components subsequently limits the ability to describe finer scale details in PDFs, and, together with the extremely undersampled q-space data used at high accelerations, likely hinders the level of detail in such reconstructions. In contrast, PINV operates directly on the training PDFs themselves, exploiting the prior information encoded in the dictionary atoms to bypass sparsity constraints. Together, these differences may contribute to the different behavior of PINV and PCA as the acceleration factor increases.

The PINV method investigated here uses a dictionary containing PDFs from a slice of fully sampled training data, without any further training. An alternative approach is to use a dictionary trained with the K-SVD algorithm (Aharon et al., 2006), which enhances the sparsity

level of PDF representations and is a fraction of the size of the PINV dictionary, allowing up to a 50% reduction in computation time. Regardless of the dictionary, reconstructions are performed using the Tikhonov-regularized pseudoinverse. Previous comparisons between PINV using a 3191-column dictionary and PINV(K-SVD) using a 258-column dictionary reported nearly identical reconstruction quality in terms of RMSE, as well as equivalent representational power between the two dictionaries (Bilgic et al., 2013). Although we did not include results from PINV(K-SVD) here, we did perform CS reconstructions and PSOCT angular error analysis using PINV(K-SVD) and observed very similar results to PINV, both in terms of RMSE with respect to fully sampled data and in terms of angular error with respect to PSOCT. Future work will investigate potential benefits of deep and nonlinear representations of PDFs that can exploit similarities present in image- and q-space domains.

Relation to previous studies

We have previously used the present PSOCT analysis framework (and sample 1A from this work) in an extensive validation study (Jones et al., 2020). In that work, we assessed the accuracy of fiber orientations estimated from various dMRI orientation reconstruction methods and sampling schemes, including DSI, single- and multi-shell. The DSI results are somewhat different between the two studies because the DSI ODF reconstruction is different. In (Jones et al., 2020), DSI reconstructions were performed with the DSI Studio toolbox (http://dsi-studio.labsolver.org), and used filtered q-space signals (Hanning window, width=16) with default parameters (*e.g.*, zero-padded 16x16x16 q-space grid, ODF integration lower/upper bounds of 0.25x and 0.75x FOV), and a ODF peak threshold of 0.01 (1%). Here, we used a different approach for ODF reconstruction, imposed a (slightly) more stringent peak threshold (5%) and included a peak separation threshold. These factors likely contributed to the ~1-2° increases in mean angular error reported here compared to previous results. Nonetheless, the angular errors in this work were similar to the best performing reconstructions in (Jones et al., 2020), namely DSI and GQI with the fully sampled DSI data and Q-ball with single- and multi-shell data, which were between ~17° and ~19°.

This work analyzed the dictionary-based CS-DSI methods introduced by Bilgic et. al. (2013), however there were several technical differences between these works that should be noted. First, in terms of data acquisition, we used DSI data from *ex vivo* human brain samples acquired at 9.4T with a 4-channel surface Rx coil, whereas Bilgic et. al. used *in vivo* DSI from the 3T Connectom system with a custom 64-channel head coil (Keil et al., 2013). These differences did not have an apparent effect on the dictionary-based CS-DSI methods. First, the CS algorithms yielded similar results when using *ex vivo* training and test data as when using *in vivo* training and test data. Whether the same would also be true when using *ex vivo* training data and *in vivo* test data (or vice versa) has yet to be investigated. Such an approach may be of interest as long, *ex vivo* acquisitions can be a way to collect very high-SNR training data. Second, while both studies generated dictionaries with PDFs from a single slice of fully sampled data, a single slice of our *ex vivo* samples covers only a small anatomical region, whereas a slice of *in vivo* data covers an entire cross-section of the brain. Given that our training samples were cut from different anatomical locations, one might expect that local microstructural differences between training and test samples might pose challenges for dictionary generalizability. However, our findings showed that high-quality reconstructions could be obtained using training and test data from different samples, indicating that our *ex vivo* dictionaries possess the representational power to generalize across samples. Indeed, the "residual" (Bilgic et al., 2013)

between *ex vivo* PDF dictionaries, *i.e.*, the energy of the part of one dictionary that cannot be represented by another, was negligibly small ($\sim 10^{-12}$), confirming that dictionaries from different samples possess equivalent representational power.

Limitations

We investigated two dictionary-based CS-DSI reconstruction methods that use discrete EAP representations and L2-regularized algorithms. However, there are various other CS-DSI methodologies that utilize other basis functions (*e.g.*, discrete cosine transform, discrete wavelet transform) or approaches to solving the underdetermined CS problem (*e.g.*, L1-regularized methods such as equality constrained or regularized Dictionary-FOCUSS). While the L2 dictionary-based CS-DSI investigated here was shown to provide comparable reconstruction quality to both fixed transforms and iterative L1 approaches (Bilgic et al., 2013), those evaluations were mostly based on RMSE with respect to fully sampled DSI data. Analyzing the fidelity of other CS-DSI algorithms using PSOCT would be of interest, however the lengthy computation times required of methods using iterative algorithms or fixed transforms would ultimately limit their utility.

This work focused primarily on validating the accuracy of CS-DSI fiber orientations extracted from ODFs. While ODFs are commonly used structures essential for tractography, they represent only a portion of the abundant information contained in diffusion PDFs. Diffusion PDFs can provide detailed descriptions of tissue microstructure, and have been previously used to characterize age-related WM demyelination (Fatima et al., 2013), delineate pathological tissue lesions in patients with multiple sclerosis (Assaf et al., 2002), and map *in vivo* axon caliber in the human brain (Hori et al., 2016). Future work will aim to validate the microstructural parameters (*e.g.*, return to origin probability, mean-squared displacement, kurtosis) derived from CS-DSI reconstructions.

It is important to recognize that PSOCT is just one of a number of techniques used to obtain reference measurements of microscopic axonal orientations for subsequent comparison to dMRI. Other approaches include extracting orientations from myelin-stained sections (Choe et al., 2012; Leergaard et al., 2010; Schilling et al., 2017; Seehaus et al., 2015) or from confocal microscopy of slices stained with DiI, a fluorescent dye (Budde & Frank, 2012). Quantification of 3D orientations has been reported with Dil stained slices (Khan et al., 2015; Schilling et al., 2016; Schilling et al., 2018). The major obstacle for these histological methods is that they rely on manual tracing or image processing to extract orientation estimates. Optical imaging techniques based on light polarization are able to provide direct measurements of fiber orientations by exploiting the intrinsic birefringence property of the tissue. Polarized light imaging (PLI) is another technique that uses birefringence to measure axonal orientations (Axer et al., 2011; Henssen et al., 2019; Mollink et al., 2017). However, unlike PSOCT, PLI requires tissue to be sectioned and mounted before imaging. This can lead to severe tissue distortions that demand a complex registration framework to correct (Ali et al., 2017; Ali et al., 2018; Majka & Wójcik, 2016). PSOCT images the tissue before slicing, greatly reducing tissue distortions and allowing accurate volumetric reconstructions.

As is the case with all validation techniques, there are also various limitations to PSOCT. A notable limitation is that the optic axis orientation measurements do not describe the 3D orientation, but rather its projection onto the imaging plane. It is possible to interrogate 3D orientations from volumetric intensity data using structure tensor analysis (Wang et al., 2011; Wang et al., 2015), similar to the aforementioned histological applications. Another approach is

to collect PSOCT optic axis measurements with multiple light incidence angles on the tissue surface and use these measurements to infer the through-plane orientation. This approach has been previously demonstrated in biological tissue (Liu et al., 2016; Ugryumova et al., 2006; Ugryumova et al., 2009).

5. Conclusion

We have demonstrated that, when utilized in an appropriate manner, dictionary-based CS-DSI reconstructions can reduce acquisition times by a factor of 3 while preserving the accuracy of fiber orientation estimates with respect to PSOCT. In particular, given an adequate SNR level of the training data, the PCA method produced high-fidelity reconstructions that reliably maintained the accuracy of fully sampled DSI data. We also demonstrated that we can tolerate a non-negligible increase in the RMSE between PDFs obtained from CS-DSI and fully sampled DSI data, without incurring a large decrease in the accuracy of the peak orientations with respect to the axonal orientations measured with optical imaging. This underscores the need for assessing reconstruction quality with respect to ground-truth measurements. Our findings confirm the viability of CS-DSI as a technique for accelerating DSI acquisitions and offer valuable insights into the appropriate implementation of CS-DSI algorithms.

Acknowledgements

This work was funded by NIH grants R01-EB021265, K99-EB023993, R01-AG057672, R01-EB019956, and U01-MH117023. It was carried out at the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital, using resources provided by the Center for Functional Neuroimaging Technologies, P41-EB015896, a P41 Biotechnology Resource Grant supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Institutes of Health. This work also involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program (S10RR025563, S10RR023401, S10RR019307, and S10RR023043).

References

Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, *54*(11), 4311-4322.

Ali, S., Rohr, K., Axer, M., Amunts, K., Eils, R., & Wörz, S. (2017). Registration of ultra-high resolution 3D PLI data of human brain sections to their corresponding high-resolution counterpart. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017),

Ali, S., Wörz, S., Amunts, K., Eils, R., Axer, M., & Rohr, K. (2018). Rigid and non-rigid registration of polarized light imaging data for 3D reconstruction of the temporal lobe of the human brain at micrometer resolution. *NeuroImage*, *181*, 235-251.

Assaf, Y., Ben-Bashat, D., Chapman, J., Peled, S., Biton, I., Kafri, M., Segev, Y., Hendler, T., Korczyn, A., & Graif, M. (2002). High b-value q-space analyzed diffusion-weighted MRI: application to multiple sclerosis. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *47*(1), 115-126.

Axer, H., Beck, S., Axer, M., Schuchardt, F., Heepe, J., Flücken, A., Axer, M., Prescher, A., & Witte, O. W. (2011). Microstructural analysis of human white matter architecture using polarized light imaging: views from neuroanatomy. *Frontiers in neuroinformatics*, *5*, 28.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994a). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B*, *103*(3), 247-254.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994b). MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, *66*(1), 259-267.

Bilgic, B., Chatnuntawech, I., Setsompop, K., Cauley, S. F., Yendiki, A., Wald, L. L., & Adalsteinsson, E. (2013). Fast dictionary-based reconstruction for diffusion spectrum imaging. *IEEE transactions on medical imaging*, *32*(11), 2022-2033.

Bilgic, B., Setsompop, K., Cohen-Adad, J., Yendiki, A., Wald, L. L., & Adalsteinsson, E. (2012). Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries. *Magnetic Resonance in Medicine*, *68*(6), 1747-1754.

Budde, M. D., & Frank, J. A. (2012). Examining brain microstructure using structure tensor analysis of histological sections. *Neuroimage*, *63*(1), 1-10.

Choe, A., Stepniewska, I., Colvin, D., Ding, Z., & Anderson, A. (2012). Validation of diffusion tensor MRI in the central nervous system using light microscopy: quantitative comparison of fiber properties. *NMR in Biomedicine*, *25*(7), 900-908.

Daducci, A., Canales-Rodrı, E. J., Descoteaux, M., Garyfallidis, E., Gur, Y., Lin, Y.-C., Mani, M., Merlet, S., Paquette, M., & Ramirez-Manzanares, A. (2013). Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion MRI. *IEEE transactions on medical imaging*, *33*(2), 384-399.

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on information theory*, *52*(4), 1289-1306.

Fatima, Z., Motosugi, U., Hori, M., Onodera, T., Ishigame, K., Yagi, K., & Araki, T. (2013). Age-related white matter changes in high b-value q-space diffusion-weighted imaging. *Neuroradiology*, *55*(3), 253-259. https://doi.org/10.1007/s00234-012-1099-4

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., & Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in neuroinformatics*, *8*, 8.

Gorodnitsky, I. F., & Rao, B. D. (1997). Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Transactions on signal processing*, *45*(3), 600-616.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS biology*, *6*(7).

Henssen, D. J., Mollink, J., Kurt, E., van Dongen, R., Bartels, R. H., Gräβel, D., Kozicz, T., Axer, M., & van Walsum, A.-M. V. C. (2019). Ex vivo visualization of the trigeminal pathways in the human brainstem using 11.7 T diffusion MRI combined with microscopy polarized light imaging. *Brain Structure and Function*, *224*(1), 159-170.

Hori, M., Kamiya, K., Nakanishi, A., Fukunaga, I., Miyajima, M., Nakajima, M., Suzuki, M., Suzuki, Y., Irie, R., Kamagata, K., Arai, H., & Aoki, S. (2016). Prospective estimation of mean axon diameter and extra-axonal space of the posterior limb of the internal capsule in patients with idiopathic normal pressure hydrocephalus before and after a lumboperitoneal shunt by using q-space diffusion MRI. *European Radiology*, *26*(9), 2992-2998. https://doi.org/10.1007/s00330-015-4162-9

Jones, R., Grisot, G., Augustinack, J., Magnain, C., Boas, D. A., Fischl, B., Wang, H., & Yendiki, A. (2020). Insight into the fundamental trade-offs of diffusion MRI from polarization-sensitive optical coherence tomography in ex vivo human brain. *NeuroImage*, *214*, 116704. https://doi.org/https://doi.org/10.1016/j.neuroimage.2020.116704

Keil, B., Blau, J. N., Biber, S., Hoecht, P., Tountcheva, V., Setsompop, K., Triantafyllou, C., & Wald, L. L. (2013). A 64-channel 3T array coil for accelerated brain MRI. *Magnetic Resonance in Medicine*, *70*(1), 248-258. https://doi.org/10.1002/mrm.24427

Khan, A. R., Cornea, A., Leigland, L. A., Kohama, S. G., Jespersen, S. N., & Kroenke, C. D. (2015). 3D structure tensor analysis of light microscopy data for validating diffusion MRI. *Neuroimage*, *111*, 192-203.

Lacerda, L. M., Sperl, J. I., Menzel, M. I., Sprenger, T., Barker, G. J., & Dell'Acqua, F. (2016). Diffusion in realistic biophysical systems can lead to aliasing effects in diffusion spectrum imaging. *Magnetic resonance in medicine*, *76*(6), 1837-1847.

Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., & Laval-Jeantet, M. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, *161*(2), 401-407.

Leergaard, T. B., White, N. S., De Crespigny, A., Bolstad, I., D'Arceuil, H., Bjaalie, J. G., & Dale, A. M. (2010). Quantitative histological validation of diffusion MRI fiber orientation distributions in the rat brain. *PloS one*, *5*(1), e8595.

Liu, C. J., Black, A. J., Wang, H., & Akkin, T. (2016). Quantifying three-dimensional optic axis using polarization-sensitive optical coherence tomography. *Journal of biomedical optics*, *21*(7), 070501.

Lustig, M., Donoho, D., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *58*(6), 1182-1195.

Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed sensing MRI. *IEEE signal processing magazine*, *25*(2), 72-82.

Maffei, C., Girard, G., Schilling, K. G., Adluru, N., Aydogan, D. B., Hamamci, A., Yeh, F.-C., Mancini, M., Wu, Y., Sarica, A., Teillac, A., Baete, S. H., Karimi, D., Lin, Y.-C., Boada, F., Richard, N., Hiba, B., Quattrone, A., Hong, Y., Shen, D., Yap, P.-T., Boshkovski, T., Campbell, J. S. W., Stikov, N., Pike, G. B., Bendlin, B. B., Alexander, A. L., Prabhakaran, V., Anderson, A., Landman, B. A., Canales-Rodríguez, E. J. Z., Barakovic, M., Rafael-Patino, J., Yu, T., Rensonnet, G., Schiavi, S., Daducci, A., Pizzolato, M., Fischi-Gomez, E., Thiran, J.-P., Dai, G., Grisot, G., Lazovski, N., Puente, A., Rowe, M., Sanchez, I., Prchkovska, V., Jones, R., Lehman, J., Haber, S., & Yendiki, A. (2020). *The IronTract challenge: Validation and optimal tractography methods for the HCP diffusion acquisition scheme* ISMRM (oral presentation),

Majka, P., & Wójcik, D. K. (2016). Possum—a framework for three-dimensional reconstruction of brain images from serial sections. *Neuroinformatics*, *14*(3), 265-278.

Menzel, M. I., Tan, E. T., Khare, K., Sperl, J. I., King, K. F., Tao, X., Hardy, C. J., & Marinelli, L. (2011). Accelerated diffusion spectrum imaging in the human brain using compressed sensing. *Magnetic Resonance in Medicine*, *66*(5), 1226-1233.

Mollink, J., Kleinnijenhuis, M., van Walsum, A.-M. v. C., Sotiropoulos, S. N., Cottaar, M., Mirfin, C., Heinrich, M. P., Jenkinson, M., Pallebage-Gamarallage, M., & Ansorge, O. (2017). Evaluating fibre orientation dispersion in white matter: comparison of diffusion MRI, histology and polarized light imaging. *Neuroimage*, *157*, 561-574.

Paquette, M., Gilbert, G., & Descoteaux, M. (2016). Optimal DSI reconstruction parameter recommendations: better ODFs and better connectivity. *NeuroImage*, *142*, 1-13.

Paquette, M., Merlet, S., Gilbert, G., Deriche, R., & Descoteaux, M. (2015). Comparison of sampling strategies and sparsifying transforms to improve compressed sensing diffusion spectrum imaging. *Magnetic resonance in medicine*, *73*(1), 401-416.

Preibisch, S., Saalfeld, S., & Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics*, *25*(11), 1463-1465.

Reese, T. G., Benner, T., Wang, R., Feinberg, D. A., & Wedeen, V. J. (2009). Halving imaging time of whole brain diffusion spectrum imaging and diffusion tractography using simultaneous image refocusing in EPI. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *29*(3), 517-522.

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: a robust approach. *Neuroimage*, *53*(4), 1181-1196.

Schilling, K., Gao, Y., Janve, V., Stepniewska, I., Landman, B. A., & Anderson, A. W. (2017). Can increased spatial resolution solve the crossing fiber problem for diffusion MRI? *NMR in Biomedicine*, *30*(12), e3787.

Schilling, K., Janve, V., Gao, Y., Stepniewska, I., Landman, B. A., & Anderson, A. W. (2016). Comparison of 3D orientation distribution functions measured with confocal microscopy and diffusion MRI. *Neuroimage*, *129*, 185-197.

Schilling, K. G., Janve, V., Gao, Y., Stepniewska, I., Landman, B. A., & Anderson, A. W. (2018). Histological validation of diffusion MRI fiber orientation distributions and dispersion. *Neuroimage*, *165*, 200-221.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., & Schmid, B. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, *9*(7), 676.

Seehaus, A., Roebroeck, A., Bastiani, M., Fonseca, L., Bratzke, H., Lori, N., Vilanova, A., Goebel, R., & Galuske, R. (2015). Histological validation of high-resolution DTI in human post mortem tissue. *Frontiers in Neuroanatomy*, *9*, 98.

Setsompop, K., Cohen-Adad, J., Gagoski, B. A., Raij, T., Yendiki, A., Keil, B., Wedeen, V. J., & Wald, L. L. (2012). Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *Neuroimage*, *63*(1), 569-580.

Setsompop, K., Fan, Q., Stockmann, J., Bilgic, B., Huang, S., Cauley, S. F., Nummenmaa, A., Wang, F., Rathi, Y., & Witzel, T. (2018). High-resolution in vivo diffusion imaging of the human brain with generalized slice dithered enhanced resolution: Simultaneous multislice (g S lider-SMS). *Magnetic resonance in medicine*, *79*(1), 141-151.

Setsompop, K., Gagoski, B. A., Polimeni, J. R., Witzel, T., Wedeen, V. J., & Wald, L. L. (2012). Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic resonance in medicine*, *67*(5), 1210-1224.

Setsompop, K., Kimmlingen, R., Eberlein, E., Witzel, T., Cohen-Adad, J., McNab, J. A., Keil, B., Tisdall, M. D., Hoecht, P., Dietz, P., Cauley, S. F., Tountcheva, V., Matschl, V., Lenz, V. H., Heberlein, K., Potthast, A., Thein, H., Van Horn, J., Toga, A., Schmitt, F., Lehne, D., Rosen, B. R., Wedeen, V., & Wald, L. L. (2013). Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *NeuroImage*, *80*, 220-233. https://doi.org/10.1016/j.neuroimage.2013.05.078

Stejskal, E. O., & Tanner, J. E. (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, *42*(1), 288-292.

Tian, Q., Rokem, A., Folkerth, R. D., Nummenmaa, A., Fan, Q., Edlow, B. L., & McNab, J. A. (2016). Q-space truncation and sampling in diffusion spectrum imaging. *Magnetic resonance in medicine*, *76*(6), 1750-1763.

Tobisch, A., Stirnberg, R., Harms, R. L., Schultz, T., Roebroeck, A., Breteler, M., & Stöcker, T. (2018). Compressed sensing diffusion spectrum imaging for accelerated diffusion microstructure MRI in long-term population imaging. *Frontiers in neuroscience*, *12*, 650.

Ugryumova, N., Gangnus, S. V., & Matcher, S. J. (2006). Three-dimensional optic axis determination using variable-incidence-angle polarization-optical coherence tomography. *Optics letters*, *31*(15), 2305-2307.

Ugryumova, N., Jacobs, J., Bonesi, M., & Matcher, S. J. (2009). Novel optical imaging technique to determine the 3-D orientation of collagen fibers in cartilage: variable-incidence angle polarization-sensitive optical coherence tomography. *Osteoarthritis and cartilage*, *17*(1), 33-42.

Wang, H., Akkin, T., Magnain, C., Wang, R., Dubb, J., Kostis, W. J., Yaseen, M. A., Cramer, A., Sakadžić, S., & Boas, D. (2016). Polarization sensitive optical coherence microscopy for brain imaging. *Optics letters*, *41*(10), 2213-2216.

Wang, H., Black, A. J., Zhu, J., Stigen, T. W., Al-Qaisi, M. K., Netoff, T. I., Abosch, A., & Akkin, T. (2011). Reconstructing micrometer-scale fiber pathways in the brain: multi-contrast optical coherence tomography based tractography. *Neuroimage*, *58*(4), 984-992.

Wang, H., Lenglet, C., & Akkin, T. (2015). Structure tensor analysis of serial optical coherence scanner images for mapping fiber orientations and tractography in the brain. *Journal of biomedical optics*, *20*(3), 036003.

Wang, H., Magnain, C., Wang, R., Dubb, J., Varjabedian, A., Tirrell, L. S., Stevens, A., Augustinack, J. C., Konukoglu, E., & Aganj, I. (2018). as-PSOCT: Volumetric microscopic imaging of human brain architecture and connectivity. *Neuroimage*, *165*, 56-68.

Wang, H., Zhu, J., Reuter, M., Vinke, L. N., Yendiki, A., Boas, D. A., Fischl, B., & Akkin, T. (2014). Cross-validation of serial optical coherence scanning and diffusion tensor imaging: a study on neural fiber maps in human medulla oblongata. *Neuroimage*, *100*, 395-404.

Wedeen, V. J., Hagmann, P., Tseng, W. Y. I., Reese, T. G., & Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic resonance in medicine*, *54*(6), 1377-1386.

Wedeen, V. J., Wang, R., Schmahmann, J. D., Benner, T., Tseng, W.-Y. I., Dai, G., Pandya, D., Hagmann, P., D'Arceuil, H., & de Crespigny, A. J. (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage*, *41*(4), 1267-1277.