RESEARCH ARTICLE

WILEY

A practical evaluation of measures derived from compressed sensing diffusion spectrum imaging

Hamsanandini Radhakrishnan[1,2] | Chenying Zhao[1,2,3,4] | Valerie J. Sydnor[1,2] | Erica B. Baller[1,2] | Philip A. Cook[5] | Damien A. Fair[6] | Barry Giesbrecht[7] | Bart Larsen[1,2] | Kristin Murtha[1,2] | David R. Roalf[2,3] | Sage Rush-Goebel[1,2] | Russell T. Shinohara[8,9] | Haochang Shou[8,9] | M. Dylan Tisdall[5] | Jean M. Vettel[7,10] | Scott T. Grafton[7] | Matthew Cieslak[1,2] | Theodore D. Satterthwaite[1,2] [ORCID]

[1]Lifespan Informatics and Neuroimaging Center, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[3]Lifespan Brain Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[4]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[5]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[6]Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, Minnesota, USA

[7]Department of Psychological and Brain Sciences, University of California, Santa Barbara, California, USA

[8]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[9]Center for Biomedical Image Computing & Analytics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[10]U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, Maryland, USA

Correspondence

Theodore D. Satterthwaite, Richards Medical Labs, 5th Floor, Pod A, 3700 Hamilton Walk,

Abstract

Diffusion Spectrum Imaging (DSI) using dense Cartesian sampling of q-space has been shown to provide important advantages for modeling complex white matter architecture. However, its adoption has been limited by the lengthy acquisition time required. Sparser sampling of q-space combined with compressed sensing (CS) reconstruction techniques has been proposed as a way to reduce the scan time of DSI acquisitions. However prior studies have mainly evaluated CS-DSI in post-mortem or non-human data. At present, the capacity for CS-DSI to provide accurate and reliable measures of white matter anatomy and microstructure in the living human brain remains unclear. We evaluated the accuracy and inter-scan reliability of 6 different CS-DSI schemes that provided up to 80% reductions in scan time compared to a full DSI scheme. We capitalized on a dataset of 26 participants who were scanned over eight independent sessions using a full DSI scheme. From this full DSI scheme, we subsampled images to create a range of CS-DSI images. This allowed us to compare the accuracy and inter-scan reliability of derived measures of white matter structure (bundle segmentation, voxel-wise scalar maps) produced by the CS-DSI and the full DSI schemes. We found that CS-DSI estimates of both bundle segmentations and voxel-wise scalars were nearly as accurate and reliable as those generated by the full DSI scheme. Moreover, we found that the accuracy and reliability of CS-DSI was higher in white matter bundles that were more reliably segmented by the full DSI scheme. As a final step, we replicated the accuracy of CS-DSI in a prospectively acquired dataset ($n = 20$, scanned once). Together, these results illustrate the utility

Matthew Cieslak and Theodore D. Satterthwaite contributed equally to this study.

Philadelphia, PA 19104, USA.
Email: sattertt@pennmedicine.upenn.edu

of CS-DSI for reliably delineating in vivo white matter architecture in a fraction of the scan time, underscoring its promise for both clinical and research applications.

KEYWORDS
compressed sensing, diffusion-weighted imaging, MRI acquisition, white matter

1 | INTRODUCTION

Diffusion-weighted magnetic resonance imaging (dMRI) has become the dominant imaging modality for noninvasively characterizing white matter architecture in humans, with myriad applications in both research and clinical practice. While a large and growing variety of dMRI sequences exist, they all sample three spatial dimensions in k-space and three additional dimensions in q-space, which is defined by the diffusion-encoding gradients (Callaghan et al., 1988). Each 3D volume in a dMRI sequence samples a coordinate in q-space, with the entire set of points defining a q-space sampling scheme. Investigators have previously explored a wide range of q-space sampling schemes (Afzali et al., 2021; Assaf et al., 2004; Assaf & Basser, 2005; Basser et al., 1994; Caruyer et al., 2013; Jensen et al., 2005; Tuch, 2004; Wedeen et al., 2005; Wu et al., 2008), with most sampling points on the surface of one or more spheres (aka "shells"). One alternative strategy densely samples q-space on a Cartesian grid, enabling the direct estimation of the diffusion ensemble average propagator (EAP), the physical process driving biologically meaningful derivatives of dMRI. This scheme is often called diffusion spectrum imaging (DSI) and can resolve crossing fibers and other complex tract architectures. However, Cartesian sampling requires long scan times, which has limited the use of DSI in translational research or clinical practice. Recently, sparse random sampling of q-space combined with compressed sensing (CS) reconstruction techniques has shown massive promise in both post-mortem and animal studies (Jones et al., 2021a; Menzel et al., 2011; Naeyaert et al., 2021; Paquette et al., 2015). However, the validity of common white matter derivatives generated by compressed sensing accelerated DSI (CS-DSI) has not been adequately tested in in vivo human data. Here, we sought to validate the ability of CS-DSI to characterize both macro- and micro-scale white matter properties in living humans by evaluating its accuracy and reliability compared to a full DSI scheme.

The Cartesian sampling schemes used in DSI provide several practical and theoretical advantages due to their ability to model the EAP directly. DSI allows direct measurement of the EAP in each voxel due to the Fourier relationship between q-space and the EAP (Wedeen et al., 2005). This relationship between the q-space signal and the EAP obviates the choice of which shells to sample (i.e., which b-values for each shell). For example, high b-value shells can be more useful for

tractography in the presence of crossing fibers (Raffelt et al., 2017; Schilling et al., 2017) while low b-value shells are required for estimating some free water parameters (Assaf & Basser, 2005; Tuch et al., 2002; Yoshiura et al., 2001). Conversely, higher b-values have lower SNR, which may impact the accuracy of ODF estimation (Han et al., 2015). While multishelled schemes can also approximate the EAP, their limited sampling cannot always capture its full complexity (Novikov et al., 2018; Tristán-Vega et al., 2023). DSI helps mitigate these limitations by attaining a nonparametric sampling of the EAP at a regular Cartesian lattice and densely sampling the q-space to avoid aliasing artifacts (Wedeen et al., 2008). Nonetheless, DSI still requires other parameters to be chosen, such as the maximum b-value or the number of grid points, and these factors will be limited by scanner hardware specifics and the amount of scan time available.

Previous research has demonstrated that DSI provides more accurate fiber reconstructions: DSI produces more accurate ODF estimates (Daducci et al., 2014) and achieves greater biological fidelity in tractography (particularly when using anatomic tracing in non-human primates as ground-truth; Maffei et al., 2021). Probability density functions derived from DSI have also exhibited heightened sensitivity to microstructural characteristics such as axon size (Hori et al., 2016), age-related white matter changes (Fatima et al., 2013), and lesions in individuals with multiple sclerosis (Assaf et al., 2002). Furthermore, measures derived from DSI acquisitions may be sensitive to complex biological changes like neuroinflammation, axonal loss, and demyelination (Wang et al., 2020, 2021; Yeh et al., 2013, 2016; Zhang et al., 2013). EAPs reconstructed from DSI have been shown to result in improved tractography (Yendiki et al., 2021) and enhanced gray-white contrast (Jones et al., 2021b). Together, available research suggests potential for DSI. However, DSI's Cartesian sampling scheme comes at the cost of a much longer acquisition time: a typical DSI scan can take upward of 20 min to acquire, limiting its applications in both time-constrained research settings and clinical practice.

One potential solution to this obstacle is CS (Candes & Wakin, 2008; Donoho, 2006). CS is a widely used technique that can reconstruct signals from undersampled data and has been very valuable for advances in diverse fields including telecommunications (Berger et al., 2010), astronomy (Bobin & Starck, 2009; Wiaux et al., 2009), radar (Baraniuk & Steeghs, 2007; Herman & Strohmer, 2009), and medical imaging (Carroll et al., 2009; Lustig

et al., 2007). In MRI, CS is often used to accelerate acquisitions by undersampling k-space (Adluru & Dibella, 2008; He et al., 2007; Jaspan et al., 2015; Lustig et al., 2008). However, recent work has shown that dMRI acquisitions can under-sample q-space and still accurately estimate EAPs and other derivatives after using CS methods. (Menzel et al., 2011; Merlet & Deriche, 2013; Michailovich et al., 2011; Paquette et al., 2015; Pu et al., 2011; Ramirez-Manzanares et al., 2007). Prior work has provided some guidance regarding optimal q-space sampling strategies (Menzel et al., 2011; Merlet & Deriche, 2013) and effective signal recovery methods (Bilgic et al., 2012; Cheng et al., 2011; Gramfort et al., 2012; Tristán-Vega & Westin, 2011; Ye et al., 2012). These sampling and recovery methods have been validated in simulated data (Menzel et al., 2011; Merlet & Deriche, 2013), post-mortem tissue (Jones et al., 2021a), and in small samples of in vivo human data ($n < 3$; Bilgic et al., 2013, 2012; Merlet & Deriche, 2013).

While these studies have underscored CS-DSI's promise, most of this work has primarily focused on sequence development and reconstruction techniques and no study to our knowledge has compared CS-DSI outputs to full DSI outputs in living humans. Signal processing theory dictates that under-sampled signals can be reconstructed with CS at a low approximation error, but MRI acquisition includes substantial inherent noise. The diffusion signal and its derivatives vary across scanners, protocols, and populations (Schilling et al., 2021b). It is unclear how the approximation error arising from under-sampling the q-space affects ODFs, and how these ODFs compare to those derived from full DSI schemes. Moreover, as popular diffusion derivatives such as tractography, bundle segmentations, or scalar maps are not mere linear transforms of the ODF, it is unclear how these errors are compounded in the interpretation of diffusion imaging outcomes. As such, important benchmarks of in vivo validity, including the accuracy and reliability of commonly used derivatives such as white matter bundles and scalar maps, have not been established. Furthermore, the trade-off between undersampling and robustness remains unknown.

We sought to establish the validity of CS-DSI in living humans. We were motivated by the possibility that CS-DSI could provide reliable measures of white matter anatomy in a fraction of the scan time required by a full DSI scan. Such an advance would allow both researchers and clinicians to harness the advantages of DSI sequences that were previously impractical to deploy. We examined six different CS-DSI schemes and evaluated whether these schemes could generate white matter bundles and scalar maps comparable to those generated by a full DSI scheme. While there are many ways to sparsely sample q-space, we focused on the homogenous angular sampling scheme (HA-SC) (Merlet, 2013). HA-SC schemes ensure homogenous angular coverage of q-space and have been shown to produce more consistent reconstructions than other sampling strategies (Jones et al., 1999; Menzel et al., 2011). To assess the relationship between CS-DSI validity and scan time, we included a range of sampling densities in our schemes that provided between 50% and 80% reduction in scan time. We capitalized on a unique dataset of 26 participants who were scanned over eight independent sessions using a full DSI scheme. From this full DSI scheme, we subsampled images to retrospectively generate synthetic CS-DSI images. Using these images, we compared CS-DSI to full DSI on the accuracy and test–retest reliability of commonly used derivatives. Finally, we replicated our results in an independent dataset of 20 participants who were prospectively scanned on a different scanner using CS-DSI acquisition schemes. As described below, we found that CS-DSI schemes offer highly accurate and reliable alternatives to the full DSI scheme, while only requiring a fraction of the scan time.

2 | METHODS

2.1 | Overview

Our full DSI scheme sampled 258 points on a Cartesian grid in q-space. The CS-DSI schemes we evaluate here (Figure 1) are all subsets of this full DSI scan. As such, we can simulate their acquisition retrospectively by extracting the volumes contained in each CS sampling scheme from a full DSI acquisition. The dataset of full DSI acquisitions included 26 participants scanned over 8 sessions (average time between sessions = 14 days), acquiring the same full 258-direction scheme at each scan session (Cieslak et al., 2018; Nakuci et al., 2023). We then subsampled six different CS-DSI schemes from this full DSI dataset. Four of these schemes (HA-SC92, HA-SC55-1, HA-SC55-2, and RAND57) were designed such that concatenating them would result in the full DSI scheme. To explore a larger range in the number of directions sampled, we generated two additional CS-DSI schemes (HA-SC92 + 55−1, HA-SC92 + 55−1) by combining pairs of these subsampled CS-DSI schemes (Figure 1). To additionally assess whether the behavior of retrospectively assembled CS-DSI schemes could be replicated in new data, we prospectively acquired four of the CS-DSI schemes that together made up the full 258-direction grid on an additional 20 participants who were each scanned once (Figure 2). For all of the CS-DSI schemes, we first extrapolated a full DSI image using an iterative L2-regularized algorithm (Bilgic et al., 2012). For brevity, we will refer to this extrapolated full DSI image as the CS-DSI image. These CS-DSI images were then reconstructed with Generalized Q-Sampling Imaging (GQI) in both datasets (Yeh et al., 2010).

We evaluated the CS-DSI image's ability to segment bundles and generate voxel-wise scalars on two criteria: accuracy and reliability. We defined accuracy as the similarity between a measure derived from a CS-DSI image and that derived from the full DSI image (the acquired image for the retrospective data, and the concatenated image for the prospective data). As described below, we evaluated accuracy both within- and across scan sessions in the multisession retrospective data; only same-scan accuracy could be evaluated in the prospective data. Inter-scan reliability was operationalized as the similarity between a given measure derived from the same diffusion scheme in different scan sessions; this could only be calculated for the multisession retrospective data. Throughout, we compared bundles using calculated Dice scores and scalar maps using Pearson correlations.
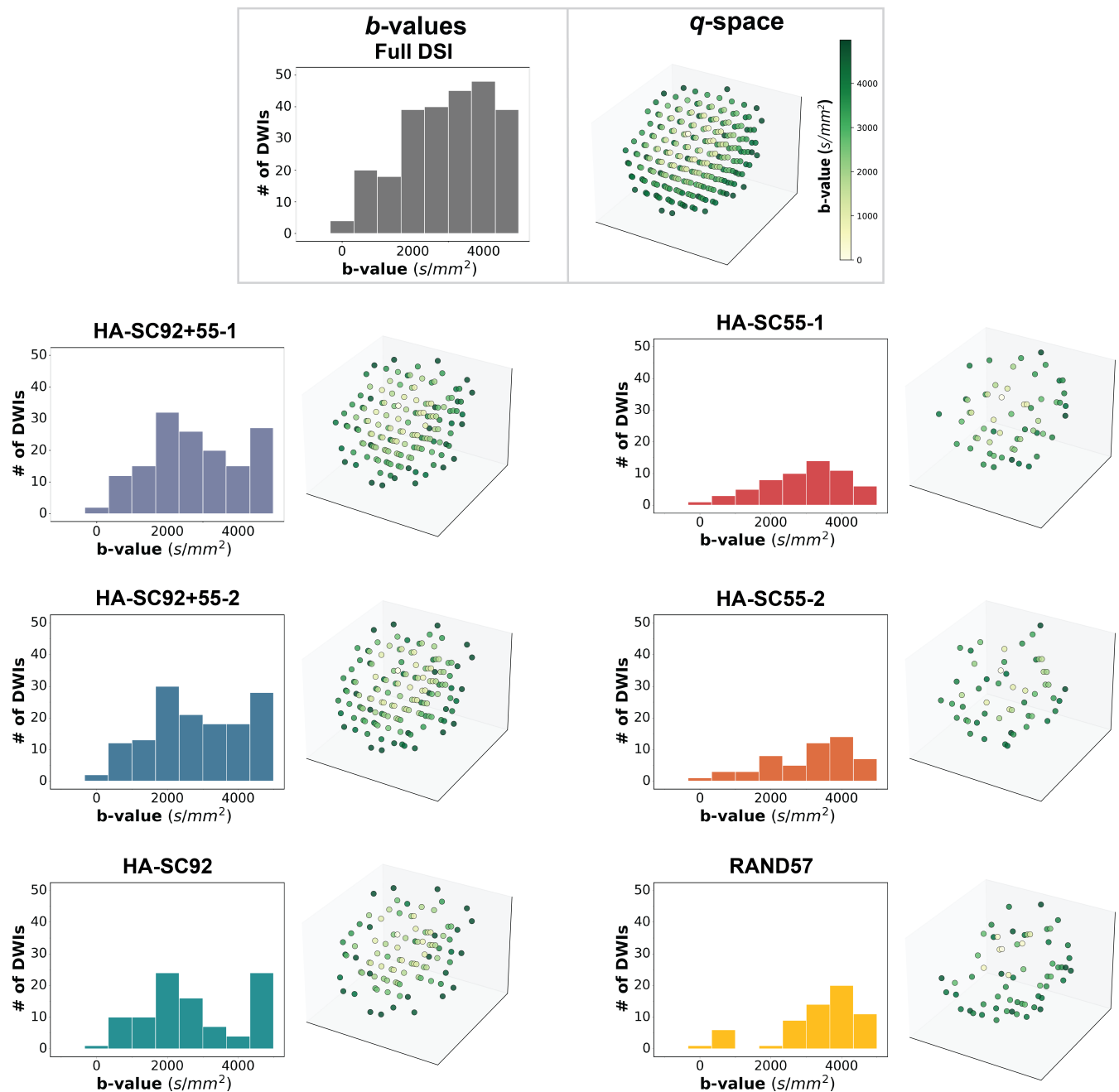
FIGURE 1 Description of the examined diffusion schemes with the histogram of b-values and distribution of b-vectors in q-space. B-vectors are scaled, and color coded by b-value. HA-SC, Homogenous Angular Sampling Scheme; RAND, Random Sampling Scheme.

2.2 | Data acquisition

2.2.1 | Retrospective full DSI data

The retrospective imaging data was extracted from a repeated measures dataset containing 8 full DSI scans (average time between sessions = 14 days) in a group of 26 healthy adults (mean age 22 ± 3.5 years, 16 Female) (Cieslak et al., 2018; Gu et al., 2015; Muldoon et al., 2016; Nakuci et al., 2023; Teich et al., 2021). CS-DSI

schemes were created by extracting volumes from the full DSI scan. The participants were scanned using a Siemens Prisma 3T MRI scanner. Fitted padding was used to minimize head movements. Each dMRI scan was acquired with a full DSI scheme (i.e., Q5, half-sphere) comprised of 258 images with variable b-values ($b_{max} = 5000$ s/mm^2) corresponding to 258 different isotropic grid points in q-space. Bipolar diffusion encoding gradients were used. The in-plane resolution and slice thickness were both set to 2 mm. The images were acquired with a TE of 100.2 ms and a TR of 4300 ms. The total scan time was

FIGURE 2 Preprocessing schematic and analytic overview.
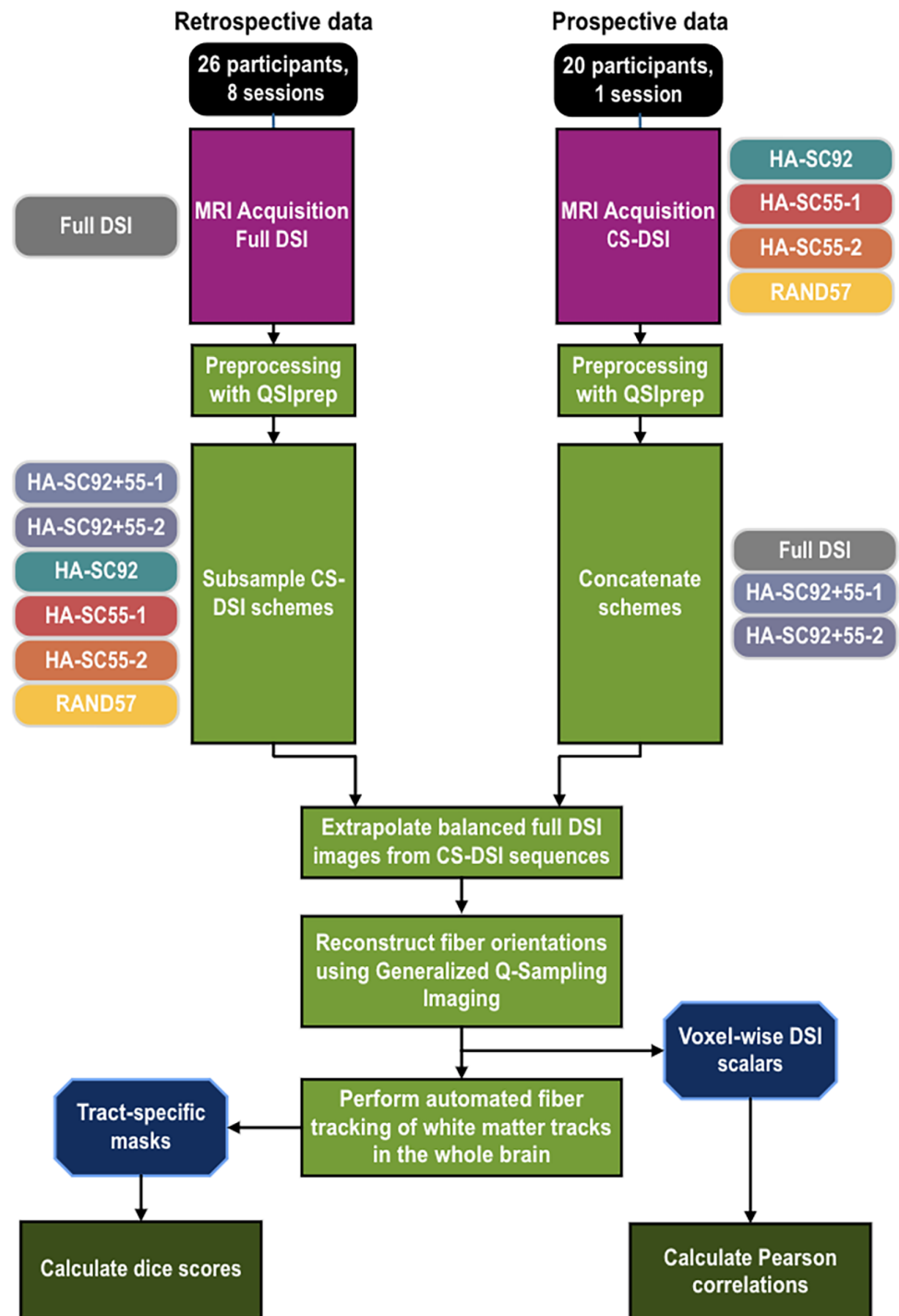
20 minutes (Table 1). Due to incomplete data, two participants were removed from the final dataset, giving a total of 24 participants. We refer to this dataset as the "retrospective" dataset.

2.2.2 | Prospective CS-DSI data

After initial analysis of retrospectively synthesized CS-DSI schemes, we sought to replicate our results with prospectively acquired CS-DSI. All participants provided informed consent before participation in this study. All experimental procedures were approved by the University of Pennsylvania Review Board. We scanned 20 healthy adults (mean age = 26.5 ± 3.5 years, 9 Female) using a Siemens Prisma 3 T MRI scanner under four different CS acquisition schemes (Figure 1). Three of the four CS schemes were homogenous angular sampling schemes (HA-SC) (Paquette et al., 2015) sampling 92, 55, and 55 directions, respectively. HA-SC schemes are sampling schemes that ensure a random but uniform angular covering of the q-space (Figure 1). Note that some points in q-space were repeated across the three HA-SC schemes to ensure consistency in the big and little deltas across scans.

TABLE 1   Acquisition time for each scheme.

Scheme	Number of unique directions (including $b = 0$)	Acquisition time (minutes)
Full DSI	258	20
HA-SC92 + 55–1[a]	146	11.9
HA-SC92 + 55–2[a]	146	11.7
HA-SC92	92	7.4
HA-SC55–1	55	4.5
HA-SC55–2	55	4.3
RAND57	57	5.4

[a]These schemes were concatenated from acquired schemes.

The fourth scheme was a random scheme (RAND) sampling 57 directions. In the RAND scheme, samples were not uniformly distributed and were rather designed such that combining the RAND scheme with the unique samples of all the acquired HA-SC schemes resulted in the balanced full DSI scheme used in the retrospective dataset (Figure 1). The exact schemes tested in this manuscript are available for download at https://github.com/PennLINC/cs_dsi/tree/main/b-tables. The in-plane resolution and slice thickness were both set to 2 mm with a multiband acceleration factor of 4. The volume images were acquired using a TE of 90 ms and a TR of 4300 ms. The acquisition times for the CS-DSI schemes were significantly lower than that for the full DSI scheme (Table 1). We refer to this dataset as the "prospective" dataset. This data is available in OpenNeuro for open access under a CC0 license (https://10.18112/openneuro.ds004737.v1.0.0).

2.3 | Preprocessing

Preprocessing was performed using *QSIPrep* 0.4.0 (Cieslak et al., 2021), which is based on *Nipype* 1.1.9 (Gorgolewski et al., 2011, 2017) on the retrospective data and *QSIPrep* 0.8.0 (*Nipype* 1.4.2) on the prospective data. *QSIprep* was run on the full DSI scan in the retrospective dataset and separately on each of the 4 directly acquired CS-DSI scans in the prospective dataset. For all full DSI and CS-DSI scans, initial motion correction was performed using only the $b = 0$ volumes. An unbiased $b = 0$ template was constructed over 3 iterations of affine registrations. The SHORELine method (Cieslak et al., 2022) was used to estimate head motion in $b > 0$ volumes. SHORELine is a cross-validated method that entails leaving out each $b > 0$ image and fitting the 3dSHORE basis function (Ozarslan et al., 2013) to all the other volumes using L2-regularization. The signal for the left-out volume serves as the registration target. A total of two iterations were run using an affine transform. Model-generated volumes were transformed into alignment with each $b > 0$ volume. Both slice-wise and whole-brain quality control measures (cross correlation and R^2) were calculated. No susceptibility distortion correction was performed. The DWI time series were resampled such that they were aligned to anterior and posterior commissures (ACPC),

generating a preprocessed DWI run in ACPC space. The accuracy of *b*-table orientation was examined by comparing fiber orientations with those of a population-averaged template (Schilling et al., 2019).

The four CS schemes used in acquiring the prospective data were designed such that combining them resulted in a full DSI image. To assess whether CS schemes longer than the ones acquired could generate results closer to the full DSI, two additional CS schemes were also generated by combining the preprocessed HA-SC92 scheme with each of the preprocessed HA-SC55 schemes (forming HA-SC92 + 55–1 and HA-SC92 + 55–2). For each participant in the retrospective data, the 6 CS-DSI schemes mentioned above were generated from the full-DSI scheme by extracting out only the relevant volumes from the full preprocessed image. This distinction is important to note as the object of comparison for CS-DSI derivatives is their corresponding full-DSI derivatives. Critically, in the retrospective data, the full-DSI is acquired in a single scan where all images are preprocessed together. The prospective full-DSI is composed of four individual scans that were each preprocessed separately and then concatenated.

2.4 | CS reconstruction

Following preprocessing, we used DSI Studio to perform fiber tracking and to derive scalar metrics. GQI requires balanced q-space sampling, so a full, balanced 258-direction DSI image was extrapolated from the CS-DSI schemes. Extrapolation was performed via CS using the 3DSHORE basis function (Ozarslan et al., 2013) with a radial order of 8 and L2 regularization. For brevity, we will refer to this extrapolated full DSI image as the CS-DSI image from now on. These GQI reconstructions with a diffusion sampling length ratio of 1.25 were used as the input for fiber tracking and anisotropy scalar estimation (Figure 2).

2.5 | Tractography and bundle segmentation

A deterministic fiber tracking algorithm (Yeh et al., 2013) was used with three augmented tracking strategies: parameter saturation, atlas-based track recognition, and topology-informed pruning (Yeh, 2020) to improve reproducibility. The exact methods used here are described in Yeh et al. (2020). Briefly, fiber tracking seeds were defined using the HCP842 tractography atlas (Yeh et al., 2018). The tracking parameters were saturated using a random generator and the generated streamlines were culled if their trajectory had a Hausdorff greater than 16 mm from the transformed atlas bundle. The track-to-voxel ratio was set to two. Topology-informed pruning (Yeh et al., 2019) was applied to the tractography with 32 iteration(s) to remove false connections. Shape analysis (Yeh, 2020) was conducted to derive shape metrics for tractography. A total of 56 fiber bundles were segmented in this way (Table S1). Each bundle was segmented 1344 times for the retrospective dataset (24 participants, 8 scan sessions, 7 diffusion schemes) and 140 times for the prospective dataset (20 participants, 1 session, 7 diffusion schemes). Note, however, that

the following bundles could not be segmented for some of the schemes in some of the participants: Left vertical occipital fasciculus (507 failures), right vertical occipital fasciculus (250 failures), left parahippocampal cingulum (118 failures), right parahippocampal cingulum (153 failures), and the left fornix (1 failure). Failures were observed evenly across all schemes and were excluded from the final analysis.

The segmented bundles were then converted into binary masks, and Dice scores were calculated between the masks for each scenario using the following formula:

$$Dice = 2 * \frac{|X \cap Y|}{|X| + |Y|}.$$

High Dice scores denoted high overlap between the two bundles examined and vice-versa (Figure 3). For each CS scheme and bundle, accuracy of bundle segmentation was defined as the Dice score between the bundle generated by the CS image and the bundle generated by the corresponding full DSI image. Two different types of accuracies were calculated: same-scan accuracy and inter-scan accuracy. Same-scan accuracy was defined as the Dice score between the bundle generated by the CS-DSI image and the corresponding bundle generated by the full DSI image (the acquired full image for the retrospective data, and the combined image for the prospective data) for a participant, within the same scan session. Hence each participant in the retrospective dataset had 8 values for same-scan accuracy (one for each scan session) for a given bundle and diffusion scheme, and participants in the prospective dataset had one value. To account for the noise between scan sessions, inter-scan accuracy was defined as

the Dice score between the bundle generated by the CS-DSI image in one scan session and the bundle generated by the full DSI image in another scan session, for the same participant. Inter-scan accuracy was calculated for each pair of scan sessions, so each participant in the retrospective dataset had 56 values for inter-scan accuracy (one for each permuted pair of scan sessions).

Similarly, inter-scan reliability for a given diffusion scheme and bundle was defined as the set of Dice scores between the bundles generated by each pair of scan sessions. Notably, inter-scan accuracy compared pairwise session differences between the CS-DSI and full DSI images, while inter-scan reliability compared pairwise session differences within a scheme. Inter-scan reliability was also calculated independently for each participant in the retrospective dataset. As above, given the number of session pairs, each participant had 28 values for reliability (one for each unique pair of scan sessions) for a given bundle and diffusion scheme (Table 2). Inter-scan metrics could not be calculated in the prospective dataset because of the absence of multisession data. For brevity, we refer to the set of these three metrics (same-scan accuracy, inter-scan accuracy and inter-scan reliability) as *validity* metrics in the rest of this paper.

2.6 | Deriving whole-brain voxel-wise maps of white matter microstructure

While macrostructural derivatives like bundles are very useful for studying white matter connectivity, sometimes more localized measures are useful. Voxel-wise scalar metrics can provide additional
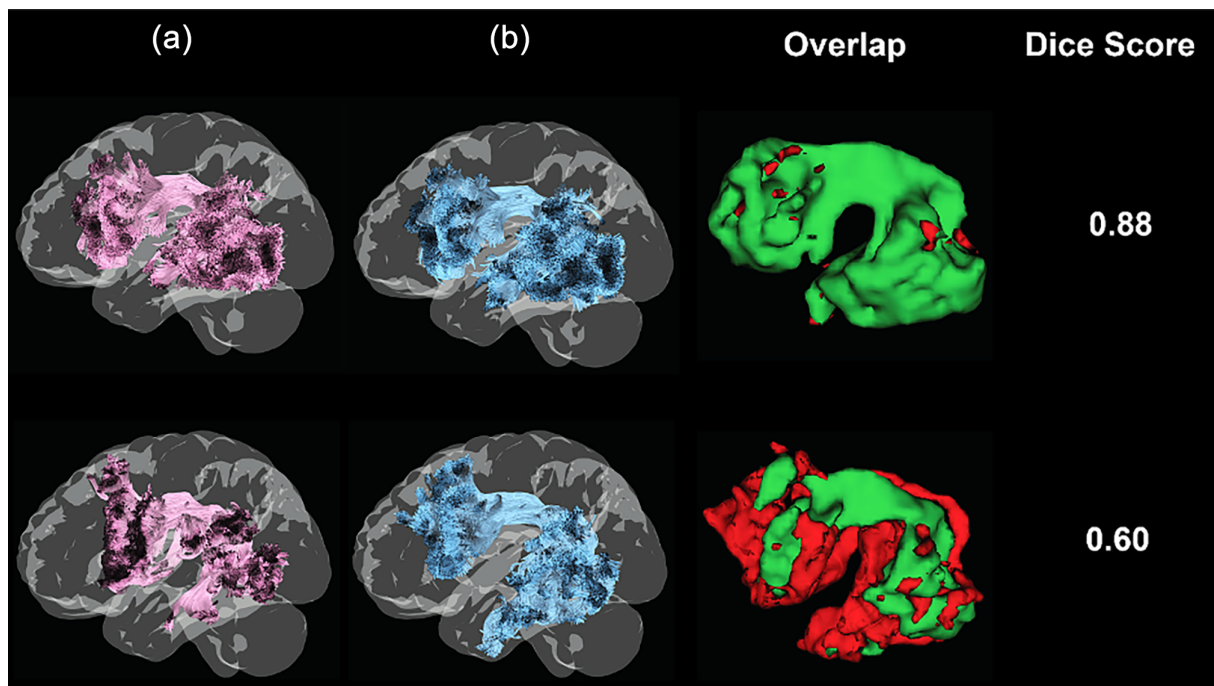
FIGURE 3 Graphical representation of range of Dice scores. The exemplar bundle shown here is the left arcuate fasciculus. The first two columns (a and b) show the left arcuate fasciculus derived from different DSI images (top row: same participant, different full DSI sessions; bottom row: same participant and session, RAND57 and full DSI). The third column shows voxels that overlap between the bundles (green) and differences (red).

TABLE 2 Definitions of validity metrics used to compare diffusion schemes.

Metric	Definition	Scheme	Datasets
Same-scan accuracy	Relationship between derivative produced by a CS-DSI image and the corresponding derivative produced by the full DSI image, within a participant and session	CS-DSI only	Both retrospective and prospective
Inter-scan accuracy	Relationship between derivative produced by a CS-DSI image in one session, and the derivative produced by the full DSI image in another session, within a participant. Calculated between each pair of sessions.	CS-DSI only	Only retrospective
Inter-scan reliability	Relationship between derivative produced by a DSI image in one session, and the derivative produced by the same DSI image in another session, within a participant. Calculated between each pair of sessions.	CS-DSI and full DSI	Only retrospective

Note: The relationship is defined by the Dice score when the derivatives are bundle segmentations and by the Pearson correlation when the derivatives are voxel-wise scalars.
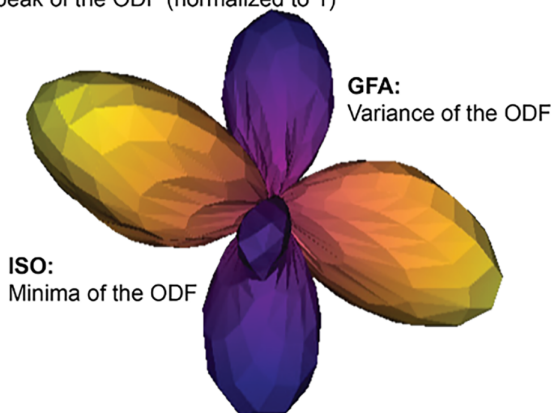
FIGURE 4 Schematic of DSI scalars examined.

insight into white matter structure and health and could capture pathology or other information at the microstructural level. GQI also allowed us to derive voxel-wise scalar maps, permitting more fine-grained comparisons (Yeh et al., 2010). We examined three ODF-based scalar metrics that quantified different fiber properties: Normalized quantitative anisotropy (NQA), generalized fractional anisotropy (GFA), and the isotropic value (ISO) (Yeh et al., 2013, Figure 4). All three metrics were calculated from the orientation distribution function (ODF) of each voxel. Quantitative anisotropy (QA) is the value of the highest peak of the ODF. QA is similar to the fractional anisotropy (FA) measure derived from tensors, except QA is defined per fiber population and so is less affected by crossing fibers. QA has arbitrary units, so we scaled the maximum QA of a subject to 1 to generate the more easily comparable normalized QA (NQA). In contrast to NQA, GFA is defined as the standard deviation between all vertices of the ODF divided by the root mean square of the vertices. Both GFA and NQA are measures of directional integrity, but GFA is more sensitive to differences in ODF shapes, as broadly different ODF structures can

have the same maxima but differing peak variances. Finally, ISO is the minima of the ODF, and mainly measures background isotropic diffusion from CSF (Yeh et al., 2013). Together, these scalars describe diverse properties of a given voxel. For each scalar metric examined, we calculated the Pearson correlation coefficient between whole-brain maps generated by CS-DSI and full DSI images. As with the Dice scores, the same-scan accuracy of each metric was defined as the Pearson correlation between the voxel-wise metric calculated from the CS-DSI image and that calculated from the full DSI image within a particular scan session. The inter-scan accuracy was the Pearson correlation between the voxel-wise metric calculated from the CS-DSI image in one scan session and that calculated by the full DSI image in a different scan session, for each pair of sessions. The inter-scan reliability was the set of Pearson correlations between voxel-wise metrics derived from each pair of scan sessions within a single diffusion scheme (see Table 2).

2.7 | Statistical analysis

Each of the validity metrics described in Table 2 was calculated for every participant, session, and bundle/scalar. The inter-scan reliability of the full DSI (from the retrospective dataset) was used as the "gold" standard for all comparisons. Therefore, for each bundle/scalar, the distributions of each of the validity metrics of CS-DSI across all participants were compared with the distribution of the inter-scan reliability of the full DSI. To determine if there was a statistically significant difference between these distributions, we compared the median difference between these distributions against a null distribution generated by permutation testing (see Supporting Information). Two-sided *p*-values were calculated across subjects per bundle for the Dice score comparisons. We corrected for multiple comparisons (56 bundles for the Dice scores and three metrics examined for the scalars) using the Benjamini/Hochberg false discovery rate method at 5% (Benjamini & Hochberg, 1995).

3 | RESULTS

3.1 | CS-DSI schemes can segment major fiber bundles

One of the main applications of dMRI is to delineate white matter bundles. We first qualitatively examined whether both retrospectively generated and prospectively acquired CS-DSI images could segment white matter bundles of varying shapes, convergent with bundles segmented by the corresponding full DSI image (i.e., the acquired full DSI image for the retrospective data, and the concatenated full DSI image for the prospective data). Inspection of data from individual participants revealed that we could adequately segment bundles from all CS-DSI images (see exemplar results in Figure 5).

Next, we quantitatively validated CS-DSI's ability to delineate bundles and calculated voxel-wise scalars, comparing CS-derived results to those derived from the full DSI image. The inter-scan reliability of the full DSI image was selected as the gold standard for comparison, as it captured noise and other session-related differences. Inter-scan reliability was calculated from the retrospective dataset and was defined as the set of relationships between derivatives produced by each pair of sessions within a given diffusion scheme (Table 2).

3.2 | Bundles segmented by CS-DSI images are similar to those segmented by their corresponding full DSI images in the same scan session

To quantitatively compare the bundles segmented by CS-DSI images and those segmented by full DSI images, we examined two measures of accuracy in the retrospective data: same-scan and inter-scan accuracy. Same-scan accuracy was defined as the Dice score between each bundle segmented by a CS-DSI image and the bundle segmented by the full DSI image for a given participant within the same scan session (Table 2). We found that all CS-DSI images showed high same-scan accuracies across fiber bundles (with median Dice scores ranging 0.72–0.82, Figure 6a). For additional context, we used full DSI inter-scan reliability as the standard of reference as test–retest differences are inevitable in neuroimaging, even when using longer full-grid schemes. We asked whether the difference between CS-DSI and full DSI bundles were comparable with test–retest differences of full DSI bundles. To this end, we compared the distributions of these same-scan accuracies of each CS-DSI scheme with the distributions of the full DSI inter-scan reliabilities across participants and sessions for every bundle generated. p-values were calculated per bundle using permutation testing on the median differences between these distributions (see Supporting Information). For the more densely sampled CS-DSI schemes, like HA-SC92 + 55−1, HA-SC92 + 55−2, and HA-SC92, the same-scan accuracy was statistically the same as the full DSI inter-scan reliability (FDR-corrected p-values >.1 for all bundles where full DSI reliability > CS-DSI accuracy, Table S2). As expected, the schemes with fewer directions sampled had an overall lower segmentation accuracy, and the RAND57 scheme performed worse than all the uniformly sampled HA-SC schemes. While the sparser sampling schemes showed significant differences across almost all bundles, the median difference between CS-DSI same-scan accuracy and full DSI inter-scan reliability was still quite small (HA-SC55−1 = 0.039, HA-SC55−2 = 0.058, RAND57 = 0.074, Figure 6a, Table S2). These results suggest that bundles from CS-DSI are comparable to bundles generated using the full DSI grid scheme and that accuracy scales with the degree of undersampling.

Not all bundles are reliably segmented, even when a full DSI image is used for tractography (Table S3). As such, we next asked whether the accuracy of CS-based bundle segmentations was related to how reliably that bundle was segmented with a full DSI scheme.
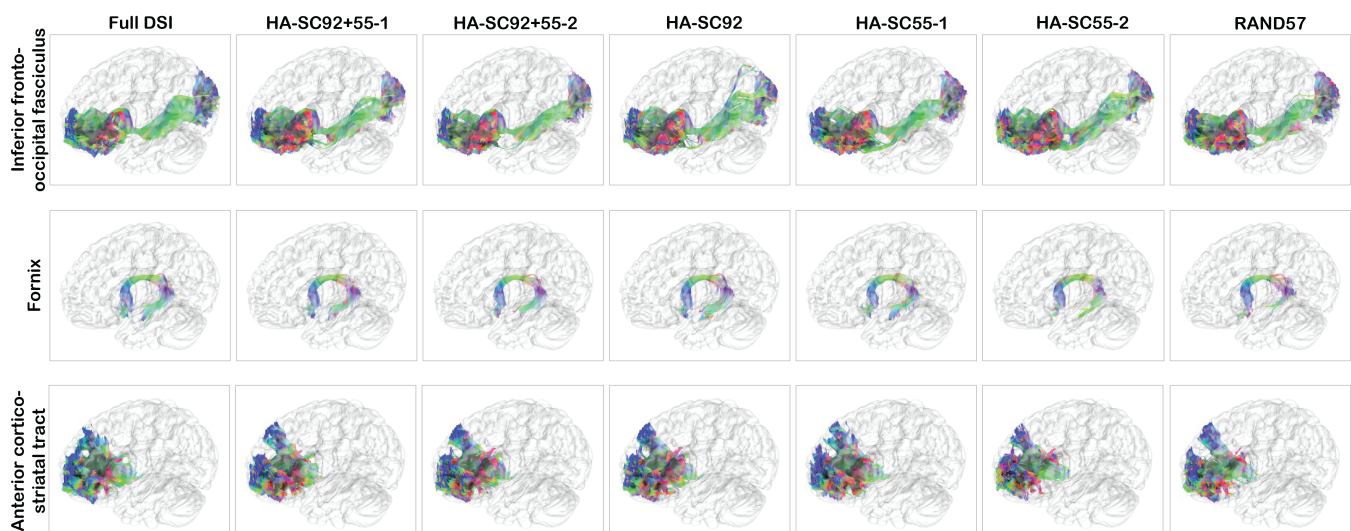
FIGURE 5 All CS images can segment both long-range and short-range bundles, comparable with those derived from a full DSI image. These figures are exemplar segmentations of different bundles in a single participant from the retrospective dataset.

Retrospective same-scan accuracy

Retrospective inter-scan accuracy
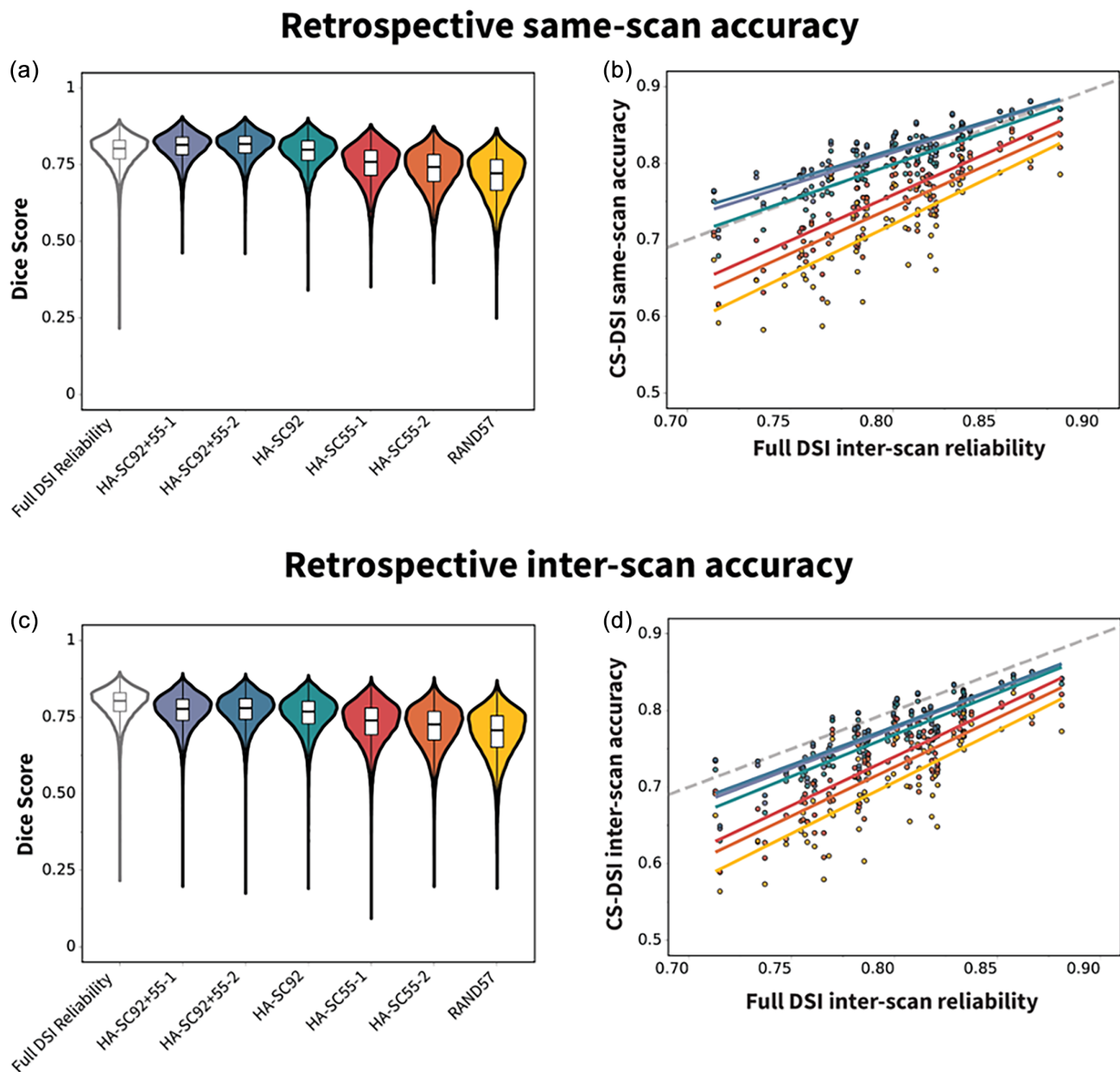
FIGURE 6 CS-DSI bundle segmentation accuracy is comparable and correlated with full DSI inter-scan reliability both within and across scan sessions. (a) Same-scan accuracy of CS-DSI schemes is comparable with inter-scan reliability of the full DSI scheme. Violins represent distributions of Dice scores across all bundles. (b) Median distribution of the same-scan accuracy for each CS-DSI scheme is highly correlated with the median distribution of the inter-scan reliability of the full DSI scheme. Here, each point on the scatter plot represents the median of Dice scores for a single bundle across participants and sessions. The gray dashed line denotes $x = y$. (c) Inter-scan accuracy of CS-DSI schemes is comparable with inter-scan reliability of the full DSI scheme. Violins represent distributions of Dice scores across all bundles. (d) Median distribution of the inter-scan accuracy for each CS-DSI scheme is highly correlated with the median distribution of the inter-scan reliability of the full DSI scheme. Here, each point on the scatter plot represents the median of Dice scores for a single bundle across participants and session pairs. The gray dashed line denotes $x = y$.

For example, the left and right reticular tracts had low same-scan accuracy across most CS-DSI schemes. We reasoned that this might be due to poorer segmentation reliability of this bundle even when using full DSI. Across bundles, we found that CS-DSI same-scan accuracy was strongly correlated with full DSI inter-scan reliability (Pearson R ranging from 0.78 to 0.89; $p < .001$ for all CS-DSI schemes, Figure 6b). This suggests that bundles that are not reliably segmented by the full DSI images were also less likely to be accurately

segmented by any of the CS-DSI images. Moreover, the more sparsely sampled CS-DSI images were more severely affected by this relationship and yielded lower Dice scores in bundles that showed low inter-scan reliability with full DSI. Bundles with high inter-scan reliability with full DSI were more accurately segmented by all CS-DSI images, including the more sparsely sampled schemes. Notably, the HA-SC92 + 55–1 and HA-SC92 + 55–2 schemes had same-scan accuracies that were *higher* than the inter-scan reliability of the full DSI for most

bundles (Figure 6b). This is likely due to the fact that differences between scan sessions were greater than differences introduced by subsampling of the same input data.

3.3 | Bundles segmented by CS-DSI images are also similar to those segmented by full DSI images acquired in a different scan session

The prior analysis compared bundles from the full DSI images to bundles generated from sub-sampled CS-DSI images within the same scan. To further examine CS-DSI segmentation accuracy while accounting for scan differences, we examined a second measure: inter-scan accuracy (Table 2). We defined inter-scan accuracy as the set of Dice scores between CS-DSI bundles and full DSI bundles in disparate sessions. Understandably, the inter-scan accuracies were lower than the same-scan accuracies on average, as they were derived from different sessions (median Dice scores ranging from 0.71 to 0.78, Table S4). This accuracy scaled with the number and uniformity of directions sampled, with dense HA-SC schemes performing with higher accuracy than the RAND scheme. While all bundles showed a significant difference between CS-DSI inter-scan accuracy and full DSI reliability for almost all diffusion schemes, the median difference between these distributions was very low (<.09, Figure 6c, Table S4). All CS-DSI schemes could still delineate very similar bundles with comparable anatomy to those delineated by full DSI. These results demonstrate that CS-DSI schemes can generate bundles similar to those segmented by the full DSI scheme in a different scan session.

We next asked if the CS-DSI inter-scan accuracy of specific bundles similarly scaled with their full DSI inter-scan reliability. We again found that bundles that were more reliably segmented by full DSI had higher inter-scan accuracies, and vice-versa (Pearson R ranging from 0.78 to 0.88; p < .001, Figure 6d). As observed previously, the slopes of this relationship scaled with sampling density, and the more sparsely sampled schemes were more severely affected by low inter-scan reliability in the full DSI scheme. These results emphasize that most bundles can be segmented by CS-DSI schemes with very high accuracies.

3.4 | The inter-scan reliability of bundles from CS-DSI is comparable to full DSI

Having demonstrated that CS-DSI schemes could segment bundles accurately, we next asked whether the CS-DSI segmented bundles were as reliable between scan sessions as bundles generated from the full DSI scheme. For each diffusion scheme, we calculated the inter-scan reliability for each bundle as the set of Dice scores between each pair of scan sessions within a participant (Table 2). We found that the inter-scan reliabilities across bundles were high for all CS-DSI schemes (medians ranging from 0.73 to 0.78). We then compared the distributions of inter-scan reliabilities for each CS-DSI scheme to the full DSI scheme across all participants within each bundle. We again found that the median difference

between these distributions scaled with sampling density (Figure 7a). However, these median differences were low (<.07, Table S5) and after correcting for multiple comparisons, no bundle showed a significant difference between CS-DSI and full DSI reliabilities for any scheme (p > .1, Table S5). These results demonstrate that CS-DSI is as reliable as full DSI for segmenting bundles.

Mirroring what we observed for bundle segmentation accuracy, we found that the inter-scan reliability of the Dice scores for all CS-DSI schemes was correlated with the inter-scan reliability of the full DSI scheme (Pearson R ranging from 0.80 to 0.94; p < .001) Figure 7b). Bundles that were most susceptible to noise between sessions in the full DSI segmentations were also less reliably segmented in the CS-DSI images over sessions. Again, the inter-scan reliability of sparsely sampled CS-DSI schemes (like HA-SC55 and RAND57) was more impacted in bundles that showed lower inter-scan reliability in the full DSI scheme.

3.5 | CS-DSI and full DSI produce highly correlated voxel-wise scalar maps in the same scan session

After showing that CS-DSI schemes studied could accurately and reliably segment white matter bundles, we next turned to voxel-wise scalar maps. We examined three different microstructural properties: NQA, GFA, and ISO (Yeh et al., 2013). We evaluated each of these voxel-wise scalars with the same validation metrics as done for bundle segmentation: same-scan accuracy, inter-scan accuracy, and inter-scan reliability (Table 2). We found that the CS-DSI derived voxel-wise scalars were all highly correlated with the full DSI derived voxel-wise-scalars, demonstrating high same-scan accuracy. Moreover, the distribution of same-scan accuracies of all the CS-DSI maps was comparable with the distribution of inter-scan reliability of the full DSI maps for all the scalars studied (Figure 8a–c). For GFA and ISO, all HA-SC images generated scalars with no significant difference between their same-scan accuracy and the full DSI inter-scan reliability distributions (Table S6). For NQA, there was no difference between the HA-SC55-1 image and full DSI. In contrast, HA-SC55-2 and RAND57 had significantly lower same-scan accuracies than full DSI inter-scan reliabilities, albeit with very low median differences (<.05). Interestingly, HA-SC92 + 55−1, HA-SC92 + 55−2, and HA-SC92 had significantly higher same-scan accuracy for NQA as compared to the full DSI inter-scan reliability (Figure 8a, Table S6), which again is likely because the CS-DSI schemes and full DSI schemes within a scan session sub-sampled the same input data. Together, these results indicate that CS-DSI and full DSI produce similar estimates of voxel-wise scalar maps within the same scan session.

3.6 | CS-DSI and full DSI produce similar estimates of voxel-wise scalar maps across scan sessions

As for the bundles, we next sought to further examine the accuracy of a scalar map of microstructure while controlling for the effect of scan

Retrospective inter-scan reliability
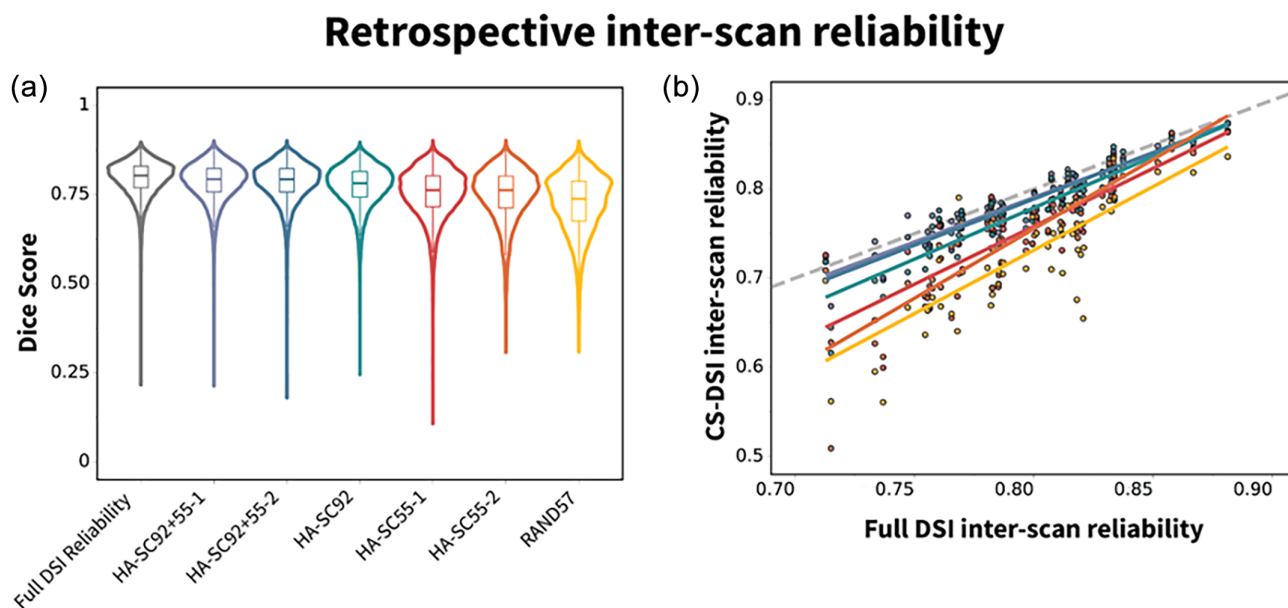
FIGURE 7 CS-DSI inter-scan reliability of bundle segmentation is comparable to and correlated with full DSI inter-scan reliability. (a) CS-DSI inter-scan reliability is comparable with full DSI inter-scan reliability. Violins represent distributions of Dice score across bundles and participants. (b) Median distribution of the inter-scan reliability for each CS-DSI scheme is highly correlated with the median distribution of the inter-scan reliability of the full DSI scheme. Here, each scatter dot represents the across-participant median of Dice scores for a single bundle. The gray dashed line denotes $x = y$.

session. Here, we defined inter-scan accuracy as the set of Pearson correlations between a CS-DSI derived voxel-wise scalar and the full DSI voxel-wise scalar in disparate sessions (Table 2). As expected, inter-scan accuracies were lower than the same-scan accuracies and scaled with the number and uniformity of directions sampled. These distributions were significantly different for all CS-DSI schemes and scalars examined ($p < .05$), but the median difference between the CS-DSI inter-scan accuracy and full DSI reliability was low for all HA-SC schemes (Figure 8d–f, Table S7). Together, these results demonstrate that CS-DSI schemes can generate whole brain voxel-wise scalars with high accuracy.

3.7 | The inter-scan reliability of CS-DSI derived scalars is very similar to full DSI

Next, we examined how the inter-scan reliability of the voxel-wise scalar maps generated by the CS-DSI images compared with that of the full DSI image. As prior, we defined inter-scan reliability as the set of Pearson R values between a voxel-wise scalar metric generated by each pair of scan sessions. As for white matter bundles, we found that the ability of a CS-DSI image to reliably estimate any of the scalar metrics decreased with decrease in q-space sampling density (Figure 8g–i). However, we found no significant differences in the inter-scan reliabilities between the CS-DSI derived scalars and the full DSI derived scalars ($p > .05$, Table S8). The median differences between these distributions were also very low across scalar metrics and diffusion schemes (<.05). Interestingly, we found that all the HA-

SC schemes had numerically (but not significantly) higher inter-scan reliabilities that the full DSI scheme for GFA, suggesting that strategic sub-sampling of q-space might be beneficial to reducing inter-scan noise in this metric. (Figure 8h). These results suggest that CS-DSI schemes can be used to rapidly measure local microstructural properties with a similar reliability as a full DSI scheme.

3.8 | CS-DSI accuracy can be replicated in an independent, prospectively acquired dataset

The results in the previous sections are from a dataset where synthetic CS-DSI data was generated retrospectively: full DSI schemes were collected, and CS-DSI schemes were subsampled from the acquired data. Furthermore, because these CS-DSI images were sub-sampled after preprocessing the full DSI image, the CS-DSI volumes benefitted from better motion correction as they could rely on points in q-space that were not a part of the synthesized subsample. We next asked whether prospectively acquired CS-DSI data without this advantage could generate bundles and scalars at similar levels of accuracy. Note that in the prospective dataset, only a single session was acquired for each CS scheme, so we could not evaluate inter-scan accuracy or inter-scan reliability of the CS-DSI schemes. For this dataset, the full DSI image was constructed by concatenating the four acquired CS-DSI images (HA-SC92, HA-SC55-1, HA-SC55-2, and RAND57). Same-scan accuracy was defined as the relationship between the CS-DSI derivative and the concatenated full DSI derivative for each scheme. Critically, we replicated the finding that all CS-
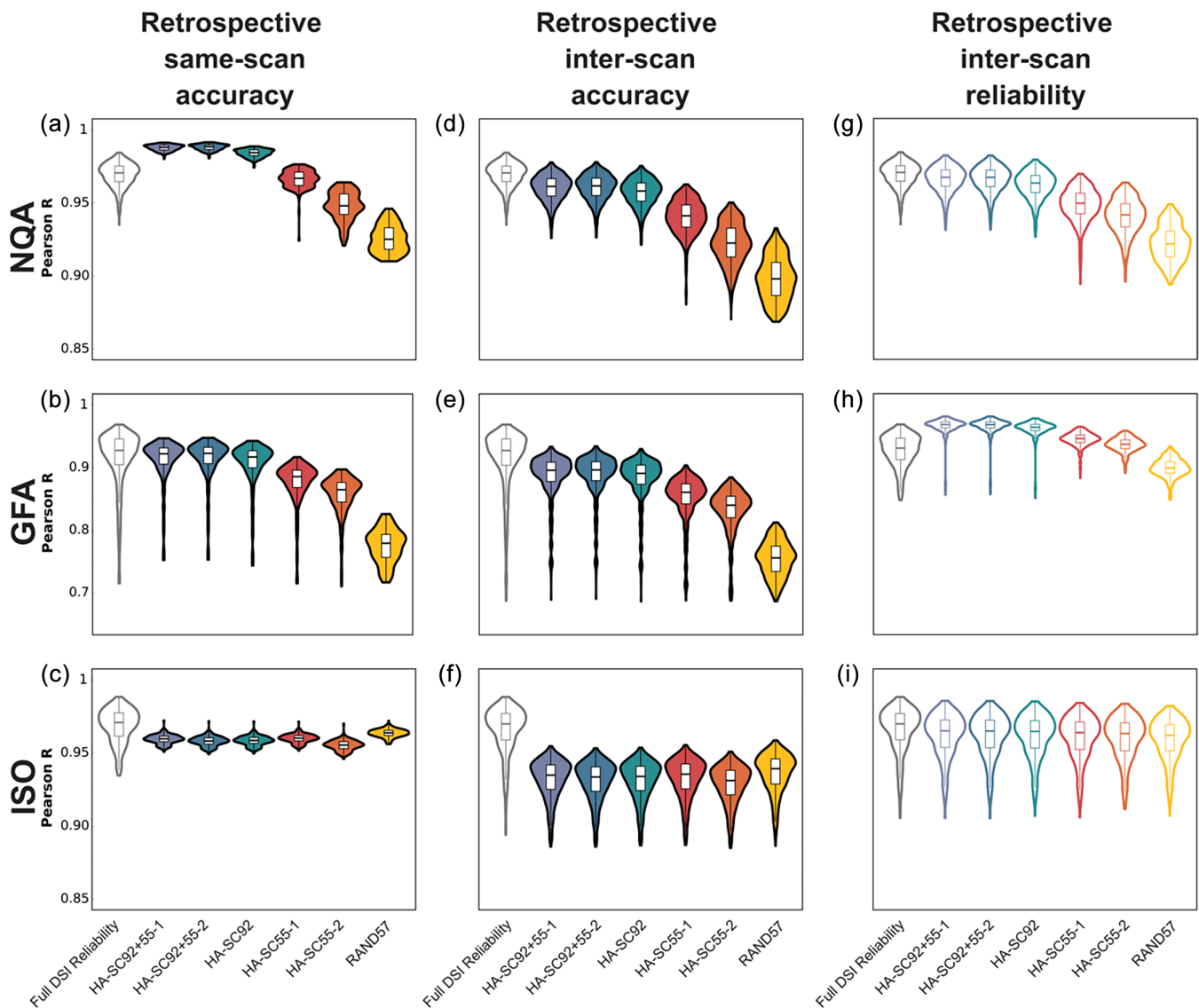
FIGURE 8 Comparing full DSI reliability with CS-DSI same-scan accuracy (a–c), inter-scan accuracy (d–f), and inter-scan reliability (g–i) when deriving whole-brain voxel-wise scalar maps. 3% of observed outlier values were excluded from each plot for clarity.

DSI schemes had high same-scan accuracies across bundles (medians ranged from 0.74 to 0.80; Figure 9a). Furthermore, we replicated the result that bundle segmentation accuracy was highly correlated with the between-scan reliability, again using the full DSI reliability from the retrospective dataset as a reference (Pearson R ranging from 0.76 to 0.81, *p* < .001, Figure 9b).

Finally, we evaluated the accuracy of voxel-wise NQA, GFA, and ISO maps in the prospective data. We found that the maps derived from CS-DSI images were all highly correlated with maps derived from the full DSI image. Here too, the distribution of Pearson correlation coefficients between the CS-DSI derived scalars and the full DSI derived scalars was comparable with inter-scan reliability of the full DSI scheme (from the retrospective data) for almost all scalars studied (Figure 10). This replication of our results in an independent dataset that was prospectively acquired on a different scanner further bolster confidence in our findings and underscore the utility of CS-DSI.

Pearson correlations of the voxel-wise scalars between the CS schemes and the combined scheme in the prospective dataset are similar to the full DSI inter-scan reliability from the retrospective data. 3% of observed outlier values were excluded from each plot for clarity.

4 | DISCUSSION

CS-DSI has been shown to be a promising strategy to reduce DSI acquisition time in post-mortem and animal studies. We examined the capacity for six different CS-DSI schemes to noninvasively quantify macrostructural and microstructural white matter properties in the human brain. We found that both CS-DSI derived bundle segmentations and voxel-wise scalars were comparable in accuracy and reliability to those derived from full DSI images. Our results demonstrate

Prospective same-scan accuracy
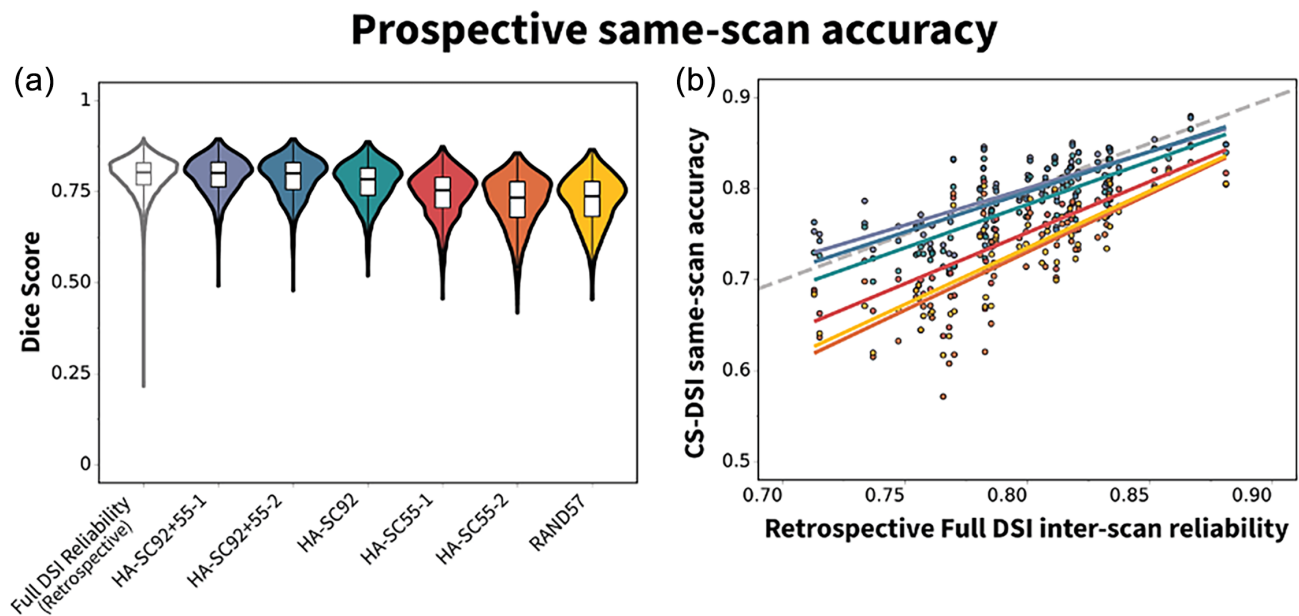
FIGURE 9 CS-DSI accuracy of bundle segmentation can be replicated in a dataset where CS-DSI images are prospectively acquired. (a) Dice scores between most CS-segmented bundles and combined DSI-segmented bundles are similar to the inter-scan reliability of full DSI bundles from the retrospective data. Violins represent distributions of Dice scores across all participants and bundles for a given scheme. (b) Median distribution of the accuracy for each CS-DSI scheme is highly correlated with the median distribution of the full DSI inter-scan reliability from the retrospective data. Here, each point on the scatter plot represents the median of Dice scores for a single bundle across participants. The gray dashed lined denotes $x = y$.

Prospective same-scan accuracy
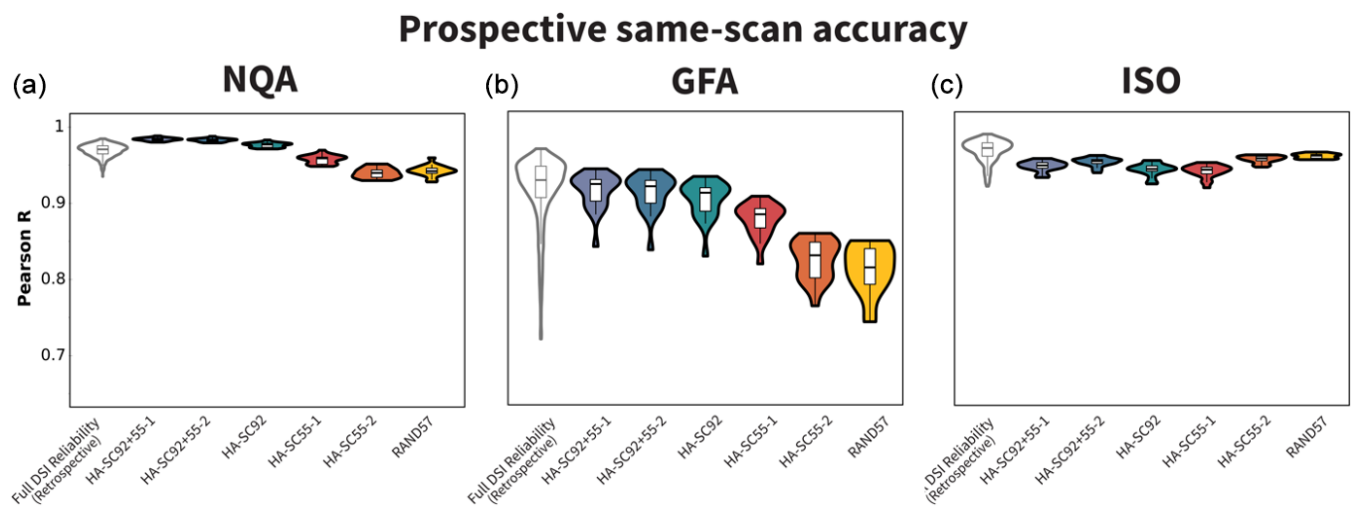
FIGURE 10 CS-DSI accuracy of scalar map generation can be replicated in a dataset where CS-DSI data is prospectively acquired.

that CS-DSI is an efficient and robust diffusion imaging technique that can accurately and reliably quantify white matter in a fraction of the acquisition time required by a full DSI scan, underscoring its clinical and research potential.

There are many approaches to construct CS schemes for dMRI. In this paper, we primarily examined HA-SC schemes with different q-space sampling densities, and a random scheme (RAND57) designed to enable concatenation to the full DSI scheme. The HA-SC scheme ensured a homogenous angular covering of the q-space and had been previously shown to provide the most efficient and robust

reconstruction of the EAP in simulated data (Paquette et al., 2015). We observed that all HA-SC schemes performed better than the RAND57 scheme, possibly because the homogenous distribution of subsampled points allowed for more effective extrapolation of the unacquired points in q-space. We also consistently observed that HA-SC accuracy and reliability scaled with the number of directions sampled, with the longer, more densely sampled schemes performing better.

To further contextualize the performance of our accuracy measures, we compared them against the inter-scan reliability of the full

DSI scheme. Inter-scan reliability reflects test–retest differences, which are inevitable in neuroimaging, even when using longer acquisition schemes. We reasoned that if differences between CS-DSI derivatives and full DSI derivatives were similar to the test–retest differences in full DSI derivatives, the corresponding CS-DSI scans could be considered as viable alternatives to full DSI. We showed that CS-DSI schemes yielded similar reliability to the full DSI scheme for both bundle segmentations and scalar quantification. Impressively, while the HA-SC92 scheme took only around a third of the acquisition time required for full DSI, this reduction in acquisition time had no negative impact on its data quality: HA-SC92 accuracies and reliabilities were not statistically different from full DSI reliabilities for almost all derivatives examined. Though HA-SC92 + 55–1 and HA-SC92 + 55–2 also performed very well, the HA-SC92 scheme had an acquisition time of only 7.4 minutes. These results suggest that the information lost from subsampling the q-space in the denser HA-SC schemes was minimal and comparable with mere test–retest differences of the full DSI. This key finding demonstrates that the HA-SC92 scheme can be successfully used as an alternative for full DSI with very few trade-offs for the derived measures we examined.

While the more sparsely sampled schemes with shorter scan times differed significantly from the full DSI scheme, we found that these differences were very small, especially for the HA-SC55-1 and HA-SC55-2 schemes. All bundles segmented by HA-SC55-1 and HA-SC55-2 were still qualitatively comparable to those segmented by full DSI. Moreover, for bundles that were more reliably segmented by full DSI (like the Corpus Callosum or the Parietal Aslant), even the more sparsely sampled schemes were not significantly less reliable than the full DSI scheme. These results suggest that scientific questions focused on these specific bundles could safely adopt a 4-minute CS-DSI scan. This feature of the sparser CS-DSI schemes is particularly valuable in contexts where scan time needs to be as low as possible, like in clinical care or in research of populations where extended scans are impractical (e.g., children, clinical populations).

These differences in HA-SC and full DSI were even smaller when computing the voxel-wise scalars. The median Pearson correlation between HA-SC and full DSI scalar maps was consistently above 0.85 for all scalars examined. This finding is particularly promising as it suggests that CS-DSI has the capacity to successfully detect specific and localized microstructural changes. The high correlation between HA-SC and full DSI scalars suggests that factors that are associated with full DSI scalars would also be associated with HA-SC scalars, allowing us to detect the same properties in a fraction of the scan time. Future studies correlating HA-SC derivatives with factors of interest like age, cognition, or clinical status, and comparing these relationships with those from full DSI will be able to further ascertain CS-DSI's utility in these contexts.

Despite being slightly longer, the RAND57 scheme did not perform as well as any of the HA-SC schemes and showed significant differences compared to the full DSI in almost all measures. These findings indicate that not all CS-DSI schemes are equivalent, and that q-space subsampling must be performed strategically to be effective. This observation also highlights the significance of our HA-SC results:

while not all CS-DSI schemes are effective, we identified a sampling strategy that can approximate full DSI performance in a fraction of the scan time.

Not only can HA-SC schemes be an effective alternative to full DSI but they might also be able to improve all studies leveraging diffusion imaging. Single shell diffusion schemes continue to be the most popular DWI acquisitions, especially in clinical workflows, due in large part to their short scan times. The HA-SC55 schemes have similar acquisition times as a traditional single shell scheme required to generate tensors (Soares et al., 2013). However, existing work suggests that DSI schemes can resolve crossing fibers and quantify more specific microstructural properties. A recent study even showed that CS-DSI was better than DTI at visualizing motor and language white matter tracks in patients with brain tumors (Young et al., 2017), further underscoring its clinical potential. Moreover, CS-DSI schemes can be effectively resampled into other HARDI schemes (Jones et al., 2021a), enabling analyses that require multishelled data like NODDI (Zhang et al., 2012) or SANDI (Palombo et al., 2020). Simultaneously, unlike multishelled schemes, CS-DSI allows direct reconstruction of the full EAP (Jones et al., 2021a), resulting in improved tractography (Yendiki et al., 2021) and better gray-white contrast (Jones et al., 2021b). While a comparison between CS-DSI and multishell schemes is beyond the scope of this work, future studies comparing the sensitivity and reliability of CS-DSI derivatives with those derived from multishelled schemes would be highly valuable.

This study has several limitations that should be acknowledged. First, the retrospective CS-DSI images were subsampled from the full DSI dataset *after* preprocessing. Within each scan session, the motion correction on the subsampled volumes could also leverage from volumes that were not a part of the examined scheme, giving the retrospectively synthesized images an advantage not available to prospectively acquired CS-DSI data. Given this advantage, along with the absence of inter-scan noise, it is unsurprising that same-scan accuracy was higher than full DSI inter-scan reliability for some of the denser CS-DSI schemes. We addressed this problem by both calculating inter-scan accuracy, as well as replicating these results in prospectively acquired CS-DSI data. Moreover, while our results demonstrate minimal differences between CS-DSI and full DSI derivatives, we cannot gauge the practical consequences of these differences. While beyond the scope of this paper, future studies examining the relationship between subsampling and correlations with metrics of interest (like age, cognition, or pathological outcomes) will help answer this question. Third, we used SHORE basis functions with L2-regularization to perform CS. While L2-regularization has been shown to be an effective way to extrapolate the full DSI signal from the CS-DSI data (Merlet & Deriche, 2010, 2013), there are many other CS-DSI strategies that use other approaches (discrete cosine transform, discrete wavelet transform; Almasri et al., 2020; Hong et al., 2011; Lai et al., 2016). Recent studies have demonstrated that adaptive dictionary-based approaches can substantially improve reconstruction (Bilgic et al., 2013, 2012; Jones et al., 2021a). While some of these improved methods may require more computational resources, they have the potential to further increase the bundle and scalar derivation validity of even the sparser CS-DSI schemes. Investigating these algorithms using

the framework presented here may further improve white matter measures derived from CS-DSI acquisitions. Fourth, it should also be noted that different versions of QSIPrep were used for preprocessing the retrospective (v0.4.0) and prospective (v0.8.0) datasets. While there are no major pipeline differences between these versions, it is possible that they might produce slightly discrepant derivatives (Schilling et al., 2021a, 2021b). However, as all our validation metrics were only calculated within datasets (i.e., Dice scores or Pearson correlations were only calculated internally within the prospective or retrospective dataset), any subtle differences of pipeline version are unlikely to impact the results. Finally, not all bundles were successfully segmented across all participants, scan sessions, and schemes. While these failures were not restricted to CS-DSI schemes, future studies evaluating probabilistic tracking algorithms could help characterize the effect of sparse q-space sampling on these missing bundles. Moreover, while other scalars are highly correlated with the GQI scalars we evaluated, future studies examining whether these CS-DSI scans can produce comparably reliable measures like DTI FA and MD will be valuable.

In conclusion, we demonstrated that CS-DSI can provide an accurate and reliable characterization of white matter properties in human brains in vivo, in a fraction of the scan time required for full DSI. These results bolster confidence in CS-DSI and will accelerate its adoption moving forward.

CONFLICT OF INTEREST STATEMENT

The data in this manuscript were collected in compliance with ethical standards. These data have not been published previously and are not under consideration for publication elsewhere. All authors have contributed significantly to this manuscript and have approved this submission. All authors have no conflicts of interests in the conduct or reporting of this research.

DATA AVAILABILITY STATEMENT

The prospectively acquired CS-DSI dataset is available in OpenNeuro for open access under a CC0 license at https://openneuro.org/datasets/ds004737/versions/1.0.0 (doi:10.18112/openneuro.ds004737.v1.0.0, reference number ds004737). The acquisition schemes tested in this manuscript are available for download at https://github.com/PennLINC/cs_dsi/tree/main/b-tables.

ORCID

Theodore D. Satterthwaite https://orcid.org/0000-0001-7072-9399

REFERENCES

Adluru, G., & Dibella, E. V. R. (2008). Reordering for improved constrained reconstruction from undersampled k-space data. *International Journal of Biomedical Imaging*, *2008*, 341684. https://doi.org/10.1155/2008/341684

Afzali, M., Pieciak, T., Newman, S., Garyfallidis, E., Özarslan, E., Cheng, H., & Jones, D. K. (2021). The sensitivity of diffusion MRI to microstructural properties and experimental factors. *Journal of Neuroscience Methods*, *347*, 108951. https://doi.org/10.1016/j.jneumeth.2020.108951

Almasri, N., Sadhu, A., & Ray Chaudhuri, S. (2020). Toward compressed sensing of structural monitoring data using discrete cosine transform. *Journal of Computing in Civil Engineering*, *34*, 04019041. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000855

Assaf, Y., & Basser, P. J. (2005). Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *Neuroimage*, *27*(1), 48–58. https://doi.org/10.1016/j.neuroimage.2005.03.042

Assaf, Y., Ben-Bashat, D., Chapman, J., Peled, S., Biton, I. E., Kafri, M., Segev, Y., Hendler, T., Korczyn, A. D., Graif, M., & Cohen, Y. (2002). High b-value q-space analyzed diffusion-weighted MRI: application to multiple sclerosis. *Magnetic Resonance in Medicine*, *47*(1), 115–126. https://doi.org/10.1002/mrm.10040

Assaf, Y., Freidlin, R. Z., Rohde, G. K., & Basser, P. J. (2004). New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, *52*(5), 965–978. https://doi.org/10.1002/mrm.20274

Baraniuk, R., & Steeghs, P. (2007). Compressive radar imaging, in: 2007 IEEE radar conference. Presented at the 2007 IEEE radar conference, pp. 128–133. https://doi.org/10.1109/RADAR.2007.374203

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*(1), 259–267. https://doi.org/10.1016/S0006-3495(94)80775-1

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*, 289–300.

Berger, C. R., Zhou, S., Preisig, J. C., & Willett, P. (2010). Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing. *IEEE Transactions on Signal Processing*, *58*, 1708–1721. https://doi.org/10.1109/TSP.2009.2038424

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*, 2022–2033. https://doi.org/10.1109/TMI.2013.2271707

Bilgic, B., Setsompop, K., Cohen-Adad, J., Wedeen, V., Wald, L. L., & Adalsteinsson, E. (2012). Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries. In *Magnetic Resonance in Medicine*. Wiley Periodicals, Inc. (Vol. 15, pp. 1–9).

Bobin, J., & Starck, J.-L. (2009). Compressed sensing in astronomy and remote sensing: A data fusion perspective. In V. K. Goyal, M. Papadakis, & D. Van De Ville (Eds.), *Proc. SPIE 7446, Wavelets XIII*, (p. 74460I). SPIE. https://doi.org/10.1117/12.830633

Callaghan, P. T., Eccles, C. D., & Xia, Y. (1988). NMR microscopy of dynamic displacements: k-space and q-space imaging. *Journal of Physics*, *21*, 820–822. https://doi.org/10.1088/0022-3735/21/8/017

Candes, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, *25*, 21–30. https://doi.org/10.1109/MSP.2007.914731

Carroll, M. K., Cecchi, G. A., Rish, I., Garg, R., & Rao, A. R. (2009). Prediction and interpretation of distributed neural activity with sparse models.

NeuroImage, *44*, 112–122. https://doi.org/10.1016/j.neuroimage.2008.08.020

Caruyer, E., Lenglet, C., Sapiro, G., & Deriche, R. (2013). Design of multi-shell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine*, *69*(6), 1534–1540. https://doi.org/10.1002/mrm.24736

Cheng, J., Merlet, S., Caruyer, E., Ghosh, A., Jiang, T., & Deriche, R. (2011). Compressive sensing ensemble average propagator estimation via L1 spherical polar Fourier imaging. Presented at the MICCAI workshop on computational diffusion MRI – CDMRI'11.

Cieslak, M., Cook, P. A., He, X., Yeh, F. C., Dhollander, T., Adebimpe, A., Aguirre, G. K., Bassett, D. S., Betzel, R. F., Bourque, J., Cabral, L. M., Davatzikos, C., Detre, J. A., Earl, E., Elliott, M. A., Fadnavis, S., Fair, D. A., Foran, W., Fotiadis, P., ... Satterthwaite, T. D. (2021). QSI-Prep: An integrative platform for preprocessing and reconstructing diffusion MRI data. *Nature Methods*, *18*, 775–778. https://doi.org/10.1038/s41592-021-01185-5

Cieslak, M., Cook, P. A., Tapera, T. M., Radhakrishnan, H., Elliott, M., Roalf, D. R., Oathes, D. J., Bassett, D. S., Tisdall, M. D., Rokem, A., Grafton, S. T., & Satterthwaite, T. D. (2022). Diffusion MRI head motion correction methods are highly accurate but impacted by denoising and sampling scheme. *bioRxiv*. https://doi.org/10.1101/2022.07.21.500865

Cieslak, M., Meiring, W., Brennan, T., Greene, C., Volz, L. J., Vettel, J. M., Suri, S., & Grafton, S. T. (2018). Compositional measures of diffusion anisotropy and asymmetry. Presented at the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 123–126. https://doi.org/10.1109/ISBI.2018.8363537

Daducci, A., Canales-Rodríguez, E. J., Descoteaux, M., Garyfallidis, E., Gur, Y., Lin, Y. C., Mani, M., Merlet, S., Paquette, M., Ramirez-Manzanares, A., Reisert, M., Reis Rodrigues, P., Sepehrband, F., Caruyer, E., Choupan, J., Deriche, R., Jacob, M., Menegaz, G., Pričkovska, V., ... Thiran, J. P. (2014). Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion MRI. *IEEE Transactions on Medical Imaging*, *33*(2), 384–399. https://doi.org/10.1109/TMI.2013.2285500

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, *52*, 1289–1306. https://doi.org/10.1109/TIT.2006.871582

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

Gorgolewski, K., Burns, C., Madison, C., Clark, D., Halchenko, Y., Waskom, M., & Ghosh, S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*, 13.

Gorgolewski, K. J., Esteban, O., Ellis, D. G., Notter, M. P., Ziegler, E., Johnson, H., Hamalainen, C., Yvernault, B., Burns, C., Manhães-Savio, A., Jarecka, D., Markiewicz, C. J., Salo, T., Clark, D., Waskom, M., Wong, J., Modat, M., Dewey, B. E., Clark, M. G., ... Ghosh, S. (2017). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. 0.13.1. https://doi.org/10.5281/zenodo.581704

Gramfort, A., Poupon, C., & Descoteaux, M. (2012). Sparse DSI: Learning DSI structure for denoising and fast imaging. In N. Ayache, H. Delingette, P. Golland, & K. Mori (Eds.), *Medical image computing and computer-assisted intervention – MICCAI 2012, Lecture Notes in Computer Science* (pp. 288–296). Springer. https://doi.org/10.1007/978-3-642-33418-4_36

Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., Medaglia, J. D., Vettel, J. M., Miller, M. B., Grafton, S. T., & Bassett, D. S. (2015 Oct). Controllability of structural brain networks. *Nature Communications*, *1*(6), 8414. https://doi.org/10.1038/ncomms9414

Han, C., Zhao, L., Zhong, S., Wu, X., Guo, J., Zhuang, X., & Han, H. (2015). A comparison of high b-value vs standard b-value diffusion-weighted magnetic resonance imaging at 3.0 T for medulloblastomas. *British Journal of Radiology*, *88*(1054), 20150220. https://doi.org/10.1259/bjr.20150220

He, L., Chang, T.-C., Osher, S., Fang, T., & Speier, P. (2007). MR image reconstruction from undersampled data by using the iterative refinement procedure. *PAMM*, *7*, 1011207. https://doi.org/10.1002/pamm.200700776

Herman, M. A., & Strohmer, T. (2009). High-resolution radar via compressed sensing. *IEEE Transactions on Signal Processing*, *57*, 2275–2284. https://doi.org/10.1109/TSP.2009.2014277

Hong, M., Yu, Y., Wang, H., Liu, F., & Crozier, S. (2011). Compressed sensing MRI with singular value decomposition-based sparsity basis. *Physics in Medicine and Biology*, *56*, 6311–6325. https://doi.org/10.1088/0031-9155/56/19/010

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. *Eur Radiol*, *26*(9), 2992–2998. https://doi.org/10.1007/s00330-015-4162-9

Jaspan, O. N., Fleysher, R., & Lipton, M. L. (2015). Compressed sensing MRI: A review of the clinical literature. *The British Journal of Radiology*, *88*, 20150487. https://doi.org/10.1259/bjr.20150487

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, *53*(6), 1432–1440. https://doi.org/10.1002/mrm.20508

Jones, D. K., Horsfield, M. A., & Simmons, A. (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, *42*, 515–525.

Jones, R., Maffei, C., Augustinack, J., Fischl, B., Wang, H., Bilgic, B., & Yendiki, A. (2021a). High-fidelity approximation of grid- and shell-based sampling schemes from undersampled DSI using compressed sensing: Post mortem validation. *Neuroimage*, *1*(244), 118621. https://doi.org/10.1016/j.neuroimage.2021.118621

Jones, R., Tian, Q., Maffei, C., Augustinack, J., Nummenmaa, A., Huang, S., & Yendiki, A. (2021b). *Generalized anisotropy profiles distinguish cortical and subcortical structures in ex vivo diffusion MRI*. Proc Intl Soc Mag Res Med.

Lai, Z., Qu, X., Liu, Y., Guo, D., Ye, J., Zhan, Z., & Chen, Z. (2016). Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform. *Medical Image Analysis Discrete Graphical Models in Biomedical Image Analysis*, *27*, 93–104. https://doi.org/10.1016/j.media.2015.05.012

Lustig, M., Donoho, D., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, *58*, 1182–1195. https://doi.org/10.1002/mrm.21391

Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed sensing MRI. *IEEE Signal Processing Magazine*, *25*, 72–82. https://doi.org/10.1109/MSP.2007.914728

Maffei, C., Lee, C., Planich, M., Ramprasad, M., Ravi, N., Trainor, D., Urban, Z., Kim, M., Jones, R. J., Henin, A., Hofmann, S. G., Pizzagalli, D. A., Auerbach, R. P., Gabrieli, J. D. E., Whitfield-Gabrieli, S., Greve, D. N., Haber, S. N., & Yendiki, A. (2021). Using diffusion MRI data acquired with ultra-high gradient strength to improve tractography in routine-quality data. *Neuroimage*, *15*(245), 118706. https://doi.org/10.1016/j.neuroimage.2021.118706

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*, 1226–1233. https://doi.org/10.1002/mrm.23064

Merlet, S. (2013). Compressive sensing in diffusion MRI. Medical Imaging. Université Nice Sophia Antipolis, English. ffNNT: ff. fftel-00908369v1f.

Merlet, S., & Deriche, R. (2010). *Compressed sensing for accelerated EAP recovery in diffusion MRI*. MICCAI p. 14. ffinria-00536278f.

Merlet, S. L., & Deriche, R. (2013). Continuous diffusion signal, EAP and ODF estimation via compressive sensing in diffusion MRI. *Medical Image Analysis*, *17*, 556–572. https://doi.org/10.1016/j.media.2013.02.010

Michailovich, O., Rathi, Y., & Dolui, S. (2011). Spatially regularized compressed sensing for high angular resolution diffusion imaging. *IEEE Transactions on Medical Imaging*, *30*, 1100–1115. https://doi.org/10.1109/TMI.2011.2142189

Muldoon, S. F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S. T., Vettel, J. M., & Bassett, D. S. (2016). Stimulation-Based Control of Dynamic Brain Networks. *PLOS Computational Biology*, *12*(9), e1005076. https://doi.org/10.1371/journal.pcbi.1005076

Naeyaert, M., Aelterman, J., Van Audekerke, J., Golkov, V., Cremers, D., Pižurica, A., Sijbers, J., & Verhoye, M. (2021). Accelerating in vivo fast spin echo high angular resolution diffusion imaging with an isotropic resolution in mice through compressed sensing. *Magnetic Resonance in Medicine*, *85*, 1397–1413. https://doi.org/10.1002/mrm.28520

Nakuci, J., Wasylyshyn, N., Cieslak, M., Elliot, J. C., Bansal, K., Giesbrecht, B., Grafton, S. T., Vettel, J. M., Garcia, J. O., & Muldoon, S. F. (2023). Within-subject reproducibility varies in multimodal, longitudinal brain networks. *Scientific Reports*, *13*, 6699. https://doi.org/10.1101/2022.05.03.490544

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magnetic Resonance in Medicine*, *79*(6), 3172–3193. https://doi.org/10.1002/mrm.27101

Ozarslan, E., Koay, C., Shepherd, T. M., Blackb, S. J., & Basser, P. J. (2013). Simple harmonic oscillator based reconstruction and estimation for three-dimensional q-space MRI.

Palombo, M., Ianus, A., Guerreri, M., Nunes, D., Alexander, D. C., Shemesh, N., & Zhang, H. (2020). SANDI: A compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. *Neuroimage*, *215*, 116835. https://doi.org/10.1016/j.neuroimage.2020.116835 Erratum in: Neuroimage, 2021 Feb 1, 226, 117612.

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*, 401–416. https://doi.org/10.1002/mrm.25093

Pu, L., Trouard, T. P., Ryan, L., Huang, C., Altbach, M. I., & Bilgin, A. (2011). Model-based compressive diffusion tensor imaging, in: 2011 Presented at the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 254–257. https://doi.org/10.1109/ISBI.2011.5872400

Raffelt, D. A., Tournier, J. D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage*, *144*(Pt A), 58–73. https://doi.org/10.1016/j.neuroimage.2016.09.029

Ramirez-Manzanares, A., Rivera, M., Vemuri, B. C., Carney, P., & Mareci, T. (2007). Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. *IEEE Transactions on Medical Imaging*, *26*, 1091–1102. https://doi.org/10.1109/TMI.2007.900461

Schilling, K. G., Nath, V., Blaber, J., Harrigan, R. L., Ding, Z., Anderson, A. W., & Landman, B. A. (2017). Effects of b-value and number of gradient directions on diffusion MRI measures obtained with Q-ball Imaging. *Proceedings of SPIE - The International Society for Optical Engineering*, *10133*, 101330N. https://doi.org/10.1117/12.2254545

Schilling, K. G., Rheault, F., Petit, L., Hansen, C. B., Nath, V., Yeh, F. C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., Fischi-Gomez, E., Pizzolato, M., Ocampo-Pineda, M., Schiavi, S., Canales-Rodríguez, E. J., Daducci, A., Granziera, C., Innocenti, G., Thiran, J. P., ... Descoteaux, M. (2021a). Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *Neuroimage*, *243*, 118502. https://doi.org/10.1016/j.neuroimage.2021.118502

Schilling, K. G., Tax, C. M. W., Rheault, F., Hansen, C., Yang, Q., Yeh, F. C., Cai, L., Anderson, A. W., & Landman, B. A. (2021b). Fiber tractography bundle segmentation depends on scanner effects, vendor effects, acquisition resolution, diffusion sampling scheme, diffusion sensitization, and bundle segmentation workflow. *Neuroimage*, *15*(242), 118451. https://doi.org/10.1016/j.neuroimage.2021.118451

Schilling, K. G., Yeh, F.-C., Nath, V., Hansen, C., Williams, O., Resnick, S., Anderson, A. W., & Landman, B. A. (2019). A fiber coherence index for quality control of B-table orientation in diffusion MRI scans. *Magnetic Resonance Imaging*, *58*, 82–89. https://doi.org/10.1016/j.mri.2019.01.018

Soares, J., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, *7*, 31.

Teich, E. G., Cieslak, M., Giesbrecht, B., Vettel, J M., Grafton, S. T., Satterthwaite, T. D., & Bassett, D. S. (2021). Crystallinity characterization of white matter in the human brain. *New Journal of Physics*, *23*(7), 073047. https://doi.org/10.1088/1367-2630/ac1286

Tristán-Vega, A., Pieciak, T., París, G., Rodríguez-Galván, J. R., & Aja-Fernández, S. (2023). HYDI-DSI revisited: Constrained non-parametric EAP imaging without q-space re-gridding. *Medical Image Analysis*, *84*, 102728. https://doi.org/10.1016/j.media.2022.102728

Tristán-Vega, A., & Westin, C.-F. (2011). Probabilistic ODF estimation from reduced HARDI data with sparse regularization. In G. Fichtinger, A. Martel, & T. Peters (Eds.), *Medical image computing and computer-assisted intervention – MICCAI 2011, Lecture Notes in Computer Science* (pp. 182–190). Springer. https://doi.org/10.1007/978-3-642-23629-7_23

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, *52*(6), 1358–1372. https://doi.org/10.1002/mrm.20279

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, *48*, 577–582. https://doi.org/10.1002/mrm.10268

Wang, C., Holly, L. T., Oughourlian, T., Yao, J., Raymond, C., Salamon, N., & Ellingson, B. M. (2021). Detection of cerebral reorganization associated with degenerative cervical myelopathy using diffusion spectral imaging (DSI). *Journal of Clinical Neuroscience*, *86*, 164–173. https://doi.org/10.1016/j.jocn.2021.01.011

Wang, N., White, L. E., Qi, Y., Cofer, G., & Johnson, G. A. (2020). Cytoarchitecture of the mouse brain by high resolution diffusion magnetic resonance imaging. *NeuroImage*, *216*, 116876. https://doi.org/10.1016/j.neuroimage.2020.116876

Wedeen, V. J., Hagmann, P., Tseng, W. Y. I., Reese, T. G., & Weisskoff, R. M. (2005). Mapping complex tissue architecture with DSI magnetic resonance imaging. *Magnetic Resonance in Medicine*, *54*, 1377–1385.

Wedeen, V. J., Wang, R. P., Schmahmann, J. D., Benner, T., Tseng, W. Y. I., Dai, G., Pandya, D. N., Hagmann, P., D'Arceuil, H., & de Crespigny, A. J. (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage*, *41*, 1267–1277. https://doi.org/10.1016/j.neuroimage.2008.03.036

Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., & Vandergheynst, P. (2009). Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, *395*, 1733–1742. https://doi.org/10.1111/j.1365-2966.2009.14665.x

Wu, Y.-C., Field, A. S., & Alexander, A. L. (2008). Computation of diffusion function measures in q -space using magnetic resonance hybrid diffusion imaging. *IEEE Transactions on Medical Imaging*, *27*(6), 858–865. https://doi.org/10.1109/TMI.2008.922696

Ye, W., Vemuri, B. C., & Entezari, A. (2012). An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators, in: 2012 9th IEEE international symposium on biomedical imaging (ISBI). *Presented at the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, *9*, 940–943. https://doi.org/10.1109/ISBI.2012.6235711

Yeh, F.-C. (2020). Shape analysis of the human association pathways. *NeuroImage*, *223*, 117329. https://doi.org/10.1016/j.neuroimage.2020.117329

Yeh, F.-C., Panesar, S., Barrios, J., Fernandes, D., Abhinav, K., Meola, A., & Fernandez-Miranda, J. C. (2019). Automatic removal of false connections in diffusion MRI tractography using topology-informed pruning (TIP). *Neurotherapeutics*, *16*, 52–58. https://doi.org/10.1007/s13311-018-0663-y

Yeh, F.-C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., Vettel, J. M., & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage*, *178*, 57–68. https://doi.org/10.1016/j.neuroimage.2018.05.027

Yeh, F.-C., Verstynen, T. D., Wang, Y., Fernández-Miranda, J. C., & Tseng, W.-Y. I. (2013). Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS One*, *8*, e80713. https://doi.org/10.1371/journal.pone.0080713

Yeh, F.-C., Vettel, J. M., Singh, A., Poczos, B., Grafton, S. T., Erickson, K. I., Tseng, W.-Y. I., & Verstynen, T. D. (2016). Differences between generalized q-sampling imaging and diffusion tensor imaging in the preoperative visualization of the nerve fiber tracts within peritumoral edema in brain. *PLoS Computational Biology*, *12*, e1005203. https://doi.org/10.1371/journal.pcbi.1005203

Yeh, F.-C., Wedeen, V. J., & Tseng, W.-Y. I. (2010). Generalized q-sampling imaging. *IEEE Transactions on Medical Imaging*, *29*, 1626–1635. https://doi.org/10.1109/TMI.2010.2045126

Yendiki, A., Aggarwal, M., Axer, M., Howard, A. F. D., van Walsum, A. V. C., & Haber, S. N. (2022). Post mortem mapping of connectional anatomy for the validation of diffusion MRI. *Neuroimage*, *(256)*, 119146. https://doi.org/10.1016/j.neuroimage.2022.119146

Yoshiura, T., Wu, O., Zaheer, A., Reese, T. G., & Sorensen, A. G. (2001). Highly diffusion-sensitized MRI of brain: Dissociation of gray and white matter. *Magnetic Resonance in Medicine*, *45*, 734–740. https://doi.org/10.1002/mrm.1100

Young, R. J., Tan, E. T., Peck, K. K., Jenabi, M., Karimi, S., Brennan, N., Rubel, J., Lyo, J., Shi, W., Zhang, Z., Prastawa, M., Liu, X., Sperl, J. I., Fatovic, R., Marinelli, L., & Holodny, A. I. (2017). Comparison of compressed sensing diffusion spectrum imaging and diffusion tensor imaging in patients with intracranial masses. *Magnetic Resonance Imaging*, *36*, 24–31. https://doi.org/10.1016/j.mri.2016.10.001

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, *61*(4), 1000–1016. https://doi.org/10.1016/j.neuroimage.2012.03.072

Zhang, H., Wang, Y., Lu, T., Qiu, B., Tang, Y., Ou, S., Tie, X., Sun, C., Xu, K., & Wang, Y. (2013). Differences between generalized q-sampling imaging and diffusion tensor imaging in the preoperative visualization of the nerve fiber tracts within peritumoral edema in brain. *Neurosurgery*, *73*, 1044–1053. https://doi.org/10.1227/NEU.0000000000000146

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Radhakrishnan, H., Zhao, C., Sydnor, V. J., Baller, E. B., Cook, P. A., Fair, D. A., Giesbrecht, B., Larsen, B., Murtha, K., Roalf, D. R., Rush-Goebel, S., Shinohara, R. T., Shou, H., Tisdall, M. D., Vettel, J. M., Grafton, S. T., Cieslak, M., & Satterthwaite, T. D. (2024). A practical evaluation of measures derived from compressed sensing diffusion spectrum imaging. *Human Brain Mapping*, *45*(5), e26580. https://doi.org/10.1002/hbm.26580