HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2022 November 11.

Published in final edited form as:

Neuroimage. 2021 December 01; 244: 118601. doi:10.1016/j.neuroimage.2021.118601.

Neural networks for parameter estimation in microstructural MRI: Application to a diffusion-relaxation model of white matter

João P. de Almeida Martins[a,b,1,*], **Markus Nilsson**[a,1], **Björn Lampinen**[c], **Marco Palombo**[d], **Peter T. While**[b,e], **Carl-Fredrik Westin**[f,g], **Filip Szczepankiewicz**[a,f,g]

[a]Department of Clinical Sciences, Radiology, Lund University, Lund, Sweden

[b]Department of Radiology and Nuclear Medicine, St. Olav's University Hospital, Trondheim, Norway

[c]Department of Clinical Sciences, Medical Radiation Physics, Lund University, Lund, Sweden

[d]Centre for Medical Image Computing and Department of Computer Science, University College London, London, United Kingdom

[e]Department of Circulation and Medical Imaging, NTNU-Norwegian University of Science and Technology, Trondheim, Norway

[f]Radiology, Brigham and Women's Hospital, Boston, MA, United States

[g]Harvard Medical School, Boston, MA, United States

Abstract

Specific features of white matter microstructure can be investigated by using biophysical models to interpret relaxation-diffusion MRI brain data. Although more intricate models have the potential to reveal more details of the tissue, they also incur time-consuming parameter estimation that may converge to inaccurate solutions due to a prevalence of local minima in a degenerate fitting landscape. Machine-learning fitting algorithms have been proposed to accelerate the parameter estimation and increase the robustness of the attained estimates. So

[*]Corresponding author at: Department of Clinical Sciences, Radiology, Lund University, Lund, Sweden. joao101martins@gmail.com (J.P. de Almeida Martins).

[1]These authors contributed equally to this work.

Credit authorship contribution statement

João P. de Almeida Martins: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Markus Nilsson:** Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Supervision, Funding acquisition. **Björn Lampinen:** Resources, Data curation, Writing – review & editing. **Marco Palombo:** Resources, Writing – review & editing, Super-vision. **Peter T. While:** Resources, Writing – review & editing, Funding acquisition. **Carl-Fredrik Westin:** Resources, Writing – review & editing, Funding acquisition. **Filip Szczepankiewicz:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition.

Code availability

MATLAB code for training and deploying the networks discussed in this work will be shared in open-source at https://github.com/filip-szczepankiewicz/deAlmeidaMartins_NIMG_2021. Trained networks discussed in this manuscript will be made available in the same repository.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.neuroimage.2021.118601.

far, learning-based fitting approaches have been restricted to microstructural models with a reduced number of independent model parameters where dense sets of training data are easy to generate. Moreover, the degree to which machine learning can alleviate the degeneracy problem is poorly understood. For conventional least-squares solvers, it has been shown that degeneracy can be avoided by acquisition with optimized relaxation-diffusion-correlation protocols that include tensor-valued diffusion encoding. Whether machine-learning techniques can offset these acquisition requirements remains to be tested. In this work, we employ artificial neural networks to vastly accelerate the parameter estimation for a recently introduced relaxation-diffusion model of white matter microstructure. We also develop strategies for assessing the accuracy and sensitivity of function fitting networks and use those strategies to explore the impact of the acquisition protocol. The developed learning-based fitting pipelines were tested on relaxation-diffusion data acquired with optimal and sub-optimal acquisition protocols. Networks trained with an optimized protocol were observed to provide accurate parameter estimates within short computational times. Comparing neural networks and least-squares solvers, we found the performance of the former to be less affected by sub-optimal protocols; however, model fitting networks were still susceptible to degeneracy issues and their use could not fully replace a careful design of the acquisition protocol.

1. Introduction

Microstructure imaging uses compartment modelling of diffusion MRI (dMRI) data with the aim to map specific tissue quantities (Alexander et al., 2019; Nilsson et al., 2013; Novikov et al., 2019). A central goal in microstructure imaging has been to estimate the volume fractions of different microstructural components such as axons (Lampinen et al., 2020, 2019; Veraart et al., 2018). Estimating volume fractions rather than signal fractions is challenging, however, because it requires the simultaneous estimation of both diffusion and relaxation properties of the different model compartments. This kind of inverse problem is sensitive to degeneracy issues (Jelescu et al., 2016; Lampinen et al., 2019), in which multiple sets of model parameters can describe the acquired data equally well. Parameter estimation can also be computationally slow, preventing real-time mapping. A potential solution is to employ machine learning to accelerate the parameter estimation process (Golkov et al., 2016). However, the current literature lacks systematic assessments of the advantages and drawbacks of this approach, which is surprising considering the exponential increase in interest for such methods.

Artificial neural networks (ANNs) and other machine learning approaches have been applied previously to accelerate the estimation of microstructure parameters from dMRI data (Barbieri et al., 2020; Bertleff et al., 2017; Golkov et al., 2016; Grussu et al., 2020; Gyori et al., 2019; Hill et al., 2021; Kaandorp et al., 2021; Nedjati-Gilani et al., 2017; Palombo et al., 2020; Reisert et al., 2017). For example, a random forest regressor has been used to fit a compartment model for white matter (WM) microstructure in the presence of water exchange (Nedjati-Gilani et al., 2017) and to fit the SANDI model for grey matter properties (Palombo et al., 2020). Reisert et al. (2017) applied machine learning to a Bayesian estimation approach which dramatically accelerated the fitting of two-and three-compartment models. Barbieri et al. (2020) applied ANNs to the intra-voxel incoherent

motion model. An important open question, however, is what impact the training strategy has on the fitting performance. This is particularly relevant when applied to non-linear multi-compartment models with many independent model parameters, which we here refer to loosely as 'high-dimensional models'. The generation of training data scales poorly with the number of model parameters, as sampling each combination of p model parameters in m steps requires m^p samples. As p increases, it is unavoidable that a finite set of samples becomes sparse in the p-dimensional model parameter space, risking selection bias. Here, we investigate the impact of different sampling patterns within this space on the performance of the neural network.

Apart from accelerating model fitting, neural networks may in principle also reduce the requirements on the imaging protocol by learning priors from training examples (Golkov et al., 2016). For example, neural networks have been used to learn a mapping between fully-sampled and sub-sampled datasets, which can in turn be used to stabilise model fitting performance against substantial degrees of data down-sampling (Alexander et al., 2017; Tian et al., 2020). However, we do not expect machine learning approaches to completely alleviate degeneracy issues. Indeed, for cases where the acquisition protocol does not provide sufficient information to resolve between different parameter values, the learning-based estimates will simply equal the mean of the model parameter distribution used for training (Reisert et al., 2017).

The aims of this study were to compare training strategies, to pro-pose tools to evaluate the performance of model fitting neural networks, and to test to what degree neural networks can solve problems with degeneracy. As a testbed, we use a high-dimensional relaxation-diffusion microstructure model of WM (Lampinen et al., 2020, 2019; Veraart et al., 2018). For this model, parameter estimation is enabled by state-of-the-art imaging protocols featuring so-called b-tensor encoding (Topgaard, 2017; Westin et al., 2016) combined with diffusion-relaxation correlations (de Almeida Martins et al., 2020; de Almeida Martins and Topgaard, 2018; Lampinen et al., 2019). We investigated the ability of neural networks to speed up model fitting, and explored the extent to which they can offset the requirements on the acquisition protocol.

2. Theory

White matter microstructure can be modelled by multiple compartments with different microstructural properties but a common orientation distribution (Alexander et al., 2019; Novikov et al., 2019). In this description, the measured signal is the convolution between an orientation distribution function (ODF) $P(\hat{\boldsymbol{n}})$ and a microstructural kernel $K(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}})$

$$S(\hat{\boldsymbol{u}}) = \int_{|\hat{\boldsymbol{n}}|=1} P(\hat{\boldsymbol{n}}) K(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}}) \mathrm{d}\hat{\boldsymbol{n}}, \quad (1)$$

where $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{u}}$ are unit vectors defining the symmetry axes of the ODF and of the diffusion encoding process, respectively. Note that the microstructural kernel depends on the relative angle between $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{u}}$, $\cos\beta = \hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}}$. In this work, we assign an effective transverse relaxation time T_2 and an apparent microscopic diffusion tensor $\mathbf{D}$ to each component, and

use exponentially decaying functions to model the effect of these microstructural properties on the relaxation-diffusion-weighted signal (Veraart et al., 2018). Under these assumptions, the microstructure kernel is written as a weighted sum of exponentials

$$K(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}}) = S_0\sum_{j=1}^{J} f_j \exp\left(-\mathbf{B}(\hat{\boldsymbol{u}}):\mathbf{D}_j(\hat{\boldsymbol{n}})\right)\exp\left(-\frac{\tau_{\mathrm{E}}}{T_{2;\,j}}\right), \tag{2}$$

corresponding to a mixture of J components with signal fraction f_j, transverse relaxation time $T_{2;\,j}$, and diffusion tensor $\mathbf{D}_j$. The colon ":" denotes the Frobenius inner product, $\mathbf{B}:\mathbf{D} = \sum_i \sum_j B_{ij}D_{ij}$. Information about $T_{2;\,j}$ and $\mathbf{D}_j$ is encoded into the signal by the echo time τ_{E} and diffusion encoding tensor $\mathbf{B}(\hat{\boldsymbol{u}})$, respectively, both of which are experimental variables. To simplify the model, we only consider axisymmetric $\mathbf{B}(\hat{\boldsymbol{u}})$ and additionally assume that the component-wise $\mathbf{D}_j$ are axisymmetric.

The convolution expressed in Eq. (1) can be simplified by factorizing both $P(\hat{\boldsymbol{n}})$ and $K(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}})$ in their spherical harmonic coefficients p_{lm} and k_{lm}, respectively:

$$P(\hat{\boldsymbol{n}}) = \sum_{l}\sum_{m} p_{lm}Y_{lm}(\hat{\boldsymbol{n}}), \tag{3}$$

and

$$K(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}}) = \sum_{l'} k_{l'0}Y_{l'0}(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}}), \tag{4}$$

where Y_{lm} are the spherical harmonics basis functions

$$Y_{lm}(\Theta,\Phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}L_l^m(\cos\Theta)\exp(im\Phi), \tag{5}$$

with the $L_l^m(x)$ term denoting the associated Legendre polynomials. The summations in Eqs. (3) are carried out for order $l = 0, 1, 2, \ldots$, and degree $m = -l, -l+1, \ldots, l$. In Eq. (4), we have taken the axial symmetry of the microstructural kernel $K(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}})$ into account (Lampinen et al., 2020; Novikov et al., 2018). Symmetry around the polar axis implies $k_{l'm'} = 0$ for either $m' \neq 0$ or odd l'. Taken together, this means that the $k_{l'm'}$ coefficients are reduced to their 0th degree terms $k_{l'0}$ (typically written as $k_{l'}$) and only even-ordered spherical harmonic terms ($l' = 0, 2, \ldots$) provide non-trivial contributions. Using the spherical harmonics addition theorem, Eq. (4) can be rewritten as

$$K(\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{n}}) = \sum_{l'} k_{l'0}\sum_{m'=-l'}^{l'} Y_{l'm'}(\hat{\boldsymbol{u}})\bar{Y}_{l'm'}(\hat{\boldsymbol{n}})\sqrt{\frac{4\pi}{2l'+1}}. \tag{6}$$

Inserting Eqs. (3) and (6) into Eq. (1) and making use of the orthonormality of the spherical harmonics basis finally yields (Driscoll and Healy, 1994; Healy et al., 1998)

$$S(\hat{\boldsymbol{u}}) = \sum_{l} \sum_{m} k_{l0} p_{lm} Y_{lm}(\hat{\boldsymbol{u}}) \sqrt{\frac{4\pi}{2l+1}}, \tag{7}$$

where $\hat{\boldsymbol{u}}$ can be parameterized by the polar and azimuthal angles, θ and ϕ, describing the orientation of $\mathbf{B}$, $\hat{\boldsymbol{u}} \equiv (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)$.

The spherical harmonic coefficients of the microstructure kernel (k_{l0}) and the ODF (p_{lm}) are estimated as the inner products between a given spherical harmonics basis function Y_{lm} and either $K(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}})$ or $P(\hat{\boldsymbol{n}})$. Due to the orthonormality of the spherical harmonics basis, the inner products are given by multiplication with the complex conjugates of the Y_{lm}, followed by integrations over the unit sphere. For the microstructural kernel, this procedure results in (Lampinen et al., 2020)

$$k_{l0} \equiv k_l = S_0 \sum_{j=1}^{J} f_j \sqrt{4\pi(2l+1)} \mathrm{I}_{lj} \exp\left(-bD_{\mathrm{I};j}\left(1 - b_\Delta D_{\Delta;j}\right)\right) \exp\left(-\frac{\tau_\mathrm{E}}{T_{2;j}}\right), \tag{8}$$

where b is the conventional b-value and b_Δ denotes the normalized anisotropy of the diffusion encoding tensor $\mathbf{B}$ (Eriksson et al., 2015). The isotropic diffusivity and the normalized diffusion anisotropy (D_I and D_Δ) are related to the axial and radial diffusivities ($D_\parallel$ and $D_\perp$) of the diffusion tensor according to $D_\mathrm{I} = \left(D_\parallel + 2D_\perp\right)/3$ and $D_\Delta = \left(D_\parallel - D_\perp\right)/3D_\mathrm{I}$ (Conturo et al., 1996); in its principal axis, a given $\mathbf{D}$ can thus be represented by a diagonal matrix parametrized as diag $\mathrm{diag}(D_\mathrm{I}(1 - D_\Delta), D_\mathrm{I}(1 - D_\Delta), D_\mathrm{I}(1 + 2D_\Delta))$. The I_{lj} factors are a function of the regular Legendre polynomials, L_l, and defined as

$$I_{lj} = \int_0^1 \exp\left(-\alpha_j x^2\right) \cdot L_l(x) \mathrm{d}x, \tag{9}$$

with $\alpha_j = 3bD_{\mathrm{I};j} b_\Delta D_{\Delta;j}$.

Different diffusion MRI models feature different numbers of components and impose different constraints on the component properties. Here we consider a two-compartment model ($J = 2$) comprising a “stick ” component (S) with $D_{\Delta;\mathrm{S}} = 1$ and a “zeppelin ” (Z) component with $D_{\Delta;\mathrm{Z}} \neq 1$. Truncating the spherical harmonic summation at the second order ($l^{\max} = 2$) then yields the signal according to

$$\begin{aligned} S(\boldsymbol{e}, \boldsymbol{m}) = S_0 \Big[& f_\mathrm{S} \exp\left(-bD_{\mathrm{I;S}}(1 - b_\Delta)\right) \\ & \times \left(\mathrm{I}_{0;\mathrm{S}} + 4\pi \mathrm{I}_{2;\mathrm{S}} \sum_m p_{2m} Y_{2m}(\theta, \phi)\right) \exp\left(-\frac{\tau_\mathrm{E}}{T_{2;\mathrm{S}}}\right) \\ & + (1 - f_\mathrm{S}) \exp\left(-bD_{\mathrm{I;Z}}\left(1 - b_\Delta D_{\Delta;\mathrm{Z}}\right)\right) \\ & \times \left(\mathrm{I}_{0;\mathrm{Z}} + 4\pi \mathrm{I}_{2;\mathrm{Z}} \sum_m p_{2m} Y_{2m}(\theta, \phi)\right) \exp\left(-\frac{\tau_\mathrm{E}}{T_{2;\mathrm{Z}}}\right) \Big], \end{aligned} \tag{10}$$

where $m \in \{-2, -1, 0, 1, 2\}$. The derivation of Eq. (10) uses the $p_{00} = Y_{00} = 1/\sqrt{4\pi}$ ODF normalization (Lampinen et al., 2020; Novikov et al., 2018). The vectors $\boldsymbol{e}$ and $\boldsymbol{m}$ capture the experiment-related parameters, $\boldsymbol{e} = (\tau_E, b, b_\Delta, \theta, \phi)$, and scalar model parameters, $\boldsymbol{m} = \left(f_S, D_{I;S}, D_{I;z}, D_{\Delta;z}, T_{2;s}, T_{2;z}, p_{20}, Re(p_{21}), \mathrm{Im}(p_{21}), Re(p_{22}), \mathrm{Im}(p_{22})\right)$, where $\mathrm{Re}(p_{lm}) = \left(p_{lm} + (-1)^m p_{l-m}\right)/2$ and $\mathrm{Im}(p_{lm}) = \left(p_{lm} - (-1)^m p_{l-m}\right)/2i$ denote the real and imaginary parts of the p_{lm} coefficients, respectively. We refer to the model expressed by Eq. (10) as the Standard Model with Relaxation (SMR). This name is chosen to mark its descendance from the "standard model" of WM microstructure (Novikov et al., 2019) and to emphasize the fact that it accounts for compartment-specific T_2 times.

The SMR model parameters can be determined by fitting Eq. (10) directly to the acquired signals (Lampinen et al., 2020). An alternative strategy is to fit to some representation of the signal, such as the spherical harmonics coefficients. Veraart et al. (2018) used a model fitting framework that effectively reduces the dimensionality of the parameter space by means of performing a rotationally invariant factorization of the voxel-wise ODFs (Novikov et al., 2018; Reisert et al., 2017). The initial step of such framework consists in projecting the measured signal onto a spherical harmonic basis

$$S(\hat{\boldsymbol{u}}) = \sum_l \sum_m S_{lm} Y_{lm}(\hat{\boldsymbol{u}}) \,. \tag{11}$$

The S_{lm} coefficients are subsequently converted to rotational invariants S_l, and fitted to the corresponding rotationally invariant terms of the $P(\hat{\boldsymbol{u}}) \otimes K(\hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{u}})$ convolution

$$S_l = p_l k_l, \tag{12}$$

where k_l is the 0th degree term of the microstructural kernel as defined by Eq. (8). The rotationally invariant coefficients, S_l and p_l, are computed from (Novikov et al., 2018)

$$x_l = \sqrt{\frac{4\pi \sum_m |x_{lm}|^2}{(2l+1)}}, \tag{13}$$

where x_{lm} are the spherical harmonics coefficients, and $x_l \equiv S_l$ or $x_l \equiv p_l$. At sufficiently low b-values, signal projections with $l > 2$ have small contributions to the measured signal Jespersen et al., 2007) and the sum in Eq. (11) is typically truncated at the second order term ($l = 2$). The fitting framework summarized by Eqs. (11)–((13) is commonly referred to as the "RotInv " approach due to its use of rotational invariants. The $l = 2$ RotInv approach condenses the five p_{2m}, $m \in \{-2, -1, 0, 1, 2\}$ parameters of the SMR model onto a single p_2 invariant capturing the orientation coherence of the sub-voxel diffusion domains, thus reducing the dimensionality of the fitting problem by four parameters.

3. Methods

3.1. Neural network architecture and training

In this work, we constructed feedforward neural networks in MATLAB R2020b (The MathWorks, Inc.), and used them to fit vectors of scalar

parameters, $\boldsymbol{m} = (f_S, D_{I;S}, D_{I;Z}, D_{\Delta;Z}, T_{2;S}, T_{2;Z}, p_{20}, \mathrm{Re}(p_{21}), \mathrm{Im}(p_{21}), \mathrm{Re}(p_{22}), \mathrm{Im}(p_{22}))$ to sets of measurements $S(\tau_E, \mathbf{B})$. We explored various network designs with different numbers of hidden nodes and/or layers before deciding on two final network architectures: an artificial neural network featuring 3 fully connected hidden layers with a decreasing number of nodes (180, 80, and 55) and a deeper/wider neural network featuring 4 fully connected hidden layers with 250 nodes each. All hidden layers were activated by hyperbolic tangent (tanh) functions and the deeper/wider network also featured batch normalization layers between the fully connected inner layers and their respective tanh activations. To distinguish the networks, we refer to them as the shallower neural network (SNN) and deeper neural network (DNN), respectively. Both SNN and DNN comprise an output layer with 11 nodes corresponding to the parameters in $\boldsymbol{m}$. The input comprised a given number (E) of signal samples acquired with a pre-defined relaxation-diffusion encoding protocol. We considered three different acquisition protocols; with $E = 164$, $E = 242$, and $E = 270$ samples ($\tau_E, \mathbf{B}$). Independent networks were trained for each protocol, meaning that 3 SNNs and 3 DNNs were evaluated. To remove the influence of S_0 from the fitting problem, we normalized the input vector to the median signal acquired at the lowest b-value and shortest echo-time.

Supervised network training was performed using a mean squared error loss

$$\mathrm{MSE} = \|\boldsymbol{m}_{\mathrm{targ}} - \boldsymbol{m}_{\mathrm{net}}\|_2^2, \tag{14}$$

where $\boldsymbol{m}_{\mathrm{targ}}$ is the ground-truth target vector, $\boldsymbol{m}_{\mathrm{net}}$ is the corresponding network output vector, and $\| \cdot \|_2$ denotes the Euclidean norm. The $\boldsymbol{m}_{\mathrm{targ}}$ parameters were rescaled between 0 and 1 using a min-max normalization strategy before being supplied to the networks. The networks were trained with sets of voxels with randomly generated model parameters and noisy signal samples $S(\tau_E, \mathbf{B})$, as detailed in Section 3.2. The SNNs were trained with a batch size of $0.5{\cdot}10^6$ and a scaled conjugate gradient optimiser. The DNNs were trained in a mini-batch fashion using a total of $5{\cdot}10^6$ training sets, a mini-batch size of $50{\cdot}10^3$, and an Adam optimiser. Throughout, training data was divided such that 75% of the original data was used to update the weights and biases and 25% was used for cross-validation. Overfitting was addressed by an early stopping method that terminated training following an increase of the MSE of the validation data for 5 (SNN) or 20 (DNN) consecutive epochs.

Network GPU training took approximately 83 min for the SNNs and 74 min for the DNNs on two parallel NVIDIA GeForce RTX 2080 SUPER, each with 8 GB of memory. Both graphic cards were installed on a high-end consumer-grade desktop computer with 32 GB memory and an 8-core Intel i9–9900k 3.6 GHz CPU with 2 threads per core.

3.2. Generating training data

We studied the impact of training data generation strategies on the network performance, including training based on uniformly sampled and real brain data. Training parameter vectors were created by two strategies:

- $\boldsymbol{m}_{\mathrm{unif}}$ was synthetically constructed by random sampling of uncorrelated uniform distributions within the bounds described in Table 1;

- $\boldsymbol{m}_{\text{brain}}$ was constructed from *in vivo* brain data by randomly sampling parameter vectors estimated from a NLLS fit of Eq. (10). This dataset contains parameter correlations found in a typical brain dataset from a healthy adult.

The $\boldsymbol{m}_{\text{brain}}$ vectors comprise the solutions of a nonlinear least-squares (NLLS) fit of Eq. (10) to *in vivo* signal data, referred to as $\boldsymbol{m}_{\text{fit}}$, together with an additional parameter set $\boldsymbol{m}_{\text{mut}}$, consisting of random mutations of the fitted solutions, given by

$$\boldsymbol{m}_{\text{mut}} = \boldsymbol{X} \circ \boldsymbol{m}_{\text{fit}}, \tag{15}$$

where '$\circ$' denotes the element-wise (Hadamard) product, and $\boldsymbol{X}$ is an 11-dimensional vector of normally distributed numbers. Each element of $\boldsymbol{X}$ is an independent and identically distributed random variable sampled from a normal distribution with mean 1 and standard deviation 0.3. The standard deviation of $\boldsymbol{X}$ was chosen following brief *in silico* experiments which revealed that virtually indistinguishable training/test results are obtained for standard deviations within the [0.2, 0.5] interval, provided all other training/network parameters are kept constant. The number of $\boldsymbol{m}_{\text{fit}}$ vectors was kept constant $\left(n_{\text{fit}} \approx 1.5 \cdot 10^5\right)$, and the total number of mutated vectors was defined as $n_{\text{mut}} = n_{\text{brain}} - n_{\text{fit}}$. The introduction of mutated parameters is a data augmentation technique, designed to simultaneously compensate for the relative low number of $\boldsymbol{m}_{\text{fit}}$ vectors and expand the $\left(f_{\text{S}}, D_{\text{I;S}}, D_{\text{I;Z}}, D_{\Delta;\text{Z}}, T_{2;\text{S}}, T_{2;\text{Z}}, p_{20}, \text{Re}(p_{21}), \text{Im}(p_{21}), \text{Re}(p_{22}), \text{Im}(p_{22})\right)$ domain of the $\boldsymbol{m}_{\text{brain}}$ parameter targets.

The training vectors, $\boldsymbol{m}_{\text{train}}$, were combinations of $\boldsymbol{m}_{\text{brain}}$ and $\boldsymbol{m}_{\text{unif}}$ parameter vectors. Using a given total number of vectors (n_{tot}) and varying number of $\boldsymbol{m}_{\text{brain}}$ parameters (n_{brain}), we modulated the fraction of *in vivo* brain data, $f_{\text{brain}} = n_{\text{brain}} / n_{\text{tot}}$, between 0 and 1 in steps of 0.05. The SNN training sets contained a total of $n_{\text{tot}} = 5 \cdot 10^5$ parameter vectors, while the DNN training sets contained $n_{\text{tot}} = 5 \cdot 10^6$. Fig. S1 in the Supporting Information shows the distribution of $\boldsymbol{m}_{\text{fit}}$, $\boldsymbol{m}_{\text{mut}}$, and $\boldsymbol{m}_{\text{unif}}$ parameters that compose a typical $n_{\text{tot}} = 5 \cdot 10^5$ SNN training dataset.

Signal data were generated from $\boldsymbol{m}_{\text{train}}$ using Eq. (10) and one of three different $(\tau_{\text{E}}, \mathbf{B})$ acquisition protocols:

- The *optimized* protocol comprises tensor-valued encoding with full relaxation-diffusion-correlation optimized for minimal SMR parameter variance (Lampinen et al., 2020)
- The *unoptimized* protocol comprises tensor-valued encoding with relaxation-diffusion-correlations restricted to low b-values (Lampinen et al., 2019). This protocol was an early attempt to design a diffusion-relaxation protocol with b-tensor encoding. It preceded the *optimized* protocol and was configured to fit into an available timeslot by following heuristics without a formal performance optimization, and was later found to yield degenerate results in white matter (Lampinen et al., 2020).

- The *LTE-only* protocol comprises diffusion-relaxation optimized for minimal SMR parameter variance but includes only linear b-tensor encoding ($b_\Delta = 1$) (Lampinen et al., 2020). Just as the *unoptimized* protocol it has been found to yield degenerate solutions in white matter.

Additional details on the various protocols can be found in their respective references and in Table S1 of the Supporting Information. We emphasize that all training data used in this study was generated using the SMR forward model, Eq. (10), rather than using raw *in vivo* brain data.

Noise was sampled from the Rice distribution and added to the ground-truth synthetic signals. Because relaxation-diffusion MRI data comprises voxels with different signal-to-noise ratio (SNR), the amplitude of the SNR at $S_0 = S(\tau_E = 0, \boldsymbol{B} = 0)$ was uniformly varied in the interval $\mathrm{SNR} \in [80, 160]$. Considering the relaxation-diffusion properties of typical healthy WM $\left(T_2 \approx 70\ \mathrm{ms}, D_I \approx 0.9\ \mu\mathrm{m}^2/\mathrm{ms}\right)$, this choice results in SNR ∈ [30, 60] at the point of maximal signal of the *optimized* protocol $\left(\tau_E = 63\ \mathrm{ms}, b = 0.1\ \mathrm{ms}/\mu\mathrm{m}^2\right)$, SNR amplitudes that are consistent with tensor-valued dMRI measurements of the *in vivo* brain (Szczepankiewicz et al., 2019a). Finally, networks were trained using $\boldsymbol{m}_{\mathrm{train}}$ vectors as targets and their corresponding *in silico* noisy signals as inputs.

3.3. Network evaluation

To find the optimal fraction of $\boldsymbol{m}_{\mathrm{unif}}$ and $\boldsymbol{m}_{\mathrm{brain}}$ parameters (adjusted by the f_{brain} parameter), we trained SNNs with varying values of f_{brain}, deployed them on *in silico* data generated from an unseen subject, and compared the various networks in terms of accuracy of the resulting parameter estimates. Network accuracy was assessed via normalized root-mean-squared errors (NRMSE) and linear correlations with ground-truth values in terms of the Pearson correlation coefficient (ρ). The NRMSE captures the absolute agreement between the target ground-truth parameters and their corresponding network estimates, whereas ρ captures the linear target-to-estimate correlation strength. The f_{brain} optimization process is discussed in more detail in section S3 of the Supporting Information. Briefly, the f_{brain} hyper-parameter controls a trade-off between accuracy to WM-relevant parameters and network generalizability, and we found $f_{\mathrm{brain}} = 0.5$ to provide an optimal balance between accuracy and generalizability. From this point onward, we concentrate on networks trained with $f_{\mathrm{brain}} = 0.5$ datasets and evaluate them in further detail using correlation plots.

The accuracy performance of an $f_{\mathrm{brain}} = 0.5$ SNN, an $f_{\mathrm{brain}} = 0.5$ DNN, and a standard NLLS solver were compared on the basis of NRMSEs and Pearson correlation coefficients. The comparison was performed using two distinct *in silico* datasets: one based on $\boldsymbol{m}_{\mathrm{fit}}$ vectors from WM and deep GM data ($\boldsymbol{m}_{\mathrm{fit;WM\text{-}like}}$), and another based on $\boldsymbol{m}_{\mathrm{unif}}$ vectors. Each dataset comprised a total of $10{\cdot}10^3$ parameter vectors and their respective *in silico* signals. The ground-truth synthetic signals were corrupted with Rician distributed noise and the SNR at the S_0 point was sampled uniformly from the [80, 160] range.

The effects of different acquisition protocols on network performance were evaluated in terms of NRMSE and sensitivity to parameter changes. The latter was gauged by modulating

the non-orientational parameters of an SMR solution $(f_S, D_{I;S}, D_{I;Z}, D_{\Delta;Z}, T_{2;S}, T_{2;Z})$ one at a time by 10% and measuring the response in all parameters. The original parameter set was based on *in vivo* data from the *corona radiata* where $f_S = 0.45$, $D_{I;S} = 0.58\ \mu m^2/ms$, $D_{I;Z} = 1.36\ \mu m^2/ms$, $D_{\Delta;Z} = 0.44$, $T_{2;S} = 69$ ms, $T_{2;Z} = 60$ ms (Lampinen et al., 2020). Subsequently, *in silico* datasets were generated for each of the 6 modulated datasets, Rice noise was added with SNR = 160 at S_0, and parameter estimation was performed with protocol-specific networks.

To investigate if the reduced parameter space of RotInv fitting impacts the performance of ANN-based fitting, we trained an SNN using rotationally invariant *in silico* datasets and the same optimal f_{brain} value found for the SMR networks. RotInv training vectors, $\boldsymbol{m}_{train;RI}$, were generated from the $\boldsymbol{m}_{train}$ vectors (Section 3.2), using Eq. (13) to convert the full SMR parameters to RotInv parameters $(f_S, D_{I;S}, D_{I;Z}, D_{\Delta;Z}, T_{2;S}, T_{2;Z}, p_2)$. The RotInv *in silico* signal data was generated in four steps: **(1)** signals were calculated using $\boldsymbol{m}$ train and Eq. (10); **(2)** noise was added to the *in silico* signal data with a SNR $\in$ [80, 160] at S_0; **(3)** S_{lm} components were estimated by projecting the noisy $S(\tau_E, \mathbf{B})$ signals to a spherical harmonics basis; and **(4)** $S_{l,\ l=\{0,2\}}$ signals were calculated from S_{lm} using Eq. (13). As with the full SMR model, training was performed using $\boldsymbol{m}_{train;RI}$ as targets and their corresponding synthetic noisy signals as ANN inputs.

Trained SMR (RotInv) networks were tested on previously unseen $\boldsymbol{m}_{unif}$ ($\boldsymbol{m}_{unif;RI}$) and $\boldsymbol{m}_{brain}$ ($\boldsymbol{m}_{brain;RI}$) synthetic datasets at an SNR $\in$ [80, 160] at S_0. Performance was compared in terms of their respective target-estimate correlations. All networks were trained/tested in a leave-one-out fashion where the training and testing $\boldsymbol{m}_{brain}$ ($\boldsymbol{m}_{brain;RI}$) datasets were generated using *in vivo* data from different subjects (Section 3.5).

3.4. In vivo *data acquisition*

We analysed data from three adult volunteers previously reported in (Lampinen et al., 2020). The study was approved by the regional ethical review board in Lund and written informed consent was obtained from all volunteers prior to scanning. Measurements were performed on a MAGNETOM Prisma 3T system (Siemens Health-care, Erlangen, Germany) using a prototype spin-echo EPI sequence that facilitates user-defined gradient waveforms for diffusion encoding (Szczepankiewicz et al., 2019a). Data were collected using echo times between 63 and 130 ms, repetition time of 3.4 s, voxel size of 2.5 mm^3, 40 slices, matrix-size of 88 × 88, in-plane and through plane acceleration factor of 2 × 2 (GRAPPA), partial-Fourier of 3/4, band-width = 1775 Hz/pixel, and “strong ” fat saturation. Diffusion encoding was performed with gradient waveforms optimized to maximize the encoding strength per unit time and to suppress concomitant field effects (Sjölund et al., 2015; Szczepankiewicz et al., 2019b). A total of 270 combinations of τ_E and **B** were used, according to the *optimized* protocol in Table S1 of the Supporting Information. Total acquisition time was 15 min.

3.5. In vivo *data processing and parameter estimation*

Prior to analysis, all *in vivo* data were corrected for eddy-currents and subject motion using ElastiX (Klein et al., 2009) with extrapolated target volumes (Nilsson et al., 2015).

Susceptibility-induced geometric distortions were corrected using the TOPUP tool in FMRIB software library (FSL) (Smith et al., 2004). Gibbs ringing artefact correction was performed according to the method described in (Kellner et al., 2016). To suppress the influence of noise, we filtered data with a 3D Gaussian kernel with a standard deviation of 0.45 times the voxel dimensions (Lampinen et al., 2020).

The SMR model parameters were estimated from a voxel-by-voxel NLLS fit of Eq. (10) to the post-processed data. The fitting process was performed with the multidimensional dMRI toolbox (https://github.com/markus-nilsson/md-dmri) (Nilsson et al., 2018), with MATLAB's built-in *lsqcurvefit* function. To remove outliers, model fitting was performed twice in each voxel and the result with lowest residual was retained (Lampinen et al., 2020). The initial guesses were sampled uniformly from the parameter bounds in Table 1. The resulting estimates were stored and used to compute *in silico* signal data following the procedure detailed in Section 3.2. NLLS fitting of a single *in vivo* brain dataset took approximately 8 h (approximately 5.5 s per voxel) on the CPU described in Section 3.1. The computations were carried out using parallel computing and multi-threading.

Finally, previously trained networks were used to estimate the parameters from Eq. (10) from *in vivo* data, which took approximately 2 and 20 s for the whole brain using the SNN and DNN, respectively. Training was performed on *in silico* $\boldsymbol{m}_{\text{train}}$ data with an optimal f_{brain} fraction. The training process followed a leave-one-out scheme, where the networks were trained on synthetic data generated from two subjects before being deployed/tested on a third, previously unseen, subject. Neural network fitting provided voxel-wise parameter maps that were compared to the ones obtained from a conventional NLLS fitting approach.

4. Results

4.1. Neural network parameter estimates

SNN-based parameter estimation was approximately 10^4 times faster than NLLS fitting on the same computer, and yielded parameters in good agreement with the ground-truth targets and preserved contrast between regions characterized by distinct (T_2, **D**) properties (Fig. 1). For example, the estimated f_S and p_2 are high in WM regions generally and highest in orientationally coherent WM regions such as the *corpus callosum*, similar to the *in-silico* ground-truth. However, a reduced contrast was observed in the $T_{2;Z}$ maps, where the distinction between WM (darker) and cortical GM (brighter) regions is more prominent in the ground-truth map. The $T_{2;Z}$ estimates are also characterized by considerable differences between ground-truth and estimated parameters in the long T_2 regions such as the lateral ventricles. The largest overall discrepancy between estimated and ground-truth parameters was found for $D_{\Delta;Z}$, likely because the signal is insensitive to it when $|D_{\Delta;Z}| < 0.5$ (Eriksson et al., 2015). Using an ANN trained on synthetic data to directly fit *in vivo* experimental data resulted in noisier maps. Nevertheless, it preserved an anatomically plausible contrast. Given the strong correlations between *in silico* ground-truth maps and network estimates, the noisier appearance of the *in vivo* parameter maps is likely because the SMR model cannot accurately represent the underlying *in vivo* data. *In vivo* SNN parameter estimates from WM

regions of interests are displayed in Table S2 of the Supporting Information, where they are additionally compared to NLLS estimates.

Fig. 2 shows that SNN-based estimates correlated well with the ground-truth parameter targets, with most parameters yielding linear correlation coefficients close to or above 0.9. The referenced figure focuses on the performance of a network trained with *in silico* $S(\tau_E, \mathbf{B})$ data generated with the *optimized* protocol and an even mix of random and WM-like samples ($f_{brain} = 0.5$), and distinguishes between performance on parameters obtained by uniform random sampling (light blue points) and parameters derived from *in vivo* non-cortical brain data (dark blue points). Red points correspond to parameter vectors derived from low component-specific signal fractions, as described in the figure caption. Poor performance is observed for low $D_{\Delta;Z}$ values, where the network yields $D_{\Delta;Z} \approx 0.3$ regardless of the underlying ground-truth. This can be attributed to an intrinsic difficulty in distinguishing between the diffusion-weighted signals of components with $|D_{\Delta;Z}| < 0.5$ components (Eriksson et al., 2015). Moreover, a poor target-to-estimate correspondence was seen for $T_{2;Z}$ -times where these were longer than the maximal echo time.

The parameter maps estimated from the deeper network are in good agreement with their respective ground-truth targets (Figs. S3 and S4 of the Supporting Information correspond to Figs. 1 and 2). In Fig. S4, we observed that DNN-based fitting resulted in slightly stronger correlations between network estimates and ground-truth parameter targets. Although *in vivo* maps from DNN and SNN are similar, differences can be found in $D_{\Delta;Z}$ and $T_{2;Z}$; the DNN produces a noisier $D_{\Delta;Z}$ and the $T_{2;Z}$ map has a higher contrast between WM and cortical GM. Both of these features are likely artefactual, and suggest that the DNN is more susceptible to differences between the SMR signal predictions and the measured *in vivo* data.

The errors and prediction-target correlations of the ANN-based estimates are compiled in Table 2, where they are also compared to a conventional NLLS solver. The NLLS, SNN, and DNN approaches all have a comparable accuracy for *in silico* datasets designed to capture WM (T_2, **D**) properties. By contrast, the function-fitting networks are observed to be more accurate than the NLLS approach for synthetic $\boldsymbol{m}_{unif}$ parameter vectors.

4.2. Effect of acquisition protocol on network accuracy and sensitivity

In this section, we focus on the relationship between acquisition protocol design and network performance. Fig. 3 shows that ANN-based fitting could partly but not completely eliminate the known fit degeneracy in the *unoptimized* and *LTE-only* protocols: ANNs based on the *optimized* protocol provide lower estimation errors (NRMSE) than the ANNs based on the other two protocols. Comparing these two protocols, we note that the *unoptimized* protocol yields relatively more accurate estimates of $D_{I;S}$, and $T_{2;Z}$, while the *LTE-only* protocol yields more accurate estimates of f_S, $D_{I;Z}$, $D_{\Delta;Z}$, $T_{2;S}$, and p_2. Fig. 3 also shows that the performance of both SNN and DNN is less affected by sub-optimal acquisition protocols than the traditional NLLS approach (Fig. 3). For the NLLS approach, the use of the *unoptimized* or *LTE-only* protocols leads to a considerable increase of the estimation errors, while only a slight increase of NRMSE is observed for the DNN or SNN approaches.

This suggests that ANNs may partly alleviate parameter estimation difficulties caused by a protocol that is inadequate in relation to the microstructure model.

Fig. 4 shows the sensitivity of the various protocols to parameter changes. Networks trained on data generated with the *optimized* protocol are sensitive to all parameters, but slightly underestimate the magnitude of the change, particularly in $D_{\Delta;Z}$. The parameter-specific modulations did not have a major effect on the estimation of the remaining unmodulated parameters. An exception was found when the underlying $T_{2;Z}$ is increased by 10%, which results in a 3% overestimation of the unchanged $D_{I;S}$. Compared to the *optimized* protocol, the *unoptimized* and *LTE-only* protocols exhibit a lower sensitivity to the small parameter modulations and appear to be unresponsive to changes in $D_{\Delta;Z}$ (both protocols) and $D_{I;S}$ (*LTE-only*). In addition to lower sensitivity, the *unoptimized* protocol also resulted in less accurate estimations of the unmodulated parameters, with a 10% modulation of f_S leading to an erroneous 7% increase in $T_{2;S}$.

4.3. Neural network fitting of rotationally invariant microstructural features

Fig. 5 A shows that training a SNN with rotational invariants results in slightly stronger correlation between target and estimated parameters (compare with the scatter plots of Fig. 2). We note a considerable improvement in accuracy at low $D_{\Delta;Z}$ values, where the constant $D_{\Delta;Z} \approx 0.3$ behaviour observed for the full SMR model (see Fig. 2) is no longer present. Applying the RotInv network to an unseen *in vivo* $S_{l=\{0,2\}}$ dataset results in parameter maps with anatomically plausible contrast (see Fig. 5 B). Consistent with the better $D_{\Delta;Z}$ accuracy performance of the RotInv approach, we note that the RotInv $D_{\Delta;Z}$ *in vivo* map has a smoother appearance and better demarcates cortical/non-cortical parenchyma than its non-rotationally invariant SMR counterpart (compare the fourth column of Figs. 1 with 5 B).

Interestingly, *in vivo* maps smoother than the ones displayed in Fig. 5 B can be attained from an ANN that was trained on unreasonably noisy *in silico* data. Fig. 6 displays the *in vivo* parameter maps obtained from a RotInv network trained with SNR $\in$ [20, 40] at S_0, which is 4 times lower than that used in Fig. 5. The resulting maps have a smooth appearance and exhibit anatomically plausible contrast. For example, regions with high f_S correspond to WM regions, the lateral ventricles are characterized by low f_S and high $D_{I;Z}$ values, and darker/brighter $D_{\Delta;Z}$ regions demarcate cortical/non-cortical parenchyma. While it is tempting to favour the seductively 'robust' maps of Fig. 6 over the noisier maps of Fig. 5 B, we note that the low-SNR RotInv network results in weak correlations between target and estimated parameters (compare the scatter plots of Fig. 6 with those of Fig. 5 B). For example, SNN-based estimates of $D_{\Delta;Z}$ may yield a smooth map that appears robust, but a closer inspection reveals that the $D_{\Delta;Z}$ estimates in WM and deep GM regions are equal to the mean of the target $D_{\Delta;Z}$ distribution and constitute an exceedingly inaccurate estimate of the underlying ground-truth. The tendency for networks to return the mean of the training parameter distribution has been reported in studies of the RotInv model (Reisert et al., 2017) and the behaviour was explained in detail by Coelho et al. (2021).

5. Discussion and conclusions

Replacing traditional NLLS solvers with function-fitting neural networks enables vastly faster parameter estimation when using high-dimensional microstructural models. On a consumer-grade desktop computer, the fitting time was reduced from hours (NNLS) to seconds (ANN). Naturally, the NNLS fitting times, based on the relatively slow trust-region-reflective algorithm, can be improved by linearizing the fitting problem (Daducci et al., 2015) or by using GPU-based solvers (Harms et al., 2017). However, while such procedures have enabled whole brain fitting of non-linear models within minutes (Daducci et al., 2015; Harms et al., 2017), we still expect the seconds-long forward pass of an ANN to provide a competitive choice in terms of computation time.

The ANN-based estimates were observed to be in good agreement with synthetic data that mimicked healthy WM as well as data that spanned the entire space of allowed model parameters. When deployed on unseen *in vivo* brain data, neural networks provide maps that are consistent with known brain anatomy and preserve contrast be-tween regions with different relaxation-diffusion properties. Our findings are encouraging and in line with recent advanced dMRI modelling studies that use machine learning techniques for parameter estimation (Barbieri et al., 2020; Bertleff et al., 2017; Golkov et al., 2016; Grussu et al., 2020; Gyori et al., 2019; Hill et al., 2021; Kaandorp et al., 2021; Nedjati-Gilani et al., 2017; Palombo et al., 2020; Reisert et al., 2017). A combination of error metrics, correlation analysis, and sensitivity matrices was found to provide a useful set of tools for quantitatively assessing parameter-specific accuracy/sensitivity and for identifying the limitations of learning-based approaches. These tools facilitate a survey of the performance across all dimensions of the SMR model, for example, revealing that $D_{\Delta;Z}$ was consistently less accurate than other parameters, as expected from previous studies that have emphasized that it is difficult to estimate (Eriksson et al., 2015; Lampinen et al., 2020, 2019). By contrast, visual inspection of ANN-based parameter maps was found to provide limited insight on the general performance of the networks. Indeed, smooth and anatomically plausible maps can be achieved even with poor network performance and data with low SNR. This is a common and deceptive pitfall that has strong implications for the evaluation of performance in machine learning approaches (Reisert et al., 2017).

We found no evidence that voxel-wise ANN-based parameter estimation can fully alleviate the degenerate fitting landscape typically present when working with biophysical models in dMRI (Jelescu et al., 2016) or replace an exhaustive sampling of all relevant experimental dimensions (Coelho et al., 2019; Lampinen et al., 2020). Fig. 3 shows worse performance in terms of parameter estimation errors (NRMSE) for the two protocols with known degeneracy problems (Lampinen et al., 2020). Similarly, Fig. 4 shows that only the optimized protocol can faithfully recover parameter-specific changes while the other two cannot. These are both signs of unresolved degeneracies. Indeed, we cannot expect good performance for *LTE-only* and *unoptimized* protocols because these protocols can yield virtually identical signal vectors for different model parameters; the inverse problem has many solutions (Lampinen et al., 2020, 2019). Nevertheless, ANN-based fitting showed an advantage compared with the traditional NLLS approach, as it yielded lower estimation errors in the degenerate cases (*unoptimized* and *LTE-only* protocols; Fig. 3). Our

interpretation is that the NLLS approach returns one out of the many solutions, whereas the ANN-based estimate tends to an average across the many solutions.

The 11-dimensional parameter space of the SMR model is difficult to sample densely and thus presents a challenge when designing training datasets that are representative of the vast fitting landscape. In this work, we addressed this challenge by constructing training data based on *in vivo* healthy brain data ($\boldsymbol{m}_{\text{brain}}$) and more naïve parameter vectors randomly sampled from the entire model parameter space ($\boldsymbol{m}_{\text{unif}}$). Networks trained exclusively with $\boldsymbol{m}_{\text{brain}}$ vectors displayed the best accuracy in terms of expected WM properties, but their domain of validity is restricted to the relatively small space spanned by $\boldsymbol{m}_{\text{brain}}$ solutions. This raises questions about their generalizability, *i.e.*, their performance in cases where atypical microscopic tissue structures are present (Alexander et al., 2019). To find a good trade-off between accuracy and generalizability we optimized the fraction of *in vivo*-based training data (f_{brain}). However, we expect that more work is needed to define a truly optimal strategy for network training.

In this study, we focused on fully connected networks that follow the design of multilayer perceptrons (MLPs), a traditional ANN class that is well-suited for regression problems (Cybenko, 1989; Hornik et al., 1989). Alternatives or complements to the fully connected ANN architecture should also be explored in future works. Promising avenues include the use of dropout (Gal and Ghahramani, 2016; Tanno et al., 2021) or deep ensemble strategies (Lakshminarayanan et al., 2016; Qin et al., 2021) as a means to derive uncertainty metrics, the use of rolled-out network structures inspired by non-learning-based iterative fitting frame-works (Ye, 2017), the use of auto-encoders (Zucchelli et al., 2021), or the use of denoising networks (Fadnavis et al., 2020; Wang et al., 2019) to minimize the amount of noise present in the data that is supplied to the function-fitting ANN. While fundamentally different network architectures may considerably boost the performance of the ANN-based fitting approach, the modest differences found between the SNN and DNN designs suggests that simply increasing the width and/or depth of the fully connected ANN architecture is not a promising avenue. Despite the potential for improvement, we note that the plots in Fig. 2 constitute an improvement over similar target-estimate correlation plots reported in (Reisert et al., 2017), where supervised learning based on polynomial regressors was used to fit a three-compartment diffusion model, and are equivalent to the correlation plots reported in more recent works on learning-based fitting of diffusion (Gyori et al., 2019; Palombo et al., 2020) and diffusion-relaxation MRI models (Grussu et al., 2020).

The fully connected ANNs we considered here are not invariant to sample rotations. The input to the ANNs is a vector of E signal samples measured at a pre-defined set of both directional (θ, ϕ) and rotationally invariant ($\tau_{\text{E}}, b, b_{\Delta}$) experimental points. The ordering in which the E measurements are provided is kept fixed and samples with similar microstructural kernels $K(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}})$ but different orientations will result in distinctive network input vectors. This places a burden on the training data generation, which has to span a sufficient set of possible tissue orientations. Using the RotInv formulation renders the network invariant to sample rotations, which considerably reduces the dimensionality of the parameter space that has to be represented in training data. The higher training efficiency likely explains its slightly higher accuracy performance relative to the full SMR networks.

Alternatives to the RotInv formulation presented in this study include a framework based on a different set of rotationally invariant features of the dMRI signal (Zucchelli et al., 2021) or fitting the full SMR model with equivariant network architectures (Cohen et al., 2018; Thomas et al., 2018).

A potential limitation of the present study is the focus on a single multi-compartment model of tissue microstructure whose range of application is mostly limited to WM and deep GM tissues. Applications for cortical GM should therefore consider models tailored to the appropriate microstructure (Palombo et al., 2020). Our decision to focus on a single model follows from previous dMRI literature which has presented the “Standard Model ” of tissue microstructure – from which our SMR model descends – as an overarching signal model that encompasses several other WM models as particular cases (Novikov et al., 2019). Fur-thermore, the “Standard Model” has been used to reveal general degeneracy problems in microstructure parameter estimation (Novikov et al., 2018). Given the generality of our model and the prevalence of degeneracies in advanced dMRI modelling, we expect the degradation of performance with less optimal protocols to also be found in alternative multi-compartment models or when using different learning-based fitting algorithms (*e.g.* : polynomial Reisert et al. 2017 or random forest Nedjati-Gilani et al. 2017, Palombo et al. 2020 regressors). However, future work is needed to fully characterize the general relationship be-tween machine learning approaches and degenerate fitting landscapes.

In conclusion, function fitting neural networks can be used to vastly accelerate parameter estimation with high-dimensional microstructural MRI models. The accuracy of ANN-based estimates was observed to degrade less with sub-optimal protocols than traditional NLLS fitting. However, the performance of function fitting networks was still observed to primarily depend on the amount of information sampled by the underlying measurements, and we found no evidence that ANN-based approaches can offset the need for a rich set of data. Therefore, machine learning methodology in MRI microstructure modelling should be matched with comprehensive data acquisition. This work presents a learning-based fitting framework, as well as tools for evaluating combinations of networks and measurement protocols in terms of error metrics, estimate-target correlation plots, and sensitivity matrices.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was financially supported by grants from the Swedish Prostate Cancer Federation, the Swedish Research Council (2016-03443, 2020-04549), eSSENCE, and Cancerfonden (2019/474). J. P. de Almeida Martins and P. T. While were supported by a grant from the Research Council of Norway (FRIPRO Researcher Project 302624) and M. Palombo by the UKRI Future Leaders Fellowship (MR/T020296/1).

References

Alexander DC, Dyrby TB, Nilsson M, Zhang H, 2019. Imaging brain microstructure with diffusion MRI: practicality and applications. NMR Biomed. 32, e3841. [PubMed: 29193413]

Alexander DC, Zikic D, Ghosh A, Tanno R, Wottschel V, Zhang J, Kaden E, Dyrby TB, Sotiropoulos SN, Zhang H, Criminisi A, 2017. Image quality transfer and applications in diffusion MRI. Neuroimage 152, 283–298. [PubMed: 28263925]

Barbieri S, Gurney-Champion OJ, Klaassen R, Thoeny HC, 2020. Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. Magn. Reson. Med 83, 312–321. [PubMed: 31389081]

Bertleff M, Domsch S, Weingärtner S, Zapp J, O'Brien K, Barth M, Schad LR, 2017. Diffusion parameter mapping with the combined intravoxel incoherent motion and kurtosis model using artificial neural networks at 3 T. NMR Biomed. 30, e3833.

Coelho S, Fieremans E, Novikov DS, 2021. How do we know we measure tissue parameters, not the prior? In: Proceedings of the Annual Meeting of the ISMRM Online conference.

Coelho S, Pozo JM, Jespersen SN, Frangi AF, 2019. Optimal experimental design for biophysical modelling in multidimensional diffusion MRI. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap PT, Khan A. (Eds.), Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Springer International Publishing, Cham, pp. 617–625.

Cohen TS, Geiger M, Köhler J, Welling M, 2018. Spherical cnns. arXiv preprint arXiv:1801.10130.

Conturo TE, McKinstry RC, Akbudak E, Robinson BH, 1996. Encoding of anisotropic diffusion with tetrahedral gradients: a general mathematical diffusion formalism and experimental results. Magn. Reson. Med 35, 399–412. [PubMed: 8699953]

Cybenko G, 1989. Approximation by superpositions of a sigmoidal function. Math. Control Signals Syst 2, 303–314.

Daducci A, Canales-Rodríguez EJ, Zhang H, Dyrby TB, Alexander DC, Thiran JP, 2015. Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. Neuroimage 105, 32–44. [PubMed: 25462697]

de Almeida Martins JP, Tax CMW, Szczepankiewicz F, Jones DK, Westin CF, Topgaard D, 2020. Transferring principles of solid-state and Laplace NMR to the field of *in vivo* brain MRI. Magn. Reson 1, 27–43.

de Almeida Martins JP, Topgaard D, 2018. Multidimensional correlation of nuclear relaxation rates and diffusion tensors for model-free investigations of heterogeneous anisotropic porous materials. Sci. Rep 8, 2488. [PubMed: 29410433]

Driscoll JR, Healy DM, 1994. Computing fourier transforms and convolutions on the 2-sphere. Adv. Appl. Math 15, 202–250.

Eriksson S, Lasic S, Nilsson M, Westin CF, Topgaard D, 2015. NMR diffusion-encoding with axial symmetry and variable anisotropy: distinguishing between prolate and oblate microscopic diffusion tensors with unknown orientation distribution. J. Chem. Phys 142, 104201. [PubMed: 25770532]

Fadnavis S, Batson J, Garyfallidis E, 2020. Patch2Self: denoising diffusion MRI with self-supervised learning. arXiv preprint arXiv:2011.01355.

Gal Y, Ghahramani Z, 2016. Dropout as a bayesian approximation: representing model uncertainty in deep learning. In: Proceedings of the International Conference on Machine Learning, pp. 1050–1059.

Golkov V, Dosovitskiy A, Sperl JI, Menzel MI, Czisch M, Sämann P, Brox T, Cremers D, 2016. q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans. Med. Imaging 35, 1344–1351. [PubMed: 27071165]

Grussu F, Battiston M, Palombo M, Schneider T, Wheeler-Kingshott CAG, Alexander DC, 2020. Deep learning model fitting for diffusion-relaxometry: a comparative study. bioRxiv preprint doi: 10.1101/2020.10.20.347625.

Gyori NG, Clark CA, Dragonu I, Alexander DC, Kaden E, 2019. *In-vivo* neural soma imaging using b-tensor encoding and deep learning. In: Proceedings of the Annual Meeting of the ISMRM. Montreal, Canada.

Harms RL, Fritz FJ, Tobisch A, Goebel R, Roebroeck A, 2017. Robust and fast non-linear optimization of diffusion MRI microstructure models. Neuroimage 155, 82–96. [PubMed: 28457975]

Healy DM, Hendriks H, Kim PT, 1998. Spherical Deconvolution. J. Multivar. Anal 67, 1–22.

Hill I, Palombo M, Santin M, Branzoli F, Philippe A−C, Wassermann D, Aigrot M−S, Stankoff B, Baron-Van Evercooren A, Felfli M, Langui D, Zhang H, Lehericy S, Petiet A, Alexander DC, Ciccarelli O, Drobnjak I, 2021. Machine learning based white matter models with permeability: an experimental study in cuprizone treated in-vivo mouse model of axonal demyelination. Neuroimage 224, 117425. [PubMed: 33035669]

Hornik K, Stinchcombe M, White H, 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366.

Jelescu IO, Veraart J, Fieremans E, Novikov DS, 2016. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. NMR Biomed. 29, 33–47. [PubMed: 26615981]

Jespersen SN, Kroenke CD, Ostergaard L, Ackerman JJ, Yablonskiy DA, 2007. Modeling dendrite density from magnetic resonance diffusion measurements. Neuroimage 34, 1473–1486. [PubMed: 17188901]

Kaandorp MPT, Barbieri S, Klaassen R, van Laarhoven HWM, Crezee H, While PT, Nederveen AJ, Gurney-Champion OJ, 2021. Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients. Magn. Reson. Med 86, 2250–2265. [PubMed: 34105184]

Kellner E, Dhital B, Kiselev VG, Reisert M, 2016. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn. Reson. Med 76, 1574–1581. [PubMed: 26745823]

Klein S, Staring M, Murphy K, Viergever MA, Pluim JP, 2009. Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging 29, 196–205. [PubMed: 19923044]

Lakshminarayanan B, Pritzel A, Blundell C, 2016. Simple and scalable predictive uncertainty estimation using deep ensembles. arXiv preprint arXiv:1612.01474.

Lampinen B, Szczepankiewicz F, Mårtensson J, van Westen D, Hansson O, Westin CF, Nilsson M, 2020. Towards unconstrained compartment modeling in white matter using diffusion-relaxation MRI with tensor-valued diffusion encoding. Magn. Reson. Med 84, 1605–1623. [PubMed: 32141131]

Lampinen B, Szczepankiewicz F, Noven M, van Westen D, Hansson O, Englund E, Martensson J, Westin CF, Nilsson M, 2019. Searching for the neurite density with diffusion MRI: challenges for biophysical modeling. Hum. Brain Mapp 40, 2529–2545. [PubMed: 30802367]

Lasič S, Szczepankiewicz F, Eriksson S, Nilsson M, Topgaard D, 2014. Microanisotropy imaging: quantification of microscopic diffusion anisotropy and orientational order parameter by diffusion MRI with magic-angle spinning of the q-vector. Front. Phys 2, 11.

Nedjati-Gilani GL, Schneider T, Hall MG, Cawley N, Hill I, Ciccarelli O, Drobnjak I, Wheeler-Kingshott CAMG, Alexander DC, 2017. Machine learning based compartment models with permeability for white matter microstructure imaging. Neuroimage 150, 119–135. [PubMed: 28188915]

Nilsson M, Szczepankiewicz F, Lampinen B, Ahlgren A, de Almeida Martins JP, Lasic S, Westin CF, Topgaard D, 2018. An open-source framework for analysis of multidimensional diffusion MRI data implemented in MATLAB. In: Proceedings of the 26th Annual Meeting of the ISMRM. Paris, France.

Nilsson M, Szczepankiewicz F, van Westen D, Hansson O, 2015. Extrapolation-based references improve motion and eddy-current correction of high B-value DWI data: application in Parkinson's disease dementia. PLoS ONE 10, e0141825. [PubMed: 26528541]

Nilsson M, van Westen D, Ståhlberg F, Sundgren PC, Lätt J, 2013. The role of tissue microstructure and water exchange in biophysical modelling of diffusion in white matter. Magn. Reson. Mater. Phy 26, 345–370.

Novikov DS, Fieremans E, Jespersen SN, Kiselev VG, 2019. Quantifying brain microstructure with diffusion MRI: theory and parameter estimation. NMR Biomed. 32, e3998. [PubMed: 30321478]

Novikov DS, Veraart J, Jelescu IO, Fieremans E, 2018. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. Neuroimage 174, 518–538. [PubMed: 29544816]

Palombo M, Ianus A, Guerreri M, Nunes D, Alexander DC, Shemesh N, Zhang H, 2020. SANDI: a compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. Neuroimage 215, 116835. [PubMed: 32289460]

Qin Y, Liu Z, Liu C, Li Y, Zeng X, Ye C, 2021. Super-Resolved q-Space deep learning with uncertainty quantification. Med. Image Anal 67, 101885. [PubMed: 33227600]

Reisert M, Kellner E, Dhital B, Hennig J, Kiselev VG, 2017. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. Neuroimage 147, 964–975. [PubMed: 27746388]

Sjölund J, Szczepankiewicz F, Nilsson M, Topgaard D, Westin C−F, Knutsson H, 2015. Constrained optimization of gradient waveforms for generalized diffusion encoding. J. Magn. Reson 261, 157–168. [PubMed: 26583528]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM, 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219. [PubMed: 15501092]

Szczepankiewicz F, Sjölund J, Ståhlberg F, Latt J, Nilsson M, 2019a. Tensor-valued diffusion encoding for diffusional variance decomposition (DIVIDE): technical feasibility in clinical MRI systems. PLoS ONE 14, e0214238. [PubMed: 30921381]

Szczepankiewicz F, Westin CF, Nilsson M, 2019b. Maxwell-compensated design of asymmetric gradient waveforms for tensor-valued diffusion encoding. Magn. Reson. Med 82, 1424–1437. [PubMed: 31148245]

Tanno R, Worrall DE, Kaden E, Ghosh A, Grussu F, Bizzi A, Sotiropoulos SN, Criminisi A, Alexander DC, 2021. Uncertainty modelling in deep learning for safer neuroimage enhancement: demonstration in diffusion MRI. Neuroimage 225, 117366. [PubMed: 33039617]

Thomas N, Smidt T, Kearnes S, Yang L, Li L, Kohlhoff K, Riley P, 2018. Tensor field networks: rotation-and translation-equivariant neural networks for 3d point clouds. arXiv preprint arXiv:1802.08219.

Tian Q, Bilgic B, Fan Q, Liao C, Ngamsombat C, Hu Y, Witzel T, Setsompop K, Polimeni JR, Huang SY, 2020. DeepDTI: high-fidelity six-direction diffusion tensor imaging using deep learning. Neuroimage 219, 117017. [PubMed: 32504817]

Topgaard D, 2017. Multidimensional diffusion MRI. J. Magn. Reson 275, 98–113. [PubMed: 28040623]

Veraart J, Novikov DS, Fieremans E, 2018. TE dependent Diffusion Imaging (TEdDI) distinguishes between compartmental T2 relaxation times. Neuroimage 182, 360–369. [PubMed: 28935239]

Wang H, Zheng R, Dai F, Wang Q, Wang C, 2019. High-field mr diffusion-weighted image denoising using a joint denoising convolutional neural network. J. Magn. Reson. Imaging 50, 1937–1947. [PubMed: 31012226]

Westin C−F, Knutsson H, Pasternak O, Szczepankiewicz F, Özarslan E, van Westen D, Mattisson C, Bogren M, O'donnell LJ, Kubicki M, 2016. Q-space trajectory imaging for multidimensional diffusion MRI of the human brain. Neuroimage 135, 345–362. [PubMed: 26923372]

Ye C, et al., 2017. Estimation of tissue microstructure using a deep network inspired by a sparse reconstruction framework. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap P-T, et al. (Eds.), Information Processing in Medical Imaging. Springer International Publishing, Cham, pp. 466–477.

Zucchelli M, Deslauriers-Gauthier S, Deriche R, 2021. Brain tissue microstructure characterization using dMRI based autoencoder neural-networks. In: Proceedings of the International MICCAI Workshop on Computational Diffusion MRI.

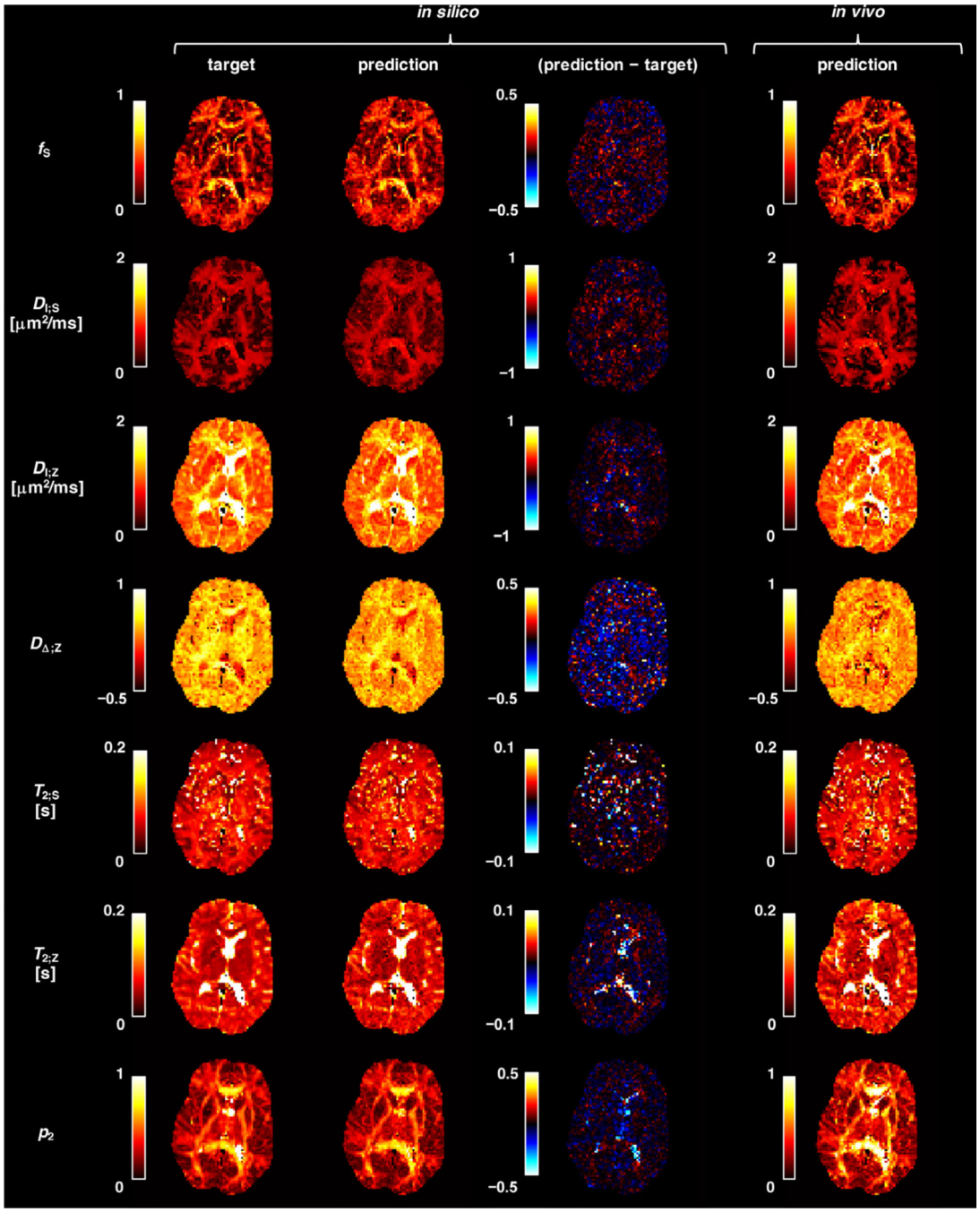

Fig. 1.

Deploying trained networks on previously unseen *in silico* and *in vivo* data provides anatomically plausible parameter maps in under 10 s (including data management times). The first and second columns compare the ground-truth targets and network predictions, respectively, of the *in silico* dataset. Difference maps are shown in the third column. Parameter maps obtained from applying a trained network on *in vivo* brain data are displayed in the fourth column.

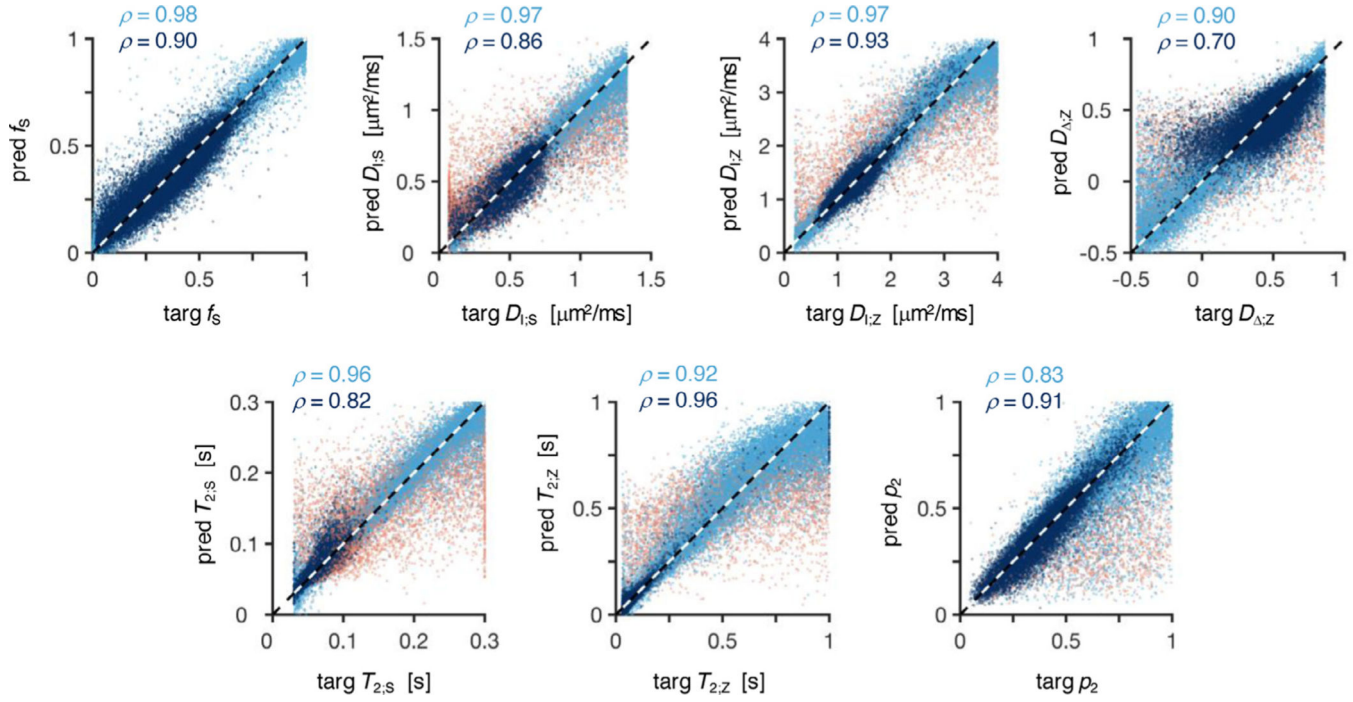

Fig. 2.

Scatter plots of ground-truth parameters *vs.* neural network predictions. Light blue points show results when the network is deployed on uniformly distributed random model parameters. The dark blue points correspond to an *in silico* dataset derived from a nonlinear least-squared fit to measured brain data where voxels within CSF and cortical GM were excluded by masking out regions where microscopic anisotropy (Lasič et al., 2014), μFA, is lower than 0.6. The red points correspond to regions where poor accuracy is expected, *i.e.*, where the signal fraction of the relevant component ("stick " or "zeppelin " depending on the parameter) accounts for less than 15% of the total signal or, for the p_2 map, parameter vectors where the "zeppelin " component accounts for more than 85% of the total signal fraction and $|D_{\Delta;Z}| < 0.4$. The inner legends show the Pearson correlation coefficients (ρ) of the blue points. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

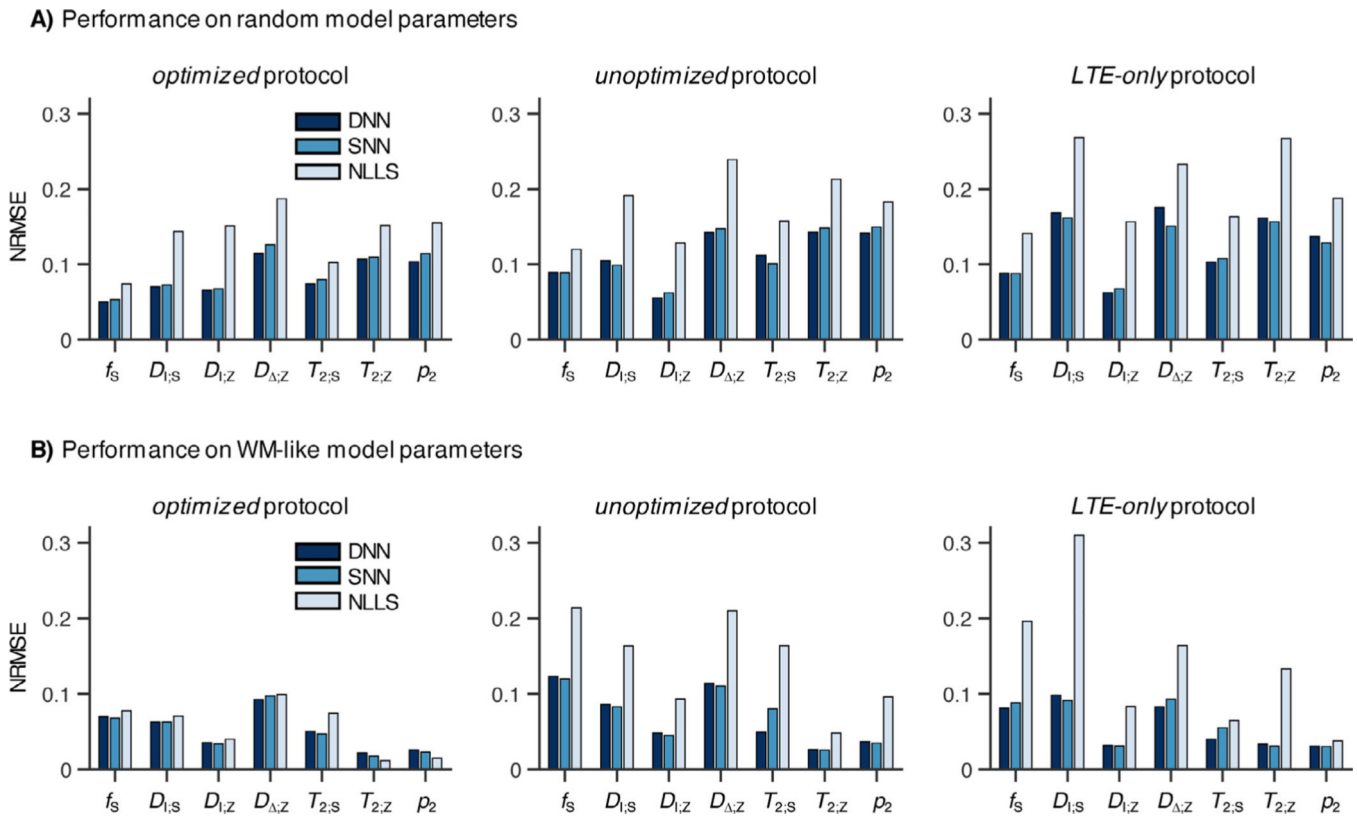

Fig. 3.

Optimized acquisition protocols result in ANN- and NLLS-based parameter estimates with smaller errors. The bar plots indicate the normalized root-mean-squared errors (NRMSE) between ground-truth and predicted parameters, for learning-based (DNN and SNN) and NLLS fitting approaches, and for in silico datasets generated with different acquisition protocols. The leftmost plots correspond to a tensor-valued (τ_E, **B**) protocol optimized for minimal parameter variance, the *optimized* protocol (Lampinen et al., 2020); the middle plots correspond to a sub-optimal tensor-valued (τ_E, **B**) protocol where relaxation-diffusion correlations are exclusively established at low b-values, the *unoptimized* protocol (Lampinen et al., 2019); the rightmost plots show the results for a (τ_E, **B**) protocol optimized for minimal parameter variance when limited to linear diffusion encoding ($b_\Delta = 1$), the *LTE-only* protocol (Lampinen et al., 2020). Panel A shows network performance on parameters sampled from a uniform distribution, and panel B shows the performance on *in silico* data based on least-squares fitting results to *in vivo* non-cortical brain tissue data.

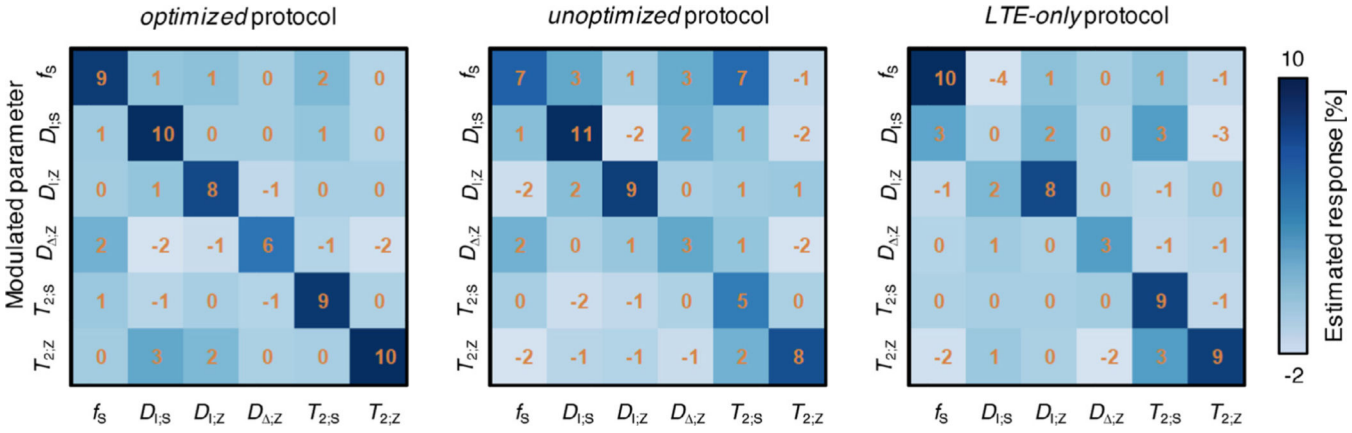

Fig. 4.

Sensitivity of acquisition protocols to 10% parameter modulations. The matrices display the relation between an induced parameter change and the observed response. When a single parameter on the y-axis is modulated by 10%, the response can be read in all other parameters along the *x*-axis. An ideal network would report a diagonal matrix with the value 10% on the diagonal, and zero otherwise. The *optimized* protocol appears sensitive in all parameters, whereas the *unoptimized* protocol lacks sensitivity $D_{\Delta;Z}$ and the *LTE-only* protocols lacks sensitivity to both $D_{\Delta;Z}$ and $D_{l;S}$.

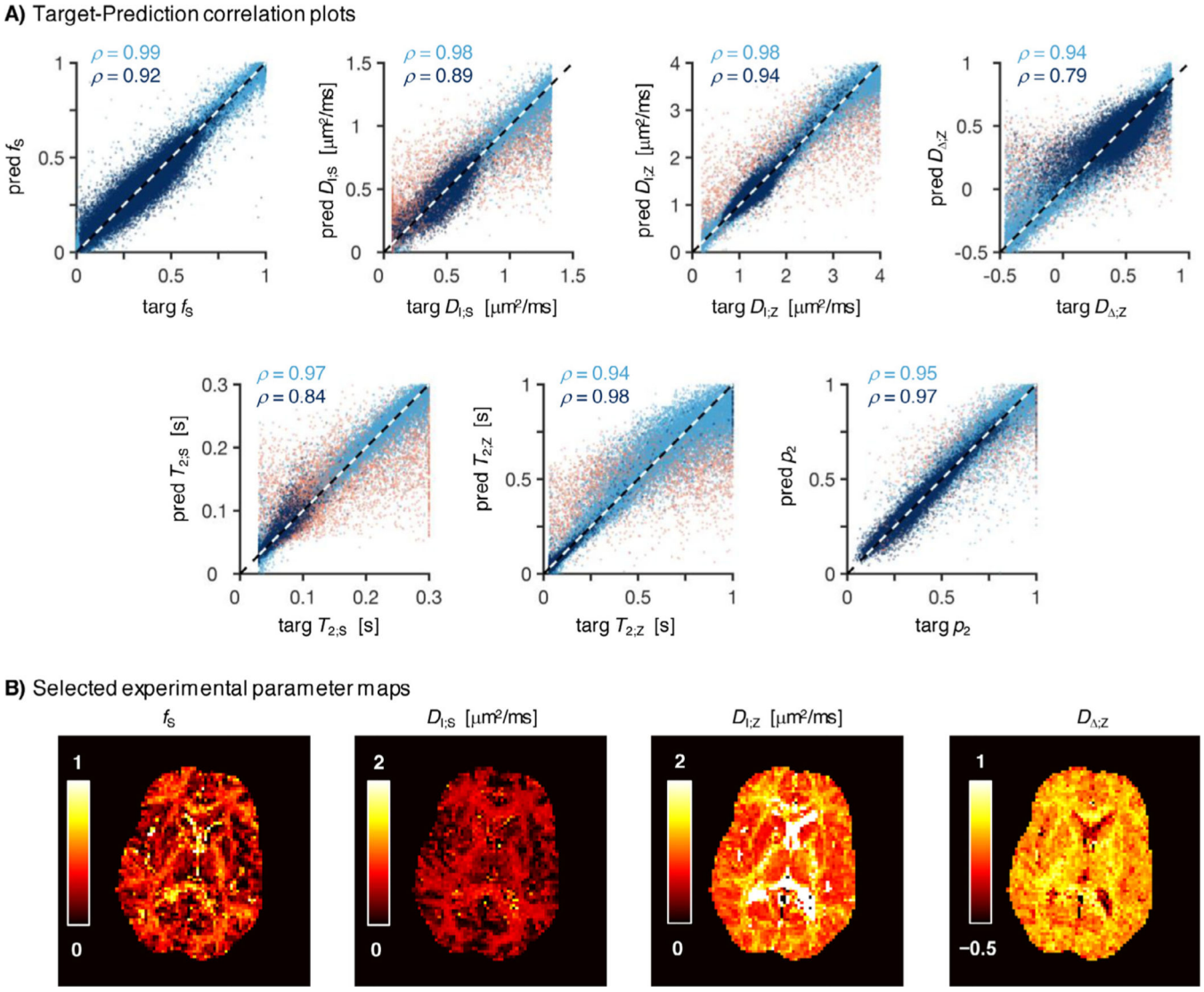

Fig. 5.

Neural network fitting of a rotationally invariant (RotInv) model results in strong target-estimate correlations and plausible maps. (A) Correlations between network-based parameter estimates and ground-truth parameter targets. The estimates were obtained from a RotInv network trained using a fraction of $f_{brain} = 0.5$ between rotationally invariant $\boldsymbol{m}_{brain}$ and $\boldsymbol{m}_{unif}$ training parameter vectors. The colour-coding and legends follow the same convention as Fig. 2. (B) Maps of microstructural diffusion parameters – f_S, $D_{l;S}$, $D_{l;Z}$ and $D_{\Delta;Z}$ – obtained from fitting a RotInv network to rotationally invariant *in vivo* brain data.

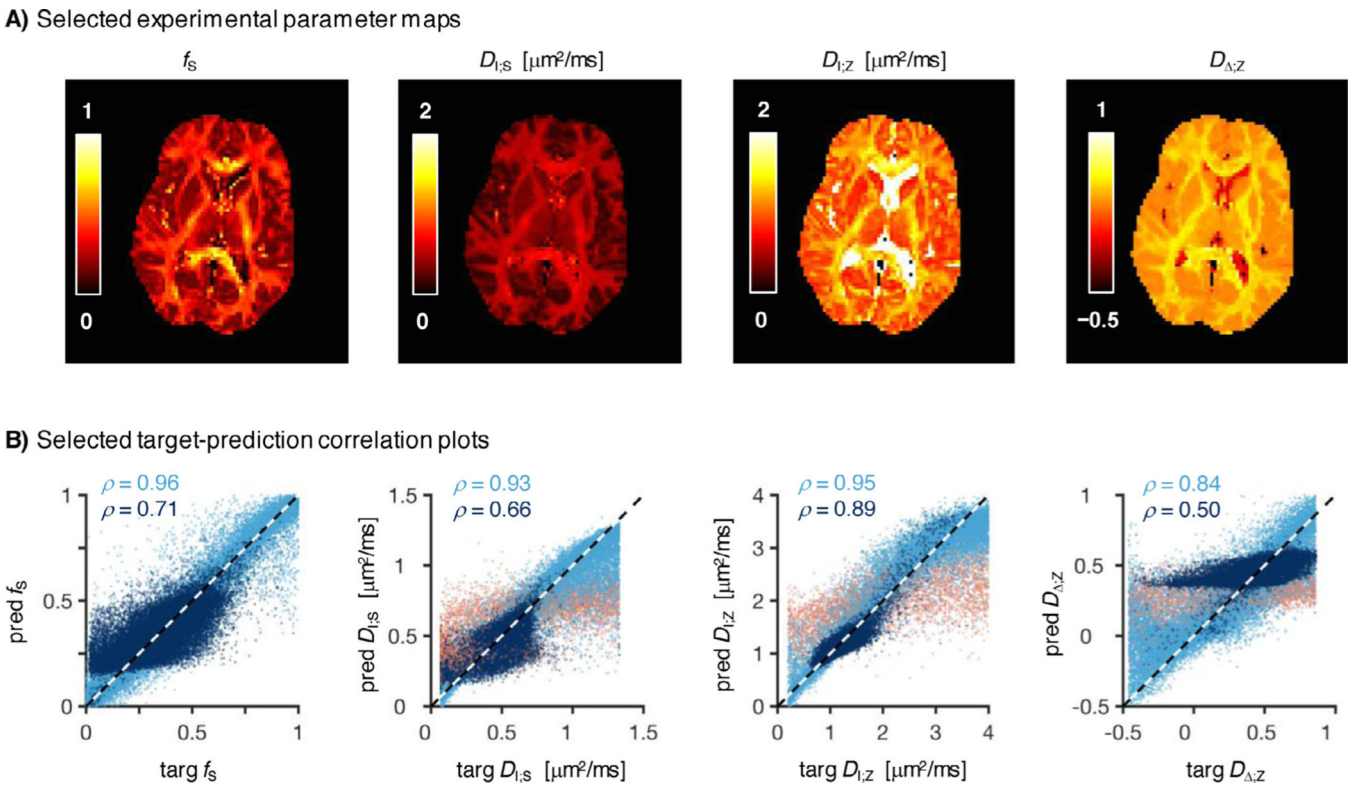

Fig. 6.

Training a neural network with an insufficient dataset may result in plausible maps but poor target-estimate correlations. **(A)** Experimental parameter maps obtained from fitting a RotInv SNN that was trained on unreasonably noisy data (SNR at S_0 in the [20, 40] range). The maps were obtained by deploying the network to rotationally invariant *in vivo* brain data. **(B)** Correlations between network-based parameter estimates and ground-truth parameter targets. The colour-coding and legends follow the same convention as Fig. 2.

Table 1

SMR parameter bounds. The diffusivity bounds were enforced by limiting $D_{\parallel;S}$, $D_{\parallel;Z}$ and $D_{\perp;Z}$ to the [0.2, 4.0] μm^2/ms interval. For $T_{2;S}$ and $T_{2;Z}$, the lower bound removes the influence of the assumedly fully-attenuated myelin water, and the large upper bound of $T_{2;Z}$ enables it to capture effects of increased values in white matter lesions (Lampinen et al., 2019) as well as possible contamination with cerebrospinal fluid which is expected to have a larger influence on the more isotropic zeppelin compartment (Lampinen et al., 2020).

Bounds	f_S	$D_{I;S}$ [μm^2/ms]	$D_{I;Z}$ [μm^2/ms]	$D_{\Delta;Z}$	$T_{2;S}$ [ms]	$T_{2;Z}$ [ms]
Minimum	0	0.07	0.2	−0.46	30	30
Maximum	1	1.33	4.0	0.86	300	1000

Table 2

Accuracy performance of DNN-, SNN- and NLLS-based fitting approaches. Performance is evaluated on synthetic data simulated from two different sets: uniformly sampled random parameters ($\boldsymbol{m}_{\text{unif}}$), and parameters derived from least-squared model fitting to *in vivo* WM and deep GM data ($\boldsymbol{m}_{\text{fit;WM-like}}$).

Metric	Dataset	Fitting method	Fitting time [s]	f_S	$D_{I;S}$	$D_{I;Z}$	$D_{\Delta;Z}$	$T_{2;S}$	$T_{2;Z}$
NRMSE	$\boldsymbol{m}_{\text{fit;WM-like}}$	NLLS	967	0.07	0.07	0.04	0.1	0.08	0.01
		SNN	0.2	0.07	0.08	0.03	0.1	0.07	0.02
		DNN	0.3	0.07	0.08	0.03	0.09	0.07	0.02
	$\boldsymbol{m}_{\text{unif}}$	NLLS	2201	0.08	0.21	0.18	0.21	0.16	0.21
		SNN	0.1	0.05	0.11	0.10	0.15	0.12	0.14
		DNN	0.5	0.05	0.11	0.10	0.14	0.11	0.14
ρ	$\boldsymbol{m}_{\text{fit;WM-like}}$	NLLS	967	0.90	0.87	0.91	0.78	0.78	0.98
		SNN	0.2	0.90	0.84	0.93	0.70	0.76	0.96
		DNN	0.3	0.91	0.85	0.92	0.73	0.77	0.95
	$\boldsymbol{m}_{\text{unif}}$	NLLS	2201	0.97	0.79	0.78	0.71	0.85	0.77
		SNN	0.1	0.98	0.93	0.94	0.86	0.92	0.87
		DNN	0.5	0.99	0.93	0.94	0.88	0.92	0.88

Neuroimage. Author manuscript; available in PMC 2022 November 11.