

Title: A generative model of realistic brain cells with application to numerical simulation of the diffusion-weighted MR signal

Authors: Marco Palombo ^{a*}, Daniel C Alexander ^a, Hui Zhang ^a

Affiliations:

^a Centre for Medical Image Computing and Dept of Computer Science, University College London, London, UK.

***Corresponding Author:** Dr. Marco Palombo, Department of Computer Science and Centre for Medical Image Computing, University College of London, Gower Street, London WC1E 6BT, United Kingdom. E-mail: mrc.palombo@gmail.com.

Abstract

To date, numerical simulations of the brain tissue have been limited by their lack of realism and flexibility. The purpose of this work is to propose a controlled and flexible generative model for brain cell morphology and an efficient computational pipeline for the reliable and robust simulation of realistic cellular structures with application to numerical simulation of intra-cellular diffusion-weighted MR (DW-MR) signal features. Inspired by the advances in computational neuroscience for modelling brain cells, we propose a generative model that enables users to simulate molecular diffusion within realistic digital brain cells, such as neurons, in a completely controlled and flexible fashion. We validate our new approach by showing an excellent match between the morphology (no statistically different 3D Sholl metrics, $P > 0.05$) and simulated intra-cellular DW-MR signal (mean relative difference $< 2\%$) of the generated digital model of brain cells and those of digital reconstruction of real brain cells from available open-access databases. We demonstrate the versatility and potential of the framework by showing a select set of examples of relevance for the DW-MR community. The computational models introduced here are useful for synthesizing intra-cellular DW-MR signals, similar to those one might measure from brain metabolites DW-MRS experiments. They also provide the foundation for a more complete simulation system that will potentially include signals from extra-cellular compartments and exchange processes, necessary for synthesizing DW-MR signals of relevance for DW-MRI experiments.

Keywords: diffusion, Monte Carlo simulation, intracellular space, cell structure, tissue microstructure, brain, diffusion-weighted MRI.

Highlights:

- We propose a generative model for controlled and flexible modelling of brain cells
- We simulate molecular diffusion inside digital cells with an unprecedented level of realism
- Relevant examples for simulation of diffusion-weighted MR signal are shown
- Excellent match between the morphology of real and digital brain cell models
- Excellent match between the DW-MR signal from real and digital brain cell models

MAIN

1. Introduction

Virtual histology is an emerging paradigm in medical imaging. The ultimate goal is to estimate microscopic tissue properties at the macroscopic scale using non-invasive imaging techniques, such as MRI. The current generation of non-invasive microstructure imaging techniques are rapidly becoming part of the mainstream package of imaging tools used routinely in clinical studies and exams (1-7). They primarily employ diffusion-weighted MRI (DW-MRI), which uses pulsed magnetic field gradients to sensitize the MR image to the displacement pattern of particles, usually water molecules, within tissue (8, 9). In particular, self-diffusion of MR visible molecules in a magnetic field gradient generates a signal loss that depends on the characteristics of the gradients as well as on tissue features that hinder or restrict the diffusion process over time, like cell membranes. Thus, it is possible to infer microscopic tissue features from the macroscopic signal loss measured by DW-MR.

Unfortunately, the relationship between the DW-MR signal and the microstructure of complex biological tissues like the brain is still not well understood. Over the last decade, a wide variety of mathematical and biophysical models have been proposed to describe this relationship (4, 7). However, the validity of their underlying model assumptions is still under debate. In fact, while the tissue microstructure is highly complex, the signal is so simple that the mapping from signal to microstructure is highly under-constrained. As a result, the more complex the system, the more challenging the identification of the most relevant features (10). Objective data-driven assessment of these assumptions, which is often hard to conduct experimentally, remains an important, yet unsolved, challenge.

In this context, advanced numerical simulations can provide a powerful tool to test the limits of a specific biophysical model or theory (11-13), aid the design of optimized experimental strategies (14-16); or even provide the basis of computational inverse models using modern machine learning techniques (17, 18). Numerical simulations and numerical phantoms play a unique role in validation that is complementary to other forms of phantoms (physical, *in vitro*, *ex vivo* and *in vivo*). Experiments with physical phantoms are expensive and time-consuming to set up, and often lack flexibility and sufficient realism. Experiments with fixed or excised-viable tissue lack a well-defined ground truth. On the other hand, accurate and complex numerical simulations can provide a well-defined ground truth mapping

between the relevant microstructural features and the diffusion MR signal. Although they necessarily represent a model of the real world based on our current understanding, they provide a framework for validation that is much more controlled (with known ground truth) than *in vitro*, *ex vivo* and *in vivo* phantoms, and is much more flexible than physical phantoms.

A key limitation of simulations to date is that they are too simplistic or inflexible: those based on geometric primitives are not realistic; and those derived from histological images lack flexibility. The state-of-the-art simulators for DW-MR signal are based on Monte-Carlo methods to simulate spins' diffusion within 3D digital models (substrates) representing the tissue (19, 20). These simulators are in principle able to manage complex substrates, but their usual applications have been limited to just configurations of simple geometric primitives such as cylinders and spheres (19, 21-28). Despite their simplicity, these models have been very useful to study different characteristics of brain white matter (WM) tissue. In fact, brain WM is comprised mostly of myelinated axonal bundles, and simple geometries like packed cylinders of poly-dispersed radii represent a valuable first-order approximation. For instance, these kind of models informed the analysis of DW-MRI measurements in both healthy and diseased conditions, helping investigating the contribution of different WM tissue features, like axonal permeability (17, 22, 28), undulations (29), beadings (30), fiber crossing (31), and so on. Few examples exist of generating more realistic substrates for WM. Some of them use complex 3D meshes reconstructed from histological images (32). Nevertheless, these approaches do not enable users to investigate, in a controlled fashion, all the possible geometrical configurations of a complex tissue microenvironment. To address this issue, a novel algorithm to design more realistic membrane geometries, better mimicking the structure of brain WM axonal bundles, has been recently proposed and embedded in the Diffusion Microscopist Simulator (DMS) simulator (33). However, there is still lack of a proper computational framework that enables the creation of realistic numerical phantoms of *any* brain tissue microenvironment.

In particular, gray matter (GM) is still one of the most challenging microarchitectures to simulate. In fact, GM is comprised of complex shaped structures like brain cells (e.g. neurons and glia) densely packed together. In order to simulate a realistic substrate mimicking GM, first of all it is necessary to have a realistic digital model for the different cells in the brain. This is extremely challenging, because brain cells are complex branched structures, comprised of different connected parts, like cell body (namely soma) and cellular projections (namely neurites). Moreover, high quality meshes, ensuring correct

connectivity between these distinct compartments, are often essential to obtain accurate simulations, resulting in an exponential increase of the computational complexity. To date, only a few attempts to simulate more realistic brain cell structures for DW-MR applications have been published (34-37). However, they still rely on a simplistic description of the cell structure, for instance as 1D branched structure (36) and disconnected (34, 37) or connected cylinders (35).

Here we present new algorithms enabling, for the first time, the construction of ultra-realistic brain cell micro-environments and the execution of diffusion simulations within them. Specifically, we introduce a novel generative model to design realistic digital substrates of brain cells. We address two main challenges in performing numerical simulation of diffusion process within realistic 3D digital representation of brain cell:

- 1) Handling the large complexity of brain cell morphology, which requires a convenient digital representation that relies on a small set of controllable features, providing realism and flexibility;
- 2) Ensuring correct connectivity between the distinct compartments comprising the system (like each branch of a dendritic tree, cellular projections and soma, different cellular entities, etc.) while minimizing the computational burden.

We validate our new approach by showing an excellent match between the morphology of the generated digital model of brain cells and those of digital reconstruction of real brain cells from available open-access databases. As consequence of the high morphological similarity of the digital representations of cellular structures at multiple length scales, we also show high similarity in the corresponding synthetic DW-MR signals.

The purpose of the present work is to introduce the new generative model and computational pipeline and provide proof-of-concept applications in DW-MR. Thus, the usage of the proposed computational framework for specific hypothesis testing and/or experimental design will be subject of future works, including eventually direct comparison with experimental data.

2. Material and Methods

Here we propose a novel generative model to design realistic digital substrates of brain cells and an efficient and robust computational pipeline that enables users to perform numerical simulation of many particles diffusing *inside* ultra-realistic brain cellular structures.

2.1 Computational pipeline

The proposed pipeline is implemented in MATLAB (the Mathworks) and Python. In particular, it is designed to interface optimally with i) open-access databases of brain cell morphology, such as NEUROMORPHO (neuromorpho.org) and the Allen Brain Atlas (<https://www.brain-map.org>), ii) CAMINO (www.camino.org.uk) for the robust and reliable synthetic DW-MR signal computation, and iii) standard toolboxes to visualize and analyse neuronal cell morphology like the TREES toolbox for MATLAB (www.treestoolbox.org).

A block diagram of the pipeline is shown in **Figure 1**, where the green blocks represent the original contribution of this work: new algorithms to construct ultra-realistic digital models of brain cells. Specifically, the cell-skeleton generator and reader blocks are implemented either to generate digital models of brain cells by using the proposed generative model (described in further details in the next section), or read digital reconstructions of real cells from experimental data or open-access databases. An efficient 3D surface mesh builder block is then designed to convert each skeletonized digital model in a 3D triangular mesh, suitable for interfacing with CAMINO or other toolboxes used in computational neuroscience, like TREES. Finally, the proposed computational pipeline is designed to accommodate a “synthetic tissue generator” module (see **Figure 1**) that can include different algorithms for context-aware packing of the digital cells, either generated or imported from real data. However, in its current implementation, the synthetic tissue generator can be used to simulate only the intra-cellular molecular diffusion. Planned future developments will enable dense cellular packing and membrane permeability, in order to simulate ultra-realistic DWI voxels of the brain tissue, also considering the extra-cellular contribution.

The modular structure of the pipeline guarantees that the synthetic substrate generator, the diffusion simulator and the DW-MR synthesizer operate independently, which can be an advantage due to the challenging task of creating a suitable complex geometry for spatial simulations. To compliment these

possibilities, compatibility with current developed standards such as Stockley-Wheal-Cole (SWC) (38), Polygon File Format (PLY) and Standard Triangle Language (STL) is also provided. Moreover, the modular structure supports straightforward parallelization of each block's tasks leading to substantial computational performance-boosts.

The pipeline accepts as input either a pre-built cellular skeleton, e.g. like those available on NEUROMORPHO, or an arbitrary skeleton built from scratch. A cellular skeleton generator based on an extension of the statistical model for complex cell morphology characterization introduced in (36) is implemented.

2.2 A generative model of brain cell morphology

To model complex cell structures with correct connectivity, we describe each cellular compartment as a branched structure (*backbone*), whose individual branch has a specific radius r_{segment} and each process (collection of branches sharing the same parent) radiates from the cell body (soma) of specific radius r_{soma} . The ensemble of backbone and sizes defines our cellular *skeleton*. The skeleton of digital cells can be either imported into the proposed framework from available public databases of real brain cell morphology or generated using the cellular skeleton generator provided. An example of 3D cell backbone and skeleton for a real Purkinje cell from NEUROMORPHO is shown in **Figure 2**.

The cell generator enables us to define realistic cellular morphology *a priori* and thus to investigate different cellular/tissue scenarios/conditions in a controlled fashion. In order to obtain realistic brain cell structures with controllable priors on the cell morphology, it is necessary to develop a generative algorithm that respects both the cell morphology priors and natural laws regulating neuronal branching. We achieve this in two steps.

First, extending the paradigm introduced in (36) from three 1D morphological features to twelve 3D ones, we incorporate a realistic 3D model of cellular structure, taking into account soma and cell fiber size, curvature, bifurcation angle and secondary structures like spines and leaflets. The overarching cellular architecture is defined by four morphometric statistics (each defined by a mean and standard deviation (SD)) accounting for the characteristic "tree-like" structure of neurons and glia: the number of projections N_{proj} leaving the soma (e.g., the dendrites or glial processes), the number of successive

embranchments (bifurcations) N_{branch} along each process, the segment length L_{segment} (in micrometers) for a given segment of process joining two successive branching points and the bifurcation angle θ . For each of these statistics, a Gaussian distribution conditioned to have positive values was assumed, parameterized by SDs $SD_{N_{\text{proj}}}$, $SD_{N_{\text{branch}}}$, $SD_{L_{\text{segment}}}$, SD_{θ} (e.g. see **Figure 1**). Note that the morphometric statistics can be defined by using any kind of distribution, e.g. not necessarily Gaussian, but also Gamma or log-normal. This can be defined by the users, according to their specific requirement/application. This statistical model defines the basic backbone, that fully respects the desired priors on the cell morphology. In order to define the final skeleton, r_{soma} and r_{segment} can be arbitrarily defined.

The skeleton is built iteratively, node after node starting from the soma:

1. A new node is drawn randomly according to the segment length and branching angle statistics;
2. A new segment, connecting the two nodes, is built.
3. The distance between the new segment and any other existing one that does not share any node with the new segment is computed by line-to-line distance, as described in (39):

- New segment $\mathbf{s}_i = \mathbf{x}_{\text{fin}}^i - \mathbf{x}_{\text{ini}}^i$, defined by initial node $\mathbf{x}_{\text{ini}}^i$ and final node $\mathbf{x}_{\text{fin}}^i$;
- N_{check} existing segments $\mathbf{s}_j |_{j=1 \dots N_{\text{check}}} = (\mathbf{x}_{\text{fin}}^j - \mathbf{x}_{\text{ini}}^j) |_{j=1 \dots N_{\text{check}}}$, defined by $(\mathbf{x}_{\text{ini}}^j, \mathbf{x}_{\text{fin}}^j) |_{j=1 \dots N_{\text{check}}}$ nodes $\neq (\mathbf{x}_{\text{ini}}^i, \mathbf{x}_{\text{fin}}^i)$;
- Distance D_{ij} between the new segment \mathbf{s}_i and the existing N_{check} segments $\mathbf{s}_j |_{j=1 \dots N_{\text{check}}}$:

$$D_{ij} = \frac{\left| (\mathbf{x}_{\text{ini}}^j - \mathbf{x}_{\text{ini}}^i) \cdot [(\mathbf{x}_{\text{fin}}^i - \mathbf{x}_{\text{ini}}^i) \times (\mathbf{x}_{\text{fin}}^j - \mathbf{x}_{\text{ini}}^j)] \right|}{\left| (\mathbf{x}_{\text{fin}}^i - \mathbf{x}_{\text{ini}}^i) \times (\mathbf{x}_{\text{fin}}^j - \mathbf{x}_{\text{ini}}^j) \right|} \Bigg|_{j=1 \dots N_{\text{check}}}$$

4. If it exists at least a case $j=j^*$ where $D_{ij^*} < r_{\text{segment}}^i + r_{\text{segment}}^{j^*}$, then the new node/segment are deleted and other node/segment are randomly generated following steps 1-2.
5. Steps 1-4 are iterated until the new node and corresponding segment are accepted;
6. The branching number statistics determines whether the cell projection stops at this new node, or if this new node is actually a branching point where the projection divides into two new segments, which will then be drawn according to the segment length and branching angle statistics.
7. On each new segment, steps 1-6 are repeated until a draw according to the branching number statistics imposes that there is no new division beyond the current node
8. Steps 1-7 are iterated according to the number of cellular projections radiating from the soma, drawn according to the projection number statistics.

Note that the condition at step 4. ensure that no segment intersection/overlap occurs in 3D between segments that do not share any node. The cellular structure can be made more and more complex by arbitrarily defining for each cell projection the ratio η between the Euclidean distance and the path distance of two connected nodes, and/or the radius of curvature R_c . Dendritic spines and/or astrocytic leaflets (37) can also be added at arbitrary density ρ_{sp} and size of head h_{sp} and neck n_{sp} .

Secondly, in order to improve the realism of the digital model for brain cells, the backbone is refined following the locally optimized graph approach proposed in (40). Specifically, a distance cost function is computed at every branching point, which is composed of two components and inspired by Cajal's laws of neuronal branching (40): 1) the Euclidean distance between that point and the nearest node in the graph located between the soma (root node) and that point, interpreted as a wiring cost; 2) the path length from the soma to the branching point under consideration, interpreted as a conduction time cost. A tunable parameter named balancing factor b_f (varying between 0 and 1) weighs these two components against each other: distance cost = $b_f \times$ wiring cost + $(1 - b_f) \times$ path length cost. With $b_f=0$ we have the shortest connection or minimum spanning tree network (using as little "wire" as possible), while with $b_f=1$ we get the entirely compartmentalized stellate structure, where each given point is directly connected to the soma. If the initial backbone represents the configuration that minimizes the distance cost function with the chosen b_f for each node, then it is preserved. Otherwise, a minimum spanning tree algorithm (41) is used to create the final backbone, that respects both the cell morphology priors and Cajal's laws of neuronal branching. Specifically, the nodes of the initial backbone are disconnected and considered as a cloud of possible source points in space. Then, the minimum spanning tree algorithm is used to search for the optimal path connecting each node with a new node chosen from the set of source points to minimize the distance cost function. Additional penalty to the distance cost function is given to the connections that do not satisfy the morphological priors on N_{proj} , N_{branch} , $L_{segment}$ and θ . An example of the initial backbone and the optimized one following this procedure is provided in *Supplementary Material, Figure S0*.

2.3 Modelling cell body and branching point

The skeleton is saved in SWC format, one of the most widely used formats to store information on cellular morphology. Once a cellular skeleton is provided (imported or generated), the framework generates a 3D surface mesh of the whole structure, taking into account individual branch and soma size (**Figure 2**). Because high quality meshes are essential to obtain accurate simulation results, instead of developing our own mesh generator, we make use of well validated open-source mesh generation software, specifically here BLENDER (<https://www.blender.org/>). We use the BLENDER “SWC Mesh” add-on (https://github.com/mcellteam/swc_mesher) to create a fine surface mesh of the cellular skeleton (**Figure 3-a**). The add-on implements a solid modelling technique known as constructive solid geometry. The technique, by design, enables the creation of a complex, smooth surface from simple primitives using Boolean operations. Briefly, MetaBalls (42) implemented in BLENDER are chosen as the primitive (see **Figure S1, a-c** in *Supplementary Material*). They are placed along the backbone with radius given by the skeleton and an optimized overlap ratio, chosen to obtain a smooth cellular surface from the superposition of all the MetaBalls. Specifically, the first MetaBall is placed at the root node with radius given by the information within the SWC file. Then, starting from the root node, one MetaBall after another is placed along the branch connecting each parent node to the corresponding child node, according to the order and radius described in the SWC file, with a certain distance between MetaBalls. The inter MetaBalls distance can be tuned in order to achieve a smooth surface upon superposition of all the MetaBalls. In **Figure S1 d-f**, an example of MetaBalls placed along three connected branches of a simple cell skeleton is shown. When the inter MetaBalls distance is greater than the MetaBalls’ diameter, each MetaBall is clearly distinguishable from the others (**Figure S1-e**). When the inter MetaBalls distance is smaller than the MetaBalls’ diameter (i.e. some degree of overlap between MetaBalls is allowed), then a single surface is generated by the superposition of the MetaBalls (**Figure S1-f**, where the overlap is 50% the MetaBalls diameter). We empirically determined that an inter MetaBalls distance lower than 25% the diameter of the MetaBalls is usually enough to obtain smooth surface from the superposition of all the MetaBalls. In **Figure S1-g** it is shown an example of the resulting mesh for a simple cellular skeleton. **Figure S1 g-o** illustrates that this procedure ensures that the 3D surface mesh has smooth transition at critical points that connect different branches and/or the soma, even at small bifurcation angle (**Figure S1 j-o**).

2.4 Surface mesh optimization

Because of the complexity of the cellular skeleton, the surface mesh obtained usually consists of millions of faces. In order to minimize the computational burden, a home-made python script is then used to refine the mesh by progressively smoothing, reducing and triangulating the mesh to reduce the number of faces to some thousands, while keeping the overall morphology unaltered (**Figure 3-a** and **b**). Of course, it is possible to reduce the number of faces even further. However, we experimentally evaluated that a few thousands faces are a good compromise between morphology preservation and memory load. This ultimately reduces the computational complexity of the diffusion process simulation step.

The mesh generator outputs the 3D surface mesh in standard file formats, including PLY and STL, ready to be fed into CAMINO to start the numerical simulation of the diffusion process and then the corresponding DW-MR signal computation.

2.5 Simulating diffusion process and DW-MR signal

With the support for the PLY format, which is compatible with CAMINO, the framework enables DW-MR signal synthesis with one of the most popular simulators of diffusion process and DW-MR signal. CAMINO (43, 44) is an open-source software toolkit for diffusion MRI processing, containing a powerful and validated Monte-Carlo based molecular diffusion simulator. Briefly, the Monte-Carlo simulator engine models the population of spins as random walkers in a 3-D environment. A specific user-defined DW-MR sequence is also modeled to induce specific phase accrual of the spins as they follow their diffusing trajectories, which produces specific DW-MR measurements. The simulation is used to synthesize a set of noise-free measurements from diffusing spins in a specified substrate and DW-MR sequence, to which noise can then be added. Further details on the specific algorithm used to simulate spins diffusion within a given substrate, and to compute the corresponding DW-MR signal can be found elsewhere (19).

We report here a select set of examples of relevance for the DW-MR community. The complex 3D surface mesh modelling the cell structure is fed into CAMINO to simulate the diffusion of 5×10^5 non-interacting spins (chosen to obtain high accuracy and precision on the synthetic DW-MR signal while limiting the computational burden, according to (19)), with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$ and Monte Carlo time step $\varepsilon_t = 20 \text{ ns}$. An illustrative Pulsed-Gradients-Spin-Echo (PGSE) experiment was simulated with: 30 b-values = 0-30 $\text{ms}/\mu\text{m}^2$ obtained by changing only the diffusion gradient strength, 256 directions

(uniformly distributed over a sphere) per b value, $\delta=3$ ms and 5 different Δ values per each set of b values: $\Delta=11, 26, 46, 76, 91$ ms. We note that these b and δ/Δ values are practically achievable with all pre-clinical scanners as well as with state-of-the-art human scanners, like the Connectom scanner (45).

From the intra-cellular direction-averaged DW-MR signal at $b=1$ ms/ μm^2 , the intracellular apparent diffusion coefficient (ADC) at each simulated diffusion time $t_d = \Delta\delta/3$ was computed as $\text{ADC}(t_d) = -\text{Ln}[S(b=1, t_d)/S_0]/1$, where S_0 is the signal at $b=0$.

The total computation time per cell was: ~ 1 hour using a single thread of a 2.4 GHz Intel Core i7; ~ 1.5 min parallelizing the computation on a high-performance computing cluster. This computation time should be considered only as indicative, since it depends on the complexity of the cell and the details of the mesh used, as well as the simulator setting, such as the number of spins.

2.6 Experiment I: studying the effect of mesh finish

To assess the bias introduced by the mesh finish, we performed the PGSE experiment for the representative meshes of a Purkinje cell in **Figure 3**, using both a complex mesh of $\sim 10^6$ triangles (**Figure 3-a**) and an optimized minimal mesh of $\sim 10^4$ triangles (**Figure 3-b**). To evaluate differences in the simulated intra-cellular signals as a function of b and ADC time dependences between the two mesh finishes, we computed the mean squared-difference (MSD) and the mean relative difference (MRD) as:

$$\text{MSD} = \langle [S_{\text{ref}}(b) - S(b)]^2 \rangle_b \quad [1]$$

$$\text{MRD} = \langle |S_{\text{ref}}(b) - S(b)| / |S_{\text{ref}}(b)| \rangle_b \quad [2]$$

where $\langle \dots \rangle_b$ denotes the mean computed over all the b values and S_{ref} is the signal computed from the simulation of spins diffusion in the finest mesh in **Figure 3-a**, while S is the signal computed from the simulation of spins diffusion in the reduced mesh in **Figure 3-b**. Similar quantity was also computed and reported for the ADC.

The computation time using a single thread of 2.4 GHz Intel Core i7 was ~ 50 hours for the finest mesh (**Figure 3-a**) and ~ 1 hour for the minimal mesh (**Figure 3-b**).

2.7 Experiment II: comparing simulation results in real and synthetic cell structure

We show the potential of our new generative model by generating individual cells using the cell generator with different parameters, in order to obtain different cell morphologies and complexity in a controlled fashion. Specifically, we show that it is possible to simulate different brain cell types, using a set of archetypical neurons identified by Cajal (46): Purkinje cell, granule cell, motor neuron, tripolar neuron, pyramidal cell, chandelier cell, spindle neuron and stellate cell (**Figure 4-a**).

The 3D models generated using the generative model are reported in **Figure 4-b** and the parameters used to generate them, obtained from trial-and-error, in **Table 1** and **2**. For comparison with histological “ground-truth”, the same cell types reported in **Figure 4-a** can also be imported from real reconstructions available on neruomorpho.org and the corresponding 3D generated models are reported in **Figure 4-c**.

From the reconstructed real and synthetic cell structures in **Figure 4-b** and **c**, we chose three very different morphologies (Purkinje cell, motor neuron and pyramidal spiny neuron), as representative of the cell structure heterogeneity characterizing the brain tissue. We have specifically chosen these three cell types because they capture the typical morphological features seen in all other neurons, as well as highlighting how different shapes neurons can take.

The overarching morphology of the three chosen cellular structures from real microscopy data and the corresponding synthetically generated ones are compared using dendrogram descriptors and the 3D Sholl analysis, as provided by the TREES toolbox. Dendrograms are frequently used to illustrate the arrangement and relationship of the nodes in a graph. Here we use dendrograms to show that the overall topology of the skeleton obtained from our generative model mirrors well that of real brain cells.

To assess the similarity of the simulated intra-cellular DW-MR signal features between real and synthetic cell structures, we performed the PGSE experiment for the meshes of the three selected cell types. To evaluate differences in the simulated intra-cellular signals as a function of b and ADC time dependences between the real and synthetic cell structures, we computed the MSD and the MRD using equations [1] and [2], with S_{ref} as the simulated intra-cellular signal from real microscopy data and S as the simulated

intra-cellular signal from the synthetic cell structure. Similar quantities were also computed and reported for the ADC.

We also assessed the effect of the natural randomness existing within each cell population. We compared the 3D Sholl metrics and intra-cellular DW-MR signals and ADCs for a set of 30 real and synthetically generated cell morphologies corresponding to the three cell populations selected. Mean and variance among each set of cells belonging to the same cell population was computed and used to compare the similarity between real and synthetic cell morphologies, as well as DW-MR signal features. The morphological features were kept constant for the 30 synthetic cells. However, note that some degree of randomness in the synthetic cells comes from the statistical nature of the generative model.

2.8 Experiment III: simulation results in synthetic cell structures

Finally, we show an example of current relevance for the DW-MR scientific community. We use the computational framework to investigate whether, in ideal conditions of infinite SNR and under the experimental conditions chosen (see *Section 2.5*), different cell types like Purkinje cells, motor neurons, and pyramidal spiny neurons, characterized by very different morphological features, may provide different intra-cellular DW-MR features.

The values of the 12 features chosen to generate the synthetic cell structures using the proposed generative model are reported in **Table 1** and **2** and correspond to the cell types: A, C and E.

To assess differences in the simulated intra-cellular DW-MR signal features between the three selected cell types, we performed the PGSE experiment using the corresponding three synthetic meshes and we computed the synthetic normalized intra-cellular signal as a function of b (for different t_d), and the ADC as a function of t_d . To evaluate differences in the simulated intra-cellular signals as a function of b and ADC time dependences between the three cell types, we computed the three MSD and the MRD between pairs of signals using equations [1] and [2], with S_{ref} as the simulated intra-cellular signal from one cell type (for example synthetic Purkinje cell) and S as the simulated intra-cellular signal from one of the remaining two (for example synthetic motor neuron). Similar quantities were also computed and reported for the ADC.

3. Results

3.1 Effect of mesh finish on the simulated DW-MR features

The logarithm of the intra-cellular direction-averaged DW-MR signals, normalized by the signal at $b=0$, for *Experiment I* is reported in **Figure 3-c** as a function of b for the two illustrative meshes chosen and $\delta/\Delta=3/11$ ms ($t_d = 10$ ms). In the inset in **Figure 3-c**, the diffusion time dependence of the intra-cellular ADCs for the two meshes considered are also reported. The MSD and MRD between the simulated DW-MR signals with the two meshes were found to be $MSD \sim 10^{-8}$ and $MRD \sim 1\%$. To obtain a $MSD < 10^{-8}$ (or equivalently a $MRD < 1\%$) between noise free and noisy signal, with a given finite SNR, we have estimated that $SNR \geq 10000$ is needed. This suggests that for simulations were $SNR < 10000$, the two signals are indistinguishable. Similar results were obtained concerning the diffusion time dependence of the ADCs computed for the two meshes.

For completeness, we report in *Supplementary Material*, **Figure S2**, the simulated signal for the two mesh finishes for long diffusion time, i.e. $\delta/\Delta=3/91$ ms ($t_d = 90$ ms). Also in this case of long diffusion time, we found $MSD < 10^{-8}$ and $MRD < 1\%$.

3.2 Comparing real and synthetic cell morphologies

Results of *Experiment II* concerning the comparison between particular instances and a set of real and synthetic cell structures are reported in **Figure 5** and **6**, respectively.

Specifically, the dendrogram descriptors for particular instances of real and synthetic cells are compared in **Figure 5-a**, while results from 3D Sholl analysis in **Figure 5-b**. From **Figure 5-a** we can see that the extent, complexity and width of the dendrograms from the synthetically generated cells match very well those from the real ones. Note, however, that individual dendrograms for each pair of synthetic *versus* real cells can look slightly different, due to the randomness inherent to the generation process, although they show the same overarching features overall. The good match between synthetic and real cells is further confirmed by **Figure 5-b**. Indeed, the distributions of 3D Sholl metrics from the individual instances of real and synthetically generated 3D cell structures are reported in **Figure 5-b** and were found to be not statistically significantly different (two-tailed t-test, $P>0.05$).

Results from the 3D Sholl analysis of a set of 30 cells belonging to the same population are reported in **Figure 6-a**, while the simulated population averaged DW-MR signals and ADCs are reported in **Figure 6-b** and **c**. We estimated the mean and variance of the population averaged distribution of Sholl metrics for the three real and synthetic cell populations (**Figure 6-a**). We found no statistically significant difference (two-tailed t-test, $P > 0.05$).

3.3 Comparing real and synthetic cell DW-MR features

The logarithm of the intra-cellular direction-averaged DW-MR signals, normalized by the signal at $b=0$, for *Experiment II* are reported in **Figure 5-c** as a function of b for the three illustrative cell structures chosen and $\delta/\Delta=3/11$ ms ($t_d = 10$ ms). The MSD and MRD between intracellular DW-MR signals computed from the simulation in real 3D cell structures and from the synthetically generated ones were found to be $\text{MSD} \sim 10^{-7}$ and $\text{MRD} < 2\%$ for all three cellular structures considered. To obtain a $\text{MSD} < 10^{-7}$ (or equivalently, a $\text{MRD} < 2\%$) between noise free and noisy signal, with a given finite SNR, we have estimated that $\text{SNR} \geq 4000$ is needed. This suggests that for simulations where $\text{SNR} < 4000$, the two signals are indistinguishable.

In the inset in **Figure 5-c**, the diffusion time dependence of the intracellular ADCs for the three cellular structures considered are reported. The MSD and MRD between intracellular ADCs computed from the simulation in real 3D cell structures and from the synthetically generated ones were found to be $\text{MSD} \sim 10^{-7}$ and $\text{MRD} < 2\%$ for all three cellular structures considered. Following the same argument as in the previous paragraph, for simulations where $\text{SNR} < 4000$ the two ADC time dependences are indistinguishable.

The simulated population averaged (from a set of 30 cells belonging to the same population) DW-MR signals and ADCs are reported in **Figure 6-b** and **c**. Also taking into account the variance of the population averaged DW-MR signals, we still found the signals from real and synthetic cell population indistinguishable. Similar results were found for the corresponding ADC time dependences.

3.4 DW-MR features from selected synthetic cell types

The intracellular direction-averaged signals as a function of b at different t_d for the three different cell types investigated in *Experiment III* are reported in **Figure 7**, while the intracellular ADC time dependence for the three different cell types are reported in the insets in **Figure 7**. The values of the 12 features chosen to generate the synthetic cell structures using the proposed generative model are reported as radar plots in **Figure 7**. We found that in the ideal case of infinite SNR and under the experimental conditions considered, the three cell types (Purkinje cells, motor neurons, or pyramidal spiny neurons) provide three different signatures in the b dependence of the intracellular direction-averaged signal and in the time dependence of the intracellular ADC.

The MSD and MRD for the simulated DW-MR signals between the Purkinje cell and motor neuron, and between the Purkinje cell and the spiny pyramidal neuron were found to be MSD = 0.15 – 0.25 and MRD = 600 – 1000 % (for t_d between 10 and 90 ms). This suggests that SNR > 2.5 is needed to be able to discriminate the signal associated with Purkinje cell from those associated with both motor neuron and spiny pyramidal neuron. On the other hand, the MSD and MRD for the simulated DW-MR signals between the motor neuron and the spiny pyramidal neuron was found to be MSD = 0.001 – 0.005 and MRD = 4 – 20 % (for t_d between 10 and 90 ms). This suggests that a much higher SNR (i.e. SNR > 40) is instead needed to be able to discriminate the signal associated with motor neuron from that associated with spiny pyramidal neuron.

However, it is important to notice that the signals coming from the three cell types are likely to become less distinguishable in real DW-MRI experiments, due to the variable contribution of the extra-cellular diffusing water and possible exchange between intra- and extra-cellular space as well as between cellular compartments.

For completeness, we report in the *Supplementary Material*, **Figure S3**, the comparison between the synthetic direction-averaged DW-MR signals for the three selected cell types for three diffusion times: t_d = 10, 45 and 90 ms. Also, in **Figure S4** and **S5**, we report and compare the synthetic direction-averaged DW-MR signals and ADC time dependences for all the cell types simulated in **Figure 4**.

4. Discussion

4.1 The first ultra-realistic simulator of brain cell structure for DW-MR

The main contribution of our work is the introduction of a new generative model of brain cell morphology (green blocks in **Figure 1**), by making substantial, and non-trivial, extension to previous work, like (36). Specifically, we incorporate a realistic 3D model of soma and cell fiber branching for the first time. Furthermore, it is the first time that the cellular structure is finely modelled in 3D, taking into account fiber size, curvature, bifurcation angle and secondary structures like spines and leaflets as well as Cajal's laws of neuronal branching.

Here, we demonstrated how this enables for the first time the design of numerical simulations for DW-MR of digitalized ultra-realistic brain cell structures, achieving a new standard in fidelity. The qualitative and quantitative similarity between the structure of real brain cells reconstructed from microscopy (and publicly available on NEUROMORPHO) and that of the synthetically generated ones, using the generative model, is evident from **Figures 5** and **6**. The presented framework enables us to use these digital reconstructions as the basis for a potentially unlimited range of simulations, each representing an *in silico* experiment. Here we show just an example of relevance for the DW-MR scientific community (**Figures 6** and **7**): the simulation suggests that different cell types, like Purkinje cells, motor neurons, or pyramidal spiny neurons, characterized by very different morphological features, may provide different outcomes of DW-MR experiments like time dependence of the intracellular ADC (**Figure 6c** and insets in **Figure 7**) or high b value dependence of the intracellular direction-averaged signal (**Figure 6b** and **7**). This suggests that it is possible to use the simulation framework to design and optimize DW-MR protocols to be the most sensitive to specific features of cell morphology, like for instance the cell body (soma) size and density (47), cell projections curvature (48), dendritic tree orientation dispersion (49), etc. In its current version, the new computational pipeline introduced here already provides MR scientists and engineers with a tool for exploring a large range of brain cell microstructure scenarios in a flexible and controlled fashion. Moreover, the remaining small differences (e.g. the branch lengths of motor neurons, see **Figure 6a**) could be reduced by further optimizing the tunable parameters of the generative model, by for example learning the 12 features directly from real reconstructions rather than from trial-and-error. We will explore ways to automate this aspect in future works.

4.2 A new generative model for controlled cell morphology modeling

This work introduces a new paradigm for ground-truth controlled studies of brain cell structure. The generative model introduced here enables users for the first time to study in a controlled fashion the link between DW-MR measurements and specific aspects of the brain cell morphology, like the complexity of the dendritic tree, the degree of curvature and branching of cell fibers, soma size contribution, short range disorder of synaptic boutons, spines and leaflets and much more. By describing the complex morphology of brain cells using only a small number of tunable features (12 in the current implementation), our generative model based strategy represents a unique tool for investigating selective structural alterations due to specific diseases, helping disentangling the impact of a particular disease on specific features of the cell morphology. Also, it may represent the basis for novel machine learning applications, helping pushing the current limits of biophysical modelling of experimental data. Finally, it is also possible to expand the variety of numerical simulators, for example by integrating also powerful molecular dynamic simulators like LAMMPS (<http://lammps.sandia.gov>), which would make it possible to realistically simulate the interaction of specific molecules within the cellular space with the cell membrane or with other molecules, for example water-water, water-proteins, water-lipid layer, metabolites-metabolites, etc.

4.3 Faster simulation of DW-MR features with optimized 3D mesh

As it is easy to imagine, performing numerical simulations of the dynamics of thousands of spins in the kind of ultra-realistic cell structures shown in **Figure 2** can be extremely computationally expensive. In order to minimize the computational burden, we included in our computational pipeline an optimized 3D mesher that allows us to reduce the number of faces comprising the 3D mesh of a single cell of a factor 10^2 - 10^3 , without compromising the performance of the simulation. In **Figure 3**, we provide an example of two meshes, one comprised of $\sim 10^6$ triangular faces (**Figure 3-a**) and an optimized one, comprised of only $\sim 10^4$ triangular faces (**Figure 3-b**). The overall morphology of the cellular mesh is perfectly preserved, even though its complexity is reduced of a factor 10^2 . **Figure 3-c** shows that the simulated DW-MR features computed in the two meshes are practically indistinguishable, but the use of the less complex mesh reduced the computational time on a single computational core by a factor ~ 50 . Optimized 3D meshing together with parallelization make possible simulations of up to thousands of different ultra-realistic cells in less than 5 minutes (these estimates are of course indicative, since the exact computational time depends on many factors like the specific characteristics of the computing

cluster used, the complexity of the cell and the details of the mesh used, as well as the simulator setting, such as the number of spins).

4.4 Potential applications

The computational pipeline introduced here represents a unique tool to test the validity of basic assumptions in current biophysical models used to estimate specific brain microstructural features such as neurites density and dispersion. Indeed, the full power of our new computational framework lies in hypothesis testing and experimental design. For example, to model the intracellular direction-averaged DW-MR signal in GM, often the dendritic tree of brain cells is modelled as independent randomly oriented sticks (50-52) or cylinders (34, 53). In reality, it is comprised of long curved and branching fibers. The computational framework introduced here can help validating the underlying hypotheses of these models showing whether or not, under specific experimental conditions, the effect of cell fibers branching and curvature is negligible. The ground-truth realistic digitalised models of brain cells introduced here can be hereafter devised for investigating many experimental questions that remain mostly unanswered, like (10): what causes the observed time dependence of intracellular biological water diffusivity along the fibers or in the gray matter? Is it varicosities, beads, synaptic boutons, undulations, or something else? Which of these structural units' changes in pathology could be detectable?

Taking into account its limitations (further described in the next section), this work potentially leads to interesting brain studies where the contribution of extra-cellular space, cell packing and membrane permeability can be considered negligible, like in metabolites DW-MR spectroscopy experiments (5, 36, 53-55). In fact, some brain metabolites such as (but not only) N-Acetyl-Aspartate, Myo-Inositol, Choline, Creatine are mostly intra-cellular and do not cross cell membrane. Consequently, the total signal in a spectroscopic voxel is the coarse-grained average of the intra-cellular signal coming from each individual cellular structure. The methods developed in this work, enable users to model each cellular compartment with high fidelity, opening the possibility to combine the contribution from different cell types together and helping the interpretation of DW-MR spectroscopy data in terms of cell morphology.

4.5 Limitations

The effects of T_1/T_2 relaxation have not been taken into account for all the performed simulations, although they clearly influence the SNR when long echo times are chosen. Also, although it is possible with CAMINO to simulate the actual level of noise corrupting the signal of real acquisitions, in this work we did not perform any systematic study of SNR impact since the aim of the work is not to study how different tissue microarchitecture impact the DW-MR signal, but it is to provide a proof-of-concept demonstration of the many potential applications offered by the novel simulation framework proposed. The study of how different cellular or tissue microarchitectures impact the DW-MR signal with realistic SNR will be a topic of future studies where the new simulation framework can be used and exploited at its full potential.

Moreover, for similar reasons, cell membrane permeability has not been considered in the simulations reported here. However, it is important to underline that it is possible to include cell membrane permeability in the simulation of the diffusion dynamics, using CAMINO. Furthermore, it is also possible to assign different cell membrane permeability to different cell types or cell subparts within the same substrate. This can be useful for studies aiming at investigating the effect of brain cell membrane permeability on water diffusion in a more realistic way, since it has been shown that, for instance, the membrane of glial cells and neurons have different permeability properties (56).

One more limitation of the proposed computational framework, in its current implementation, is that it can be used to study many different brain cells, but only considering them as independent and non-interacting parts of the whole brain tissue. How we can extend the current framework to multiple packed cells, including extra-cellular space, is addressed in the following section. Here, we would like to underline how, even in its currently limited implementation, the proposed computational framework already represents a valuable tool for the MRI community. It is the very first framework for ultra-realistic cellular structure simulation, designed specifically for the MRI community. As such, it represents a first step, from which starting a collaborative effort to push further the current limitations of numerical simulations, and opens the way towards a unique alternative to expensive physical phantoms and invasive/destructive histology sampling.

4.6 Future perspectives: towards an ultra-realistic simulator of the brain tissue for DW-MR

The morphologies of the brain cells are highly diverse and variant (see **Figure 4**). The variance, presumably, originates also from their constraining arrangement in a densely packed brain substrate. Thus, in order to generate a digital representation of the brain tissue realistically, it is necessary to develop an efficient context-aware cell packing algorithm. Indeed, being able to perform Monte Carlo simulations in realistic extra-cellular space taking also into account membrane permeability is an important step to be able to compare simulated data to acquired DW-MR data. We have estimated that to reach a cellular packing fraction equal to 80% (i.e. 20% volume fraction occupied by the extra-cellular space), for simulating a realistic DW-MR voxel of $200 \times 200 \times 200 \mu\text{m}^3$, a number of about 100 realistic cellular structures is needed. Multi-scale approaches such as (57, 58) can be very useful to accomplish this task by maintaining reasonable computation times. This represents the main direction of our future work.

5. Conclusion

In this work, we presented two main new contributions:

- The introduction of a novel generative model to design *controlled* and *flexible* realistic 3D digital substrates of brain cells;
- The design of a computational pipeline to perform *reliable* and *robust* numerical simulation of diffusion process within realistic 3D digital representation of brain cell, with an exemplar implementation and its performance analysed.

We showed that this computational framework addresses for the first time a major challenge in performing numerical simulation of diffusion process within realistic 3D digital representation of brain cell: ensuring correct connectivity between the distinct compartments comprising the system (like each branch of a dendritic tree, cellular projections and soma, different cellular entities, etc.) minimizing the computational burden.

We demonstrated the versatility and potentiality of the computational framework by showing a select set of examples of relevance for the DW-MR community. Current limitations were identified and discussed as well as possible future applications and developments.

This work represents the first necessary step towards enabling the investigation of how the morphology of brain cells may influence the DW-MR signal and derived metrics like the ADC. However, to perform such kinds of numerical studies, there are still some important open challenges, like dense cellular packing, extra-cellular space and membrane permeability, which go behind the scope of the current study and represent the main direction of future work.

Acknowledgment

We are grateful to Ross Callaghan for useful discussions about the use of MetaBalls and Blender “SWC mesh” add-on. This work was supported by EPSRC (EP/G007748, EP/I027084/01, EP/L022680/1, EP/M020533/1, N018702).

References

1. Assaf Y, Johansen-Berg H, & Thiebaut de Schotten M (2017) The role of diffusion MRI in neuroscience. *Nmr Biomed*.
2. Johansen-Berg H & Behrens TE (2013) *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy* (Academic Press).
3. Le Bihan D & Johansen-Berg H (2012) Diffusion MRI at 25: exploring brain tissue structure and function. *Neuroimage* 61(2):324-341.
4. Alexander DC, Dyrby TB, Nilsson M, & Zhang H (2017) Imaging brain microstructure with diffusion MRI: practicality and applications. *Nmr Biomed*.
5. Palombo M, Shemesh N, Ronen I, & Valette J (2017) Insights into brain microstructure from in vivo DW-MRS. *Neuroimage*.
6. Basser PJ & Jones DK (2002) Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *Nmr Biomed* 15(7-8):456-467.
7. Jones DK (2010) *Diffusion MRI* (Oxford University Press).
8. Callaghan PT (1991) *Principles of nuclear magnetic resonance microscopy* (Clarendon Press ; Oxford University Press).
9. Grebenkov DS (2007) NMR survey of reflected Brownian motion. *Reviews of Modern Physics* 79(3):1077-1137.
10. Novikov DS, Kiselev VG, & Jespersen SN (2018) On modeling. *Magn Reson Med* 79(6):3172-3193.
11. Panagiotaki E, *et al.* (2012) Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. *Neuroimage* 59(3):2241-2254.
12. Drobnjak I, Zhang H, Ianus A, Kaden E, & Alexander DC (2016) PGSE, OGSE, and sensitivity to axon diameter in diffusion MRI: Insight from a simulation study. *Magn Reson Med* 75(2):688-700.

13. Nilsson M, Lasic S, Drobnjak I, Topgaard D, & Westin CF (2017) Resolution limit of cylinder diameter estimation by diffusion MRI: The impact of gradient waveform and orientation dispersion. *Nmr Biomed*.
14. Alexander DC (2008) A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magn Reson Med* 60(2):439-448.
15. Drobnjak I, Siow B, & Alexander DC (2010) Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. *J Magn Reson* 206(1):41-51.
16. Ianus A, Drobnjak I, & Alexander DC (2016) Model-based estimation of microscopic anisotropy using diffusion MRI: a simulation study. *Nmr Biomed* 29(5):672-685.
17. Nedjati-Gilani GL, *et al.* (2017) Machine learning based compartment models with permeability for white matter microstructure imaging. *Neuroimage* 150:119-135.
18. Graham MS, Drobnjak I, & Zhang H (2018) A supervised learning approach for diffusion MRI quality control with minimal training data. *Neuroimage* 178:668-676.
19. Hall MG & Alexander DC (2009) Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI. *IEEE Trans Med Imaging* 28(9):1354-1364.
20. Yeh CH, *et al.* (2013) Diffusion microscopist simulator: a general Monte Carlo simulation system for diffusion magnetic resonance imaging. *PLoS One* 8(10):e76626.
21. Nilsson M, *et al.* (2009) On the effects of a varied diffusion time in vivo: is the diffusion in white matter restricted? *Magn Reson Imaging* 27(2):176-187.
22. Fieremans E, Novikov DS, Jensen JH, & Helpert JA (2010) Monte Carlo study of a two-compartment exchange model of diffusion. *Nmr Biomed* 23(7):711-724.
23. Ford JC & Hackney DB (1997) Numerical model for calculation of apparent diffusion coefficients (ADC) in permeable cylinders--comparison with measured ADC in spinal cord white matter. *Magn Reson Med* 37(3):387-394.
24. Peled S (2007) New perspectives on the sources of white matter DTI signal. *IEEE Trans Med Imaging* 26(11):1448-1455.
25. Balls GT & Frank LR (2009) A simulation environment for diffusion weighted MR experiments in complex media. *Magn Reson Med* 62(3):771-778.
26. Landman BA, *et al.* (2010) Complex geometric models of diffusion and relaxation in healthy and damaged white matter. *Nmr Biomed* 23(2):152-162.
27. Ramirez-Manzanares A, Cook PA, Hall M, Ashtari M, & Gee JC (2011) Resolving axon fiber crossings at clinical b-values: an evaluation study. *Med Phys* 38(9):5239-5253.
28. Nilsson M, *et al.* (2010) Evaluating the accuracy and precision of a two-compartment Karger model using Monte Carlo simulations. *J Magn Reson* 206(1):59-67.
29. Nilsson M, Latt J, Stahlberg F, van Westen D, & Hagglatt H (2012) The importance of axonal undulation in diffusion MR measurements: a Monte Carlo simulation study. *Nmr Biomed* 25(5):795-805.
30. Budde MD & Frank JA (2010) Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke. *Proc Natl Acad Sci U S A* 107(32):14472-14477.
31. Rensonnet G, Scherrer B, Warfield SK, Macq B, & Taquet M (2018) Assessing the validity of the approximation of diffusion-weighted-MRI signals from crossing fascicles by sums of signals from single fascicles. *Magn Reson Med* 79(4):2332-2345.

32. Panagiotaki E, *et al.* (2010) High-fidelity meshes from tissue samples for diffusion MRI simulations. *Med Image Comput Comput Assist Interv* 13(Pt 2):404-411.
33. Ginsburger K, *et al.* (2018) Improving the Realism of White Matter Numerical Phantoms: A Step toward a Better Understanding of the Influence of Structural Disorders in Diffusion MRI. *Front Phys* 6.
34. Hansen MB, Jespersen SN, Leigland LA, & Kroenke CD (2013) Using diffusion anisotropy to characterize neuronal morphology in gray matter: the orientation distribution of axons and dendrites in the NeuroMorpho.org database. *Front Integr Neurosci* 7:31.
35. Nguyen DV, Li JR, Grebenkov DS, & Le Bihan D (2014) A finite element method to solve the Bloch-Torrey equation applied to diffusion magnetic resonance imaging. *J. Comp. Phys.* 263:283-302
36. Palombo M, *et al.* (2016) New paradigm to assess brain cell morphology by diffusion-weighted MR spectroscopy in vivo. *Proc Natl Acad Sci U S A* 113(24):6671-6676.
37. Palombo M, Ligneul C, Hernandez-Garzon E, & Valette J (2017) Can we detect the effect of spines and leaflets on the diffusion of brain intracellular metabolites? *Neuroimage*.
38. Stockley EW, Cole HM, Brown AD, & Wheal HV (1993) A system for quantitative morphological measurement and electronic modelling of neurons: three-dimensional reconstruction. *J Neurosci Methods* 47(1-2):39-51.
39. Gellert W, Gottwald S, Hellwich M, Kästner H, & Küstner H (1990) The VNR Concise Encyclopedia of Mathematics. (Springer Netherlands,, Dordrecht), p 1 online resource.
40. Cuntz H, Forstner F, Borst A, & Hausser M (2010) One Rule to Grow Them All: A General Theory of Neuronal Branching and Its Practical Application. *Plos Comput Biol* 6(8).
41. Prim RC (1957) Shortest Connection Networks and Some Generalizations. *At&T Tech J* 36(6):1389-1401.
42. Blinn JF (1982) A Generalization of Algebraic Surface Drawing. *ACM Transactions on Graphics* 1(3):235-256.
43. Hall MG, Nedjati-Gilani G, & Alexander DC (2017) Realistic voxel sizes and reduced signal variation in Monte-Carlo simulation for diffusion MR data synthesis. *arXiv preprint arXiv:1701.03634*.
44. Cook PA, *et al.* (2006) Camino: open-source diffusion-MRI reconstruction and processing. In *14th scientific meeting of the international society for magnetic resonance in medicine (Vol. 2759, p. 2759)*. Seattle WA, USA.
45. Jones DK, *et al.* (2018) Microstructural imaging of the human brain with a 'super-scanner': 10 key advantages of ultra-strong gradients for diffusion MRI. *Neuroimage*.
46. Ramón y Cajal S, Pasik P, & Pasik T (1999) *Texture of the nervous system of man and the vertebrates* (Springer, Wien) pp xl,631 p.
47. Palombo M, Shemesh N, Ianus A, Alexander DC, & Zhang H (2018) A compartment based model for non-invasive cell body imaging by diffusion MRI. *Proc. Intl. Soc. Mag. Reson. Med.*; #1096.
48. Ozarslan E, Yolcu C, Herberthson M, Knutsson H, & Westin CF (2018) Influence of the size and curvedness of neural projections on the orientationally averaged diffusion MR signal. *Front Phys* 6.

49. Zhang H, Schneider T, Wheeler-Kingshott CA, & Alexander DC (2012) NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* 61(4):1000-1016.
50. Kaden E, Kruggel F, & Alexander DC (2016) Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magn Reson Med* 75(4):1752-1763.
51. Kaden E, Kelm ND, Carson RP, Does MD, & Alexander DC (2016) Multi-compartment microscopic diffusion imaging. *Neuroimage* 139:346-359.
52. Jespersen SN, Kroenke CD, Ostergaard L, Ackerman JJ, & Yablonskiy DA (2007) Modeling dendrite density from magnetic resonance diffusion measurements. *Neuroimage* 34(4):1473-1486.
53. Palombo M, Ligneul C, & Valette J (2017) Modeling diffusion of intracellular metabolites in the mouse brain up to very high diffusion-weighting: Diffusion in long fibers (almost) accounts for non-monoexponential attenuation. *Magnet Reson Med* 77(1):343-350.
54. Ligneul C & Valette J (2017) Probing metabolite diffusion at ultra-short time scales in the mouse brain using optimized oscillating gradients and "short"-echo-time diffusion-weighted MRS. *Nmr Biomed* 30(1).
55. Valette J, Ligneul C, Marchadour C, Najac C, & Palombo M (2018) Brain Metabolite Diffusion from Ultra-Short to Ultra-Long Time Scales: What Do We Learn, Where Should We Go? *Front Neurosci-Switz* 12.
56. Yang DM, *et al.* (2018) Intracellular water preexchange lifetime in neurons and astrocytes. *Magnet Reson Med* 79(3):1616-1627.
57. Hall MG, Nedjati-Gilani G, & Alexander DC (2017) Realistic voxel sizes and reduced signal variation in Monte-Carlo simulation for diffusion MR data synthesis. *arXiv preprint* arXiv:1701.03634
58. Grebenkov DS (2011) A fast random walk algorithm for computing the pulsed-gradient spin-echo signal in multiscale porous media. *J Magn Reson* 208(2):243-255.

Tables

Table 1. Set of tunable parameters describing the morphology of the 3D cellular structures in **Figure 4b**. N_{proj} : number of cellular projections radiating from the soma; N_{branch} : number of consecutive bifurcations; $L_{segment}$: length of each individual segment comprising the cellular projections (in μm); θ : bifurcation angle (in radians); R_c : radius of curvature of individual segments (in μm); η : direct over path ratio of individual segments; b_f : balancing factor in the extension to the minimum spanning tree algorithm.

Cell type	N_{proj} (mean \pm SD)	N_{branch} (mean \pm SD)	$L_{segment}$ (mean \pm SD)	θ (mean \pm SD)	R_c (mean \pm SD)	η (mean \pm SD)	b_f
A purkinje cell	1 \pm 0	10 \pm 2	10 \pm 5	$\pi/6 \pm \pi/8$	500 \pm 100	0.80 \pm 0.1	0.5
B granule cell	2 \pm 0	3 \pm 1	20 \pm 10	$\pi/8 \pm \pi/16$	100 \pm 20	0.70 \pm 0.1	0.5
C motor neuron	20 \pm 5	6 \pm 1	80 \pm 20	$\pi/4 \pm \pi/16$	500 \pm 100	0.85 \pm 0.1	0.5
D tripolar neuron	2 \pm 0	3 \pm 1	30 \pm 10	$\pi/4 \pm \pi/8$	500 \pm 100	0.85 \pm 0.1	0.5
E pyramidal cell	10 \pm 5	4 \pm 1	60 \pm 30	$\pi/8 \pm \pi/8$	500 \pm 100	0.95 \pm 0.1	0.5
F chandelier cell	10 \pm 5	3 \pm 1	40 \pm 10	$\pi/4 \pm \pi/16$	100 \pm 20	0.70 \pm 0.1	0.5
G spindle neuron	2 \pm 0	2 \pm 1	30 \pm 15	$\pi/16 \pm \pi/32$	100 \pm 20	0.90 \pm 0.1	0.5
H stellate cell	8 \pm 2	3 \pm 1	40 \pm 10	$\pi/4 \pm \pi/16$	50 \pm 10	0.60 \pm 0.1	0.5

Table 2. Set of tunable parameters describing the sizes of the 3D cellular structures in **Figure 4b**. r_{soma} : radius of the cell body, namely soma (in μm); $r_{segment}$: radius of each individual segment comprising the cellular projections (in μm); ρ_{sp} : density of spines/leaflets (in μm^{-1}); h_{sp} : radius and length of the spines/leaflets head (in μm); n_{sp} : radius and length of the spines/leaflets neck (in μm).

Cell type	r_{soma}	$r_{segment}$ (mean \pm SD)	ρ_{sp} (mean \pm SD)	h_{sp} (mean \pm SD)	n_{sp} (mean \pm SD)
A purkinje cell	25	1 \pm 0.5	0 \pm 0	n. a.	n. a.
B	8	0.5 \pm 0.1	0 \pm 0	n. a.	n. a.

granule cell					
C	10	0.5 ± 0.1	0 ± 0	n. a.	n. a.
motor neuron					
D	8	0.5 ± 0.1	0 ± 0	n. a.	n. a.
tripolar neuron					
E	25	0.5 ± 0.1	20 ± 10	0.5 ± 0.25	0.12 ± 0.06
pyramidal cell					
F	5	0.5 ± 0.1	0 ± 0	n. a.	n. a.
chandelier cell					
G	12	0.5 ± 0.1	0 ± 0	n. a.	n. a.
spindle neuron					
H	10	0.5 ± 0.1	0 ± 0	n. a.	n. a.
stellate cell					

Figures captions

Figure 1. (Center) Schematic of the proposed computational framework. The toolbox we developed is comprised of the green blocks. It accepts as input either a pre-built cellular skeleton, like those available on NEUROMORPHO (neuromorpho.org), or an arbitrary skeleton built from scratch using a novel generative model. (Left) We show some examples of using the digital cell generator to generate cell structures of increasing complexity in a controlled fashion from morphometric statistics *priors*: different values of mean and standard deviation of L_{segment} , N_{branch} and N_{process} , and fixed values of θ , radius of branches and soma. (Right) The toolbox outputs a 3D mesh of the whole digital cell (in the figure the scale-bars refer to 100 μm). This 3D is saved in suitable format for interfacing with CAMINO or other toolboxes used in computational neuroscience, like TREES.

Figure 2. Example of 3D cell backbone, skeleton and mesh for a reconstructed Purkinje cell from the NEUROMORPHO database (neuromorpho.org). The color encodes cellular fibers and soma diameter (in μm).

Figure 3. a) complete triangular mesh ($\sim 10^6$ triangular faces) obtained from the 3D mesher block from the reconstructed Purkinje cell reported in **Figure 2**. b) reduced mesh ($\sim 10^4$ triangular mesh) after progressive smoothing, reduction and triangulation of the more complex mesh in a). c) Direction-averaged normalized DW-MR signal as a function of b for the two meshes in a) and b), as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with 30 b -values = 0-30 $\text{ms}/\mu\text{m}^2$ obtained by changing only the diffusion gradient strength, 256 directions (uniformly distributed over a sphere) per b value, $\delta/\Delta = 3/11$ ms. c) - inset, the ADC as function of diffusion time as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with $b = 1 \text{ ms}/\mu\text{m}^2$, 256 directions (uniformly distributed over a sphere), $\delta = 3$ ms and 5 different t_d values per each set of b values: $t_d = 10, 25, 45, 75, 90$ ms.

Figure 4. a) Set of archetypical neurons identified by Cajal (46): (A) Purkinje cell, (B) granule cell, (C) motor neuron, (D) tripolar neuron, (E) pyramidal cell, (F) chandelier cell, (G) spindle neuron and (H) stellate cell. b) 3D digital model of the same cell types in a) as generated using the generative model with the parameters in **Table 1** and **2**.

c) 3D reconstruction from histological data of the same cell types in a) from open-source database neuromorpho.org.

Figure 5. a) Comparison of the dendrograms of the morphology of three selected cell types (Purkinje cells, motor neuron and spiny pyramidal neuron) obtained from 3D reconstruction from histological data of real cells (named Real) and the digital model obtained using the generative model (named Synthetic). b) Comparison of the probability density distributions of 3D Sholl metrics for the real and synthetic cells. c) Comparison of the direction-averaged normalized DW-MR signal as a function of b for the three real and synthetic cells, as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with 30 b -values = 0-30 $\text{ms}/\mu\text{m}^2$ obtained by changing only the diffusion gradient strength, 256 directions (uniformly distributed over a sphere) per b value, $\delta/\Delta = 3/11$ ms. c) - *inset*, Comparison of the ADC as function of diffusion time as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with $b = 1 \text{ ms}/\mu\text{m}^2$, 256 directions (uniformly distributed over a sphere), $\delta = 3$ ms and 5 different t_d values per each set of b values: $t_d = 10, 25, 45, 75, 90$ ms.

Figure 6. a) Comparison of the probability density distributions of 3D Sholl metrics for a set of 30 real and synthetic cells per cell type. Red lines represent average distributions among the set of cells. Mean μ and variance σ of the average distributions are also reported for comparison b) Comparison of the mean direction-averaged normalized DW-MR signal (error bars represent standard deviation) for each set of cell populations as a function of b for the three real and synthetic cells, as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with 30 b -values = 0-30 $\text{ms}/\mu\text{m}^2$ obtained by changing only the diffusion gradient strength, 256 directions (uniformly distributed over a sphere) per b value, $\delta/\Delta = 3/11$ ms. c) Similar comparison of the ADC as function of diffusion time (mean \pm standard deviation over the set of cells per each population) as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with $b = 1 \text{ ms}/\mu\text{m}^2$, 256 directions (uniformly distributed over a sphere), $\delta = 3$ ms and 5 different t_d values per each set of b values: $t_d = 10, 25, 45, 75, 90$ ms.

Figure 7. Example of possible application of the proposed simulation framework to investigate the DW-MR signal features originating from three very different digital cell models (corresponding to Purkinje cells, motor neuron and spiny pyramidal neuron) obtained using the controlled and flexible generative model using the parameters value reported in the radar plots and in **Table 1** and **2**. Direction-averaged normalized DW-MR signal as a function of b and ADC as function of diffusion time (*inset*) for the three digital model of brain cells, as computed from the simulation of 5×10^5 non-interacting spins, with diffusivity $D_0 = 2 \mu\text{m}^2/\text{ms}$, and a PGSE sequence with 30 b -values = 0-30 $\text{ms}/\mu\text{m}^2$ obtained by changing only the diffusion gradient strength, 256 directions (uniformly distributed over a sphere) per b value, $\delta = 3$ ms and 5 different t_d values per each set of b values: $t_d = 10, 25, 45, 75, 90$ ms.

Figures

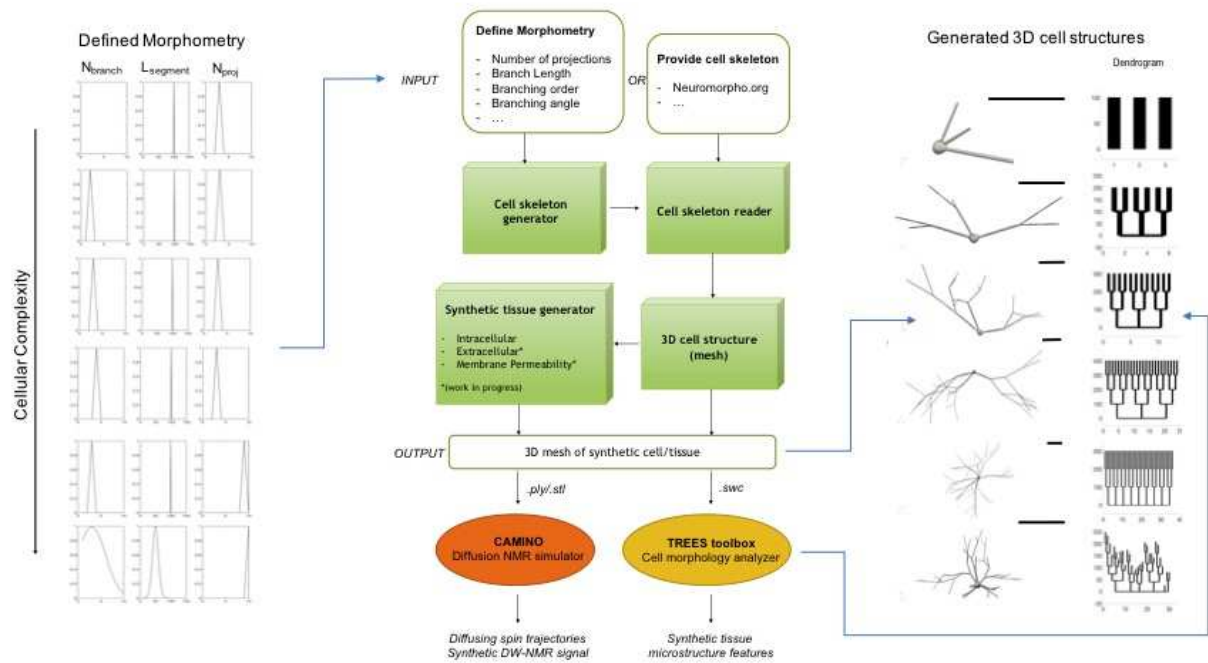


Figure 1.

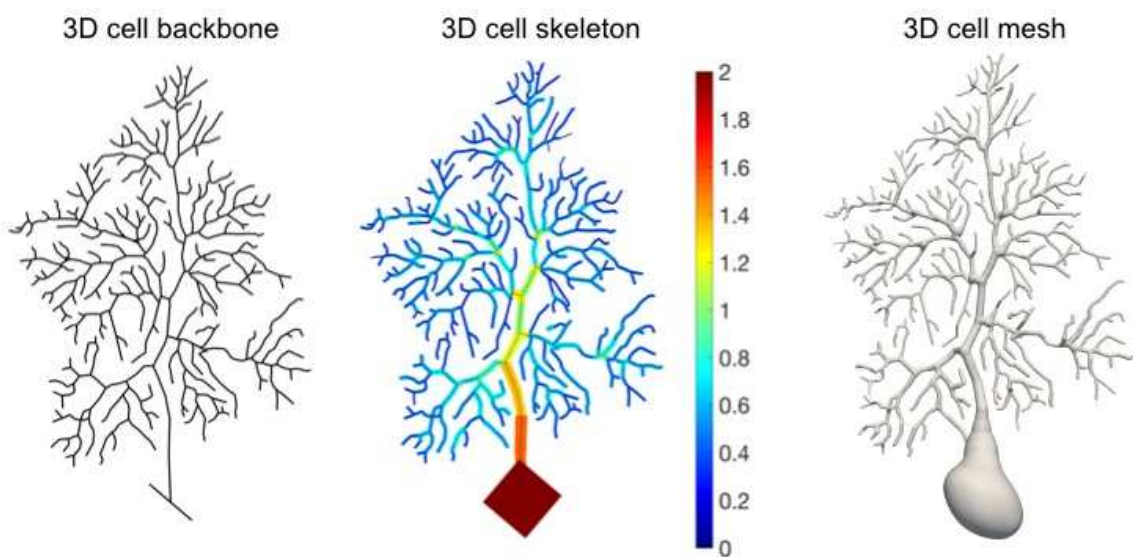


Figure 2.

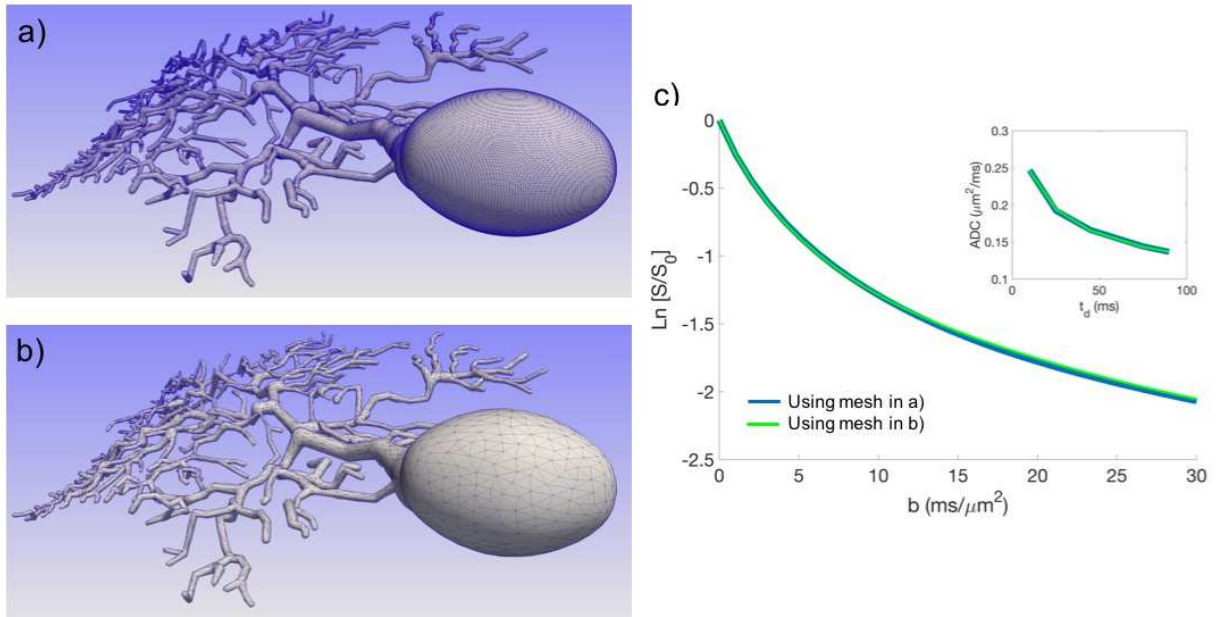


Figure 3.

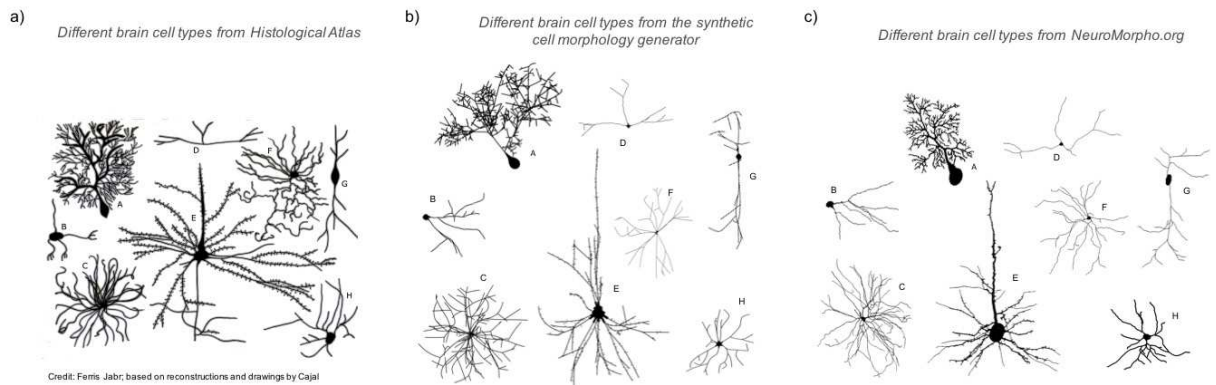


Figure 4.

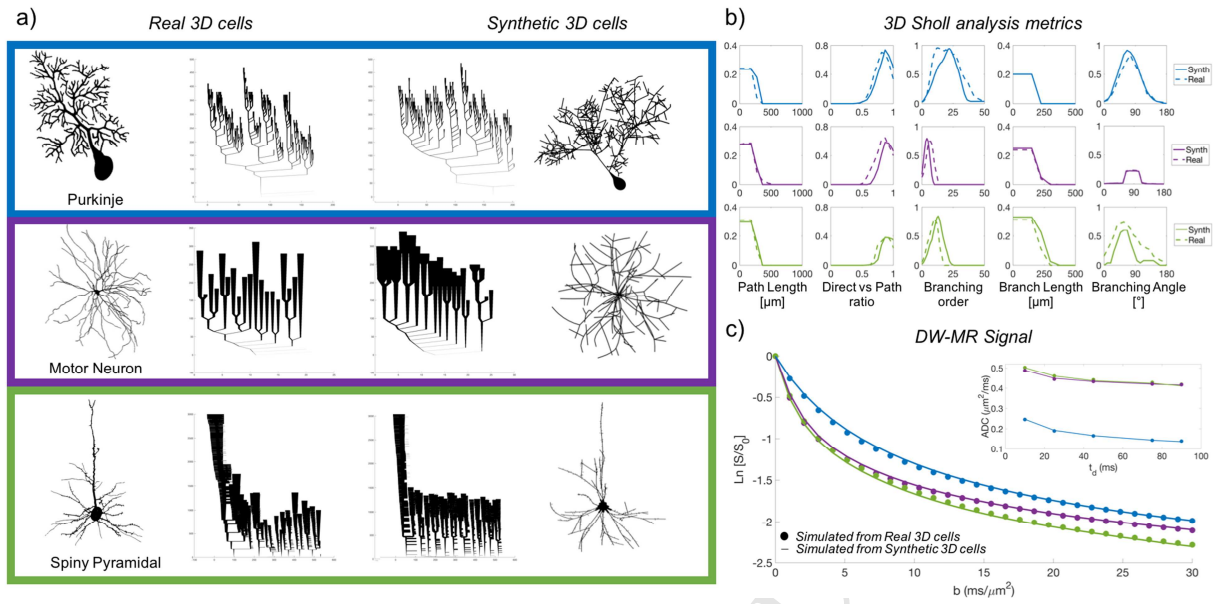


Figure 5.

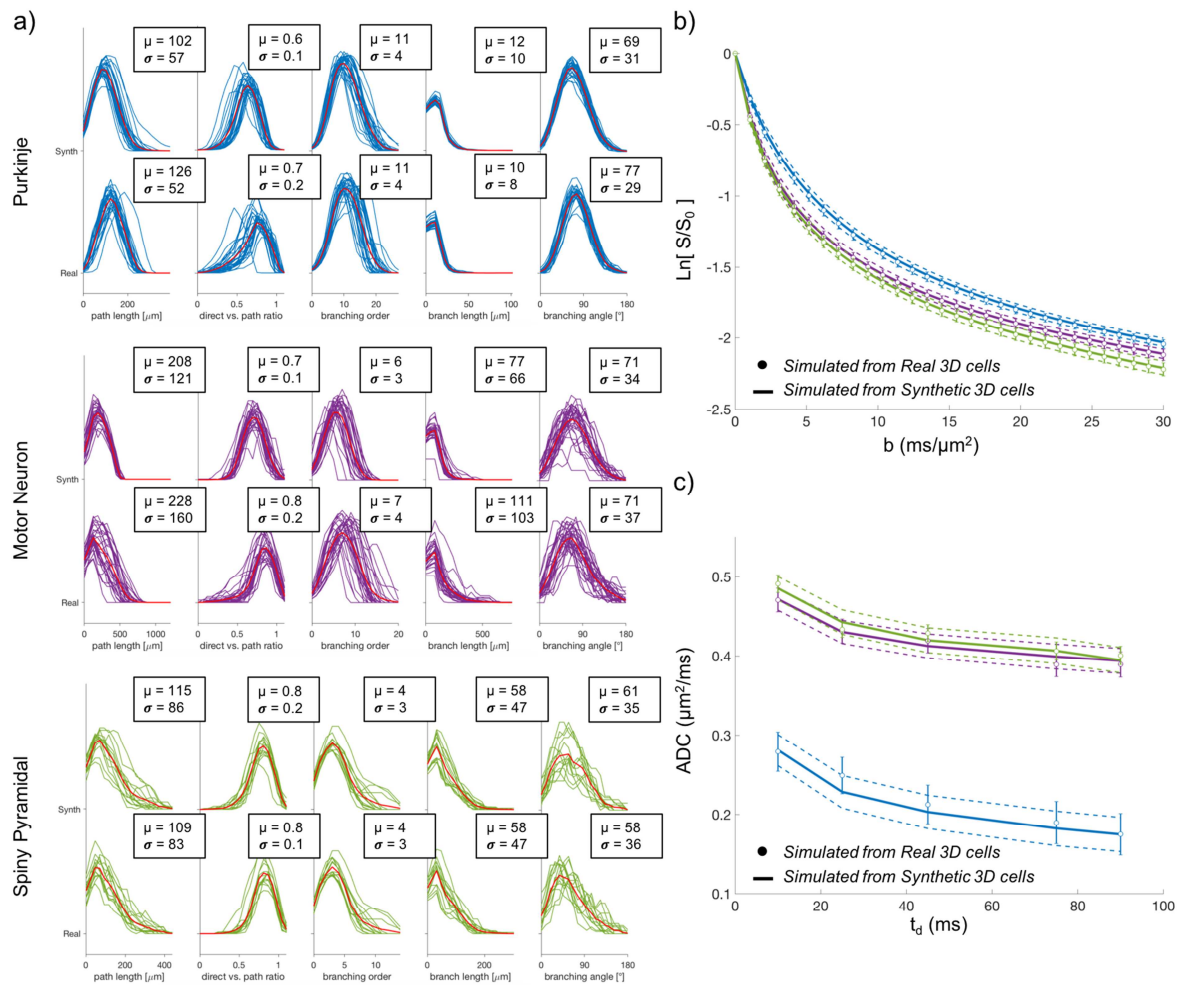


Figure 6.

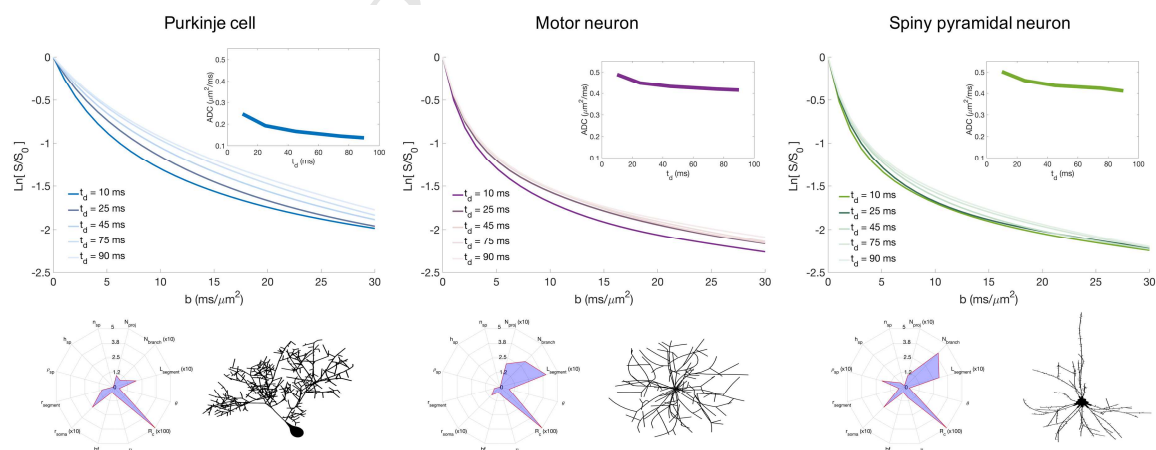


Figure 7.