MRS-Sim: Open-Source Framework for Simulating In Vivo Clinical MR Spectroscopy Data

J. T. LaMaster[1], D. Das[2], G. Oeltzschner[3], J. Crane[4], Y. Li[4], and B. H. Menze[5]
[1]Technical University of Munich, Faculty of Computer Science, [2]Massachusetts Institute of Technology, [3]Johns Hopkins University, School of Medicine, [4]Univeristy of California San Francisco, Center for Intelligent Imaging, [5]Univeristy of Zurich, Department of Quantitative Biomedicine

Background: Deep learning research for MRS is a budding field with many promising applications. The greatest challenge currently is organizing enough data for deep learning (DL) models. Clinical datasets tend to be very small while DL projects can require 1E5 to 5E5 or more individual samples. Data augmentation eases this problem, but does not alleviate it entirely. A natural solution is to train on inexpensive synthetic data and apply the model to clinical data. This only works if the DL model learns a high-dimensional distribution of the training data that is similar enough to that of the clinical data. Currently this technique fails due to insufficient modeling, by standard physics models, of spectral components and the variation that arises during acquisition. Additionally, while there are many simulator frameworks for basis sets, there is almost nothing available for clinical data simulation meaning that they must be re-developed individually for every lab.

Method: Standard physics models modulate simulated basis lines, apply lorentzian line broadening and gaussian noise, and then a baseline offset. The results more closely approximate the output of a fitting model than clinical data. This framework proposes a much richer physics model, Eqn. 1, that reverse engineers the spectral fitting process. The first step is generating the basis set. This model uses MARSS[1] to simulate the basis sets. The MARSS framework generates very high quality basis lines that accurately capture the spatial 3D nature of voxels in their simulations. Next, the physics model improvements include modeling Voigt lineshapes, zero- and first-order phase offsets, a frequency shift for metabolites and a second one for lipid and macromolecular signals, a scaling factor to account for slight differences in spectral spacing, B0 inhomogeneities, and a rich generator for baseline and residual water components. This second generator uses a bounded pseudo-random walk followed by smoothing function with one set of parameters for the broad, smooth nature of baselines and another set for the more flexible, irregular residual water region.

$$F(\omega) = \mathcal{F}\left(e^{i(\phi_0+\phi_1(\nu_{ref}))}e^{-i\Delta ft} * \left(\sum_n^N M_n e^{-2i\pi\nu_n t}e^{-t/T_2^*}e^{-i\Phi} * e^{-d_n t}e^{-g_n t^2}e^{-it\Delta\omega} + baseline + \mathcal{H}(\mathcal{N}(0, \mathrm{snr}_0, l))\right)\right) \quad \textbf{Eqn. 1}$$

Examples:

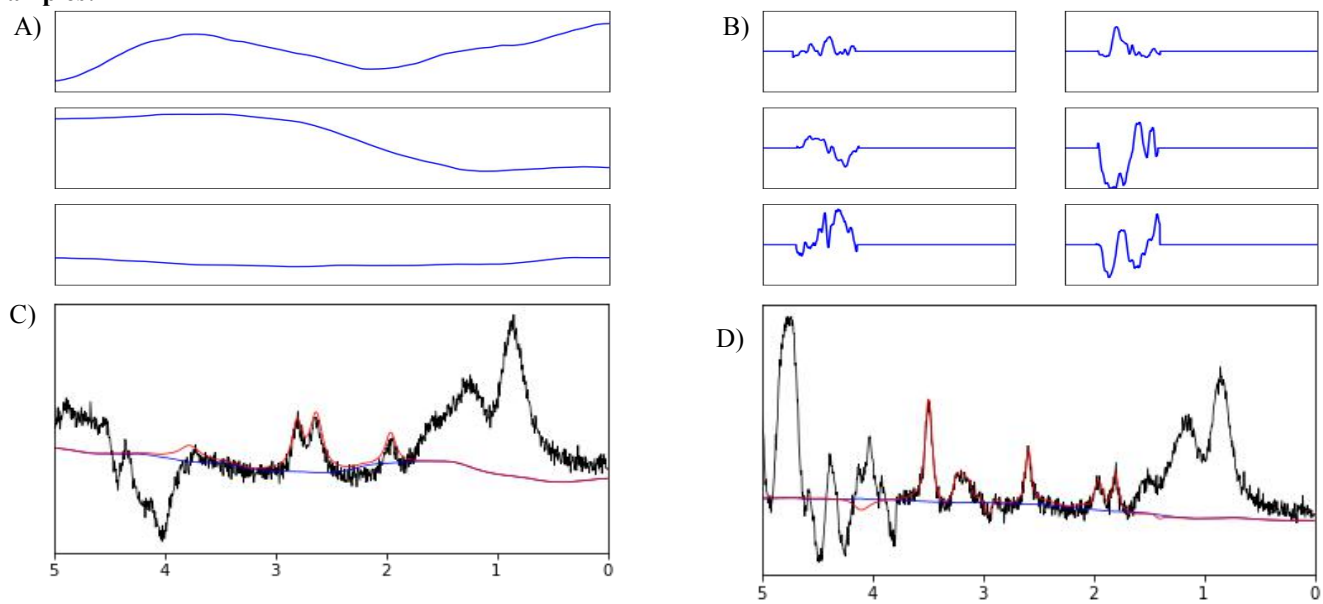

Figure 1: Generator samples. **A**) Baselines; **B**) Residual Water [6,0] ppm; **C**) Long Echo (3 lines); and **D**) Short Echo (7 lines)

Discussion: This work presents an open-source framework to simulate clinical MRSI data with the primary goal of facilitating deep learning research for MRS applications. This simulator uses very accurate basis lines and a rich, highly modular physics model. Parameters are sampled to optimize the performance of deep learning models. Due to the modularity of the implementation, it is very easy to turn artifacts on and off as well as generate spectra for different nuclei, different applications, and use different basis sets. The model is run in the terminal via a command line function and a .json config file. It is accompanied by a Jupyter notebook manual giving an overview of the model, how to use it, and how to customize the dataset. Similar to how accurate basis spectra improve spectral fitting, accurate training data improves the performance deep learning models. This framework will enable groups to explore deep learning research, even with limited clinical data.

References:

[1] Landheer, K, Swanberg, KM, Juchem, C. Magnetic resonance Spectrum simulator (MARSS), a novel software package for fast and computationally efficient basis set simulation. NMR in Biomedicine. 2021; 34:e4129. https://doi.org/10.1002/nbm.4129