Neural-Network-based Homogenized Model for Ferromagnetic Laminated Cores

F. Purnode[1], L. Denis[1], F. Henrotte[1], G. Louppe[1], and C. Geuzaine[1]

[1]Department of Electrical Engineering and Computer Science, University of Liège, Belgium
florent.purnode@uliege.be

The finite element modeling of hysteresis and eddy currents in ferromagnetic laminated cores is intricate and costly. In consequence, magnetic losses are often only evaluated a posteriori in electrical machine modeling. This leads however to potentially inaccurate results, as the effect of magnetic losses on the field computation is neglected. Deep learning offers an efficient solution to this problem. Two neural network architectures able to replicate hysteretic behaviors are discussed in this paper. Their fast and accurate inference makes them good candidates to serve as material laws in multiscale finite element simulations.

***Index Terms*—Ferromagnetic materials, Machine learning, Magnetic losses, Numerical simulation**

I. Introduction

MAGNETIC LOSSES result from the complex interplay between eddy currents, skin effect, saturation, and hysteresis. In previous work [1], we introduced an algebraic parametric material law to represent the homogenized behavior of the irreversible ferromagnetic material in macroscale Finite Element (FE) simulations. A Neural Network (NN) was trained to determine the best parameter values for this law in each finite element. The macroscale Jacobian matrix $\partial\vec{H}/\partial\vec{B}$ of the Newton-Raphson iteration scheme can then be calculated easily and exactly by an explicit differentiation of the algebraic material law. This approach allows accounting for magnetic losses during the simulation, and determining their spatial distribution. However, it is limited to the periodic case and by the *a priori* fixed form of the material law. These limitations can be removed by replacing the algebraic material law by a specifically trained NN, used to directly predict sequences of the magnetic field $\vec{H}$ from sequences of the magnetic flux density $\vec{B}$, or vice versa.

II. Neural network architectures

Recurrent Neural Networks (RNNs) handle data sequentially. They are able to capture temporal dependencies by maintaining a hidden state, updated dynamically as new inputs are processed. Various architectures have been introduced to improve their abilities to retain long-term dependencies, such as the Long Short-Term Memory (LSTM) [2]. The Gated Recurrent Unit (GRU) [3] was proposed as a simpler variant of LSTMs, shown to be more efficient [4]. GRUs have also shown promising results in other fields, e.g. in [5] within multiscale elasto-plastic simulations. In this work, we implement a GRU, constituted of two layers with a hidden size of 256. It can be trained to predict $\vec{H}$ sequences from $\vec{B}$ sequences, or vice versa. In both cases, the model processes the input sequences one time step at a time, progressively building the output sequences.

The transformer [6] has become the state-of-the-art for sequence-to-sequence modeling, and has been successfully applied to time series forecasting [7]. The classical transformer uses an encoder-decoder architecture, where the encoder creates a context-aware representation of the input, from which the decoder progressively builds the output sequence. In case the input and output have the same length, an encoder-only architecture can be used, which has already shown excellent performance in time series prediction [8]. Self-attention mechanisms allow them to capture input relationships and handle dependencies in parallel, so that all time-steps can be processed simultaneously. In this work, we implement an encoder-only transformer consisting of a stack of four standard transformer encoder layers, with an embedding dimension of 128, eight heads, and a feedforward hidden size of 2048.

III. Dataset Generation

The mesoscopic field distribution in laminated stacks can be accurately resolved by solving a one-dimensional (1D) FE magnetodynamic problem solving for eddy currents and accounting for hysteresis using the energy-based model [9]. Arbitrarily large datasets to train and test the NNs are built by first generating a set of artificial two-dimensional $\vec{H}$ sequences, and then associating each of them with the corresponding homogenized $\vec{B}$ solution of the 1D lamination model. The artificial $\vec{H} = [H_x(t_i), H_y(t_i)]$ sequences are generated as the sum of up to 50 harmonics, along both $x-$ and $y-$axes, evaluated at 1000 time steps t_i. Each $\vec{H}$ sequence is characterized by a fundamental frequency randomly chosen between 1 and 100 Hz, and a reference amplitude H_0 randomly chosen between 1 and 10^4 A/m. For each harmonic H^n including the fundamental, with n the harmonic number, the amplitudes are taken randomly between 0 and H_0/n. The spatial orientation and the phase shift between the x- and y-components are also chosen randomly. Finally, a DC bias of magnitude chosen randomly between $-H_0$ and H_0 for the $x-$ and $y-$axes is added to 50% of the $\vec{H}$ sequences.

A training, validation and test datasets with respectively $5\,10^5$, 10^4 and 10^4 ($\vec{H}$, $\vec{B}$) sequences are generated this way in 21 hours using an AMD EPYC 7763 CPU.

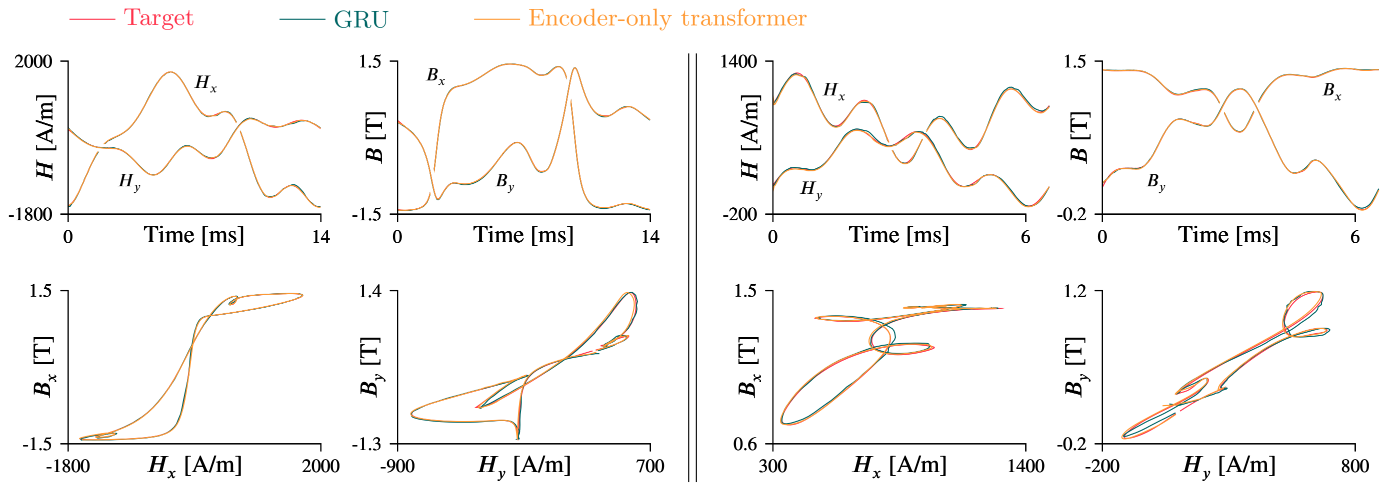

Fig. 1. Predictions made by the GRU and by the encoder-only transformer on two typical sequences from the test dataset. Both architectures provide a quasi-indistinguishable match with the target, regardless of whether $\vec{H}$ sequences are obtained from target $\vec{B}$ sequences, or vice versa.

TABLE I
AVERAGE MSE ON THE TEST DATASET AFTER TRAINING

Model	$\vec{H} \to \vec{B}$	$\vec{B} \to \vec{H}$
GRU	$7.7\,10^{-5}$	$4.9\,10^{-7}$
Encoder-only T.	$2.1\,10^{-4}$	$9.0\,10^{-7}$

IV. Training, Inference and Results

Both the GRU and the encoder-only transformer are trained by minimizing the component-wise Mean Squared Error (MSE) between the predicted and the target output sequences, both of them being first scaled down by the highest amplitude present in the training dataset. We used the Adam optimizer with a learning rate of $2\,10^{-4}$, and batches of 32 sequences. In order to train the NNs for non-periodic data, it is fed at training with sub-sequences of 500 consecutive time steps randomly extracted out of the full periodic sequences composing the training dataset. To speed up the inference of the encoder-only transformer, a reduced number of time steps can be used without significant loss in accuracy. Notably, at inference, overlapping subsets of 100 time steps are sequentially processed. The encoder-only transformer predicts the first 100 time steps using the first subset, time step 101 using the second subset, and so on.

Each training is performed on a single GPU node (NVIDIA A100 40GB) of the Lucia cluster, and runs for $2\ 10^{6}$ iterations. Training the GRU takes 12 hours, while training the encoder-only transformer takes 32 hours. On a laptop computer equipped with an M1 chip, predicting a single time step for 10^4 sequences takes approximately 0.2 s for the GRU and 3.4 s for the encoder-only transformer.

Once the training is achieved, the average MSE obtained over the test dataset (i.e., predictions obtained for input sequences that have not been used during the training) is reported in Table I. Figure 1 also demonstrates the capability of the GRU and of the encoder-only transformer to accurately reproduce the exact ferromagnetic cycles.

V. Conclusion and Perspectives

The GRU and the encoder-only transformer have proven able to accurately learn the homogenized response of ferromagnetic laminated cores. Using automatic differentiation, the exact Jacobian matrix $\partial\vec{H}/\partial\vec{B}$ is also easily obtained. In the full paper, the NNs will be directly used as material law in a macroscopic nonlinear magnetodynamic FE model of electrical machines with laminated cores.

Acknowledgment

The present research benefited from computational resources made available on Lucia, the Tier-1 supercomputer of the Walloon Region, infrastructure funded by the Walloon Region under the grant agreement n°1910247. The work of Louis Denis was supported by the F.R.S-FNRS.

References

[1] F. Purnode, F. Henrotte, G. Louppe, C. Geuzaine, "Neural network-based simulation of fields and losses in electrical machines with ferromagnetic laminated cores," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 37, no. 2, 2024.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, p. 1735–1780, 1997.

[3] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *Proceedings of the 2014 conference EMNLP*, pp. 1724–1734, 2014

[4] J. Chung *et al.*, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint*, arXiv:1412.3555, 2014.

[5] L. Wu, V. D. Nguyen, N. G. Kilingar, and L. Noels, "A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths," *Comp. Meth. Appl. Mech. Eng.*, vol. 369, pp. 113234, 2020.

[6] A. Vaswani *et al.*, "Attention is all you need," *NIPS*, vol. 30, 2017.

[7] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, L. Sun, "Transformers in time series: a survey," *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pp. 6778–6786, 2023.

[8] G. Zerveas *et al.*, "A transformer-based framework for multivariate time series representation learning," *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, pp. 2114–2124, 2021.

[9] F. Henrotte, S. Steentjes, K. Hameyer, C. Geuzaine, "Iron loss calculation in steel laminations at high frequencies," *IEEE transactions on magnetics*, vol. 50, no. 2, pp. 333—336, 2014.