

Acceleration of Multi-Scale LTS Magnet Simulations with Neural Network Surrogate Models

Louis Denis, Julien Dular, Vincent Nuttens, Mariusz Wozniak, Benoît Vanderheyden, and Christophe Geuzaine

Abstract—While the prediction of AC losses during transients is critical for designing large-scale low-temperature superconducting (LTS) magnets, brute-force finite-element (FE) simulation of their detailed geometry down to the length scale of the conductors is a computational challenge. Multi-scale methods, balancing between a coarse approximation of the fields at the scale of the magnet and a detailed description at the scale of the conductors, are promising approaches to reduce the computational load while keeping a sufficient accuracy.

In this work, we introduce a neural network approach to accelerate multi-scale magneto-thermal simulations of LTS magnets by replacing costly single-turn FE models with neural network surrogates. The neural network architecture is presented and discussed, together with an automated procedure for generating simulation data for its training. The resulting accelerated multi-scale model is used to simulate current ramp-up procedures for the IBA S2C2 magnet. The surrogate-based multi-scale model is compared with a conventional multi-scale model based on a composite wire-in-channel FE model. The surrogate model is shown to reproduce single-turn filament hysteresis, inter-filament coupling, and eddy losses, while the computational time of the multi-scale method is reduced by a factor of 800.

Index Terms—AC Losses, Finite-elements, Low-temperature superconductors, Multi-scale modeling, Neural networks

I. INTRODUCTION

DESIGNING and optimizing operation of large-scale Low-Temperature Superconducting (LTS) magnets for industrial applications requires fast and reliable AC loss simulations. Because the main loss mechanisms (filament hysteresis, inter-filament coupling and eddy losses) occur at the scale of individual conductors [1], [2], [3], an accurate estimation of the AC losses involves the calculation of the magneto-thermal response of the system from the scale of single conductors to the macroscopic scale of the magnet. While recent parallelization methods, such as domain decomposition [4] or GPU-optimized algorithms [5], can accelerate resolutions, full-scale finite-element (FE) simulations are currently out of reach for desktop computers.

In this context, computational homogenization [6] via multi-scale resolution techniques [7], [8] can be employed to reduce simulation times. The main idea is to couple a macroscopic FE model describing averaged physical fields with local

mesoscopic-scale models based on a detailed geometry, in order to capture, e.g., the relationship between local loss mechanisms and the large-scale temperature distribution in LTS magnets. In a recent work [9], we introduced a magneto-thermal multi-scale approach (MSA) that combined an homogenized 3D FE LTS magnet model with meso-scale models estimating AC local losses at the level of single filaments (SF). The MSA was shown to increase the accuracy of predicted peak AC losses with respect to state-of-the-art analytical approximations [9], [2].

In the present work, the MSA is first extended by modelling the actual composite superconductor geometry at the mesoscopic scale, replacing the SF model in [9] and enabling the representation of the different loss mechanisms with the FE method. While twisted superconductors are here modelled efficiently using the CATI method [10], their simulation in a full multi-scale resolution remains computationally expensive as it still requires high-performance computing facilities. In a second step, substantial acceleration of multi-scale simulations is achieved by replacing the mesoscopic-scale FE model with a neural network (NN) surrogate, an approach that showed promising results in computational mechanics [11]. Such NN surrogates represent an alternative to classical model order reduction techniques, such as energy-based models [12], [13] recently adapted to superconductors [14], [15].

The use of machine learning-based surrogates in applied superconductivity has grown rapidly in recent years [16] and offers interesting perspectives in modelling and optimizing superconducting systems [17], [18], [19], [20], [21]. Among others, physics-informed NNs are being investigated as alternatives to complete FE solvers [22], [23]. Here, we are specifically interested in the time-evolution of AC losses in composite superconductors. To this end, we opt for a NN surrogate based on Gated Recurrent Units (GRUs) [24] designed to represent history-dependent phenomena such as hysteresis in superconductors. Our approach differs from previous NN surrogates for AC losses in single conductors [25], [26], [27], [28], [29], mostly designed for periodic excitations and limited to averaged loss predictions. It also contrasts with [30], in which an advanced NN surrogate model for time-dependent AC loss prediction was proposed at the system level of superconducting motors. Moreover, our approach learns the direct mapping between time sequences of fields, avoiding the intermediate step of fitting coefficients in a parametric constitutive law [31], [14].

Building upon the simplified MSA presented in [9], where comparison with analytical approximations and experimental data can be found, the present paper focuses on acceler-

L. Denis is a research fellow of the Fonds de la Recherche Scientifique - FNRS. The work of J. Dular and M. Wozniak is partially supported by the CERN HFM program. (*Corresponding author: L. Denis*)

L. Denis, B. Vanderheyden, and C. Geuzaine are with the Department of Electrical Engineering and Computer Science, Institut Montefiore B28 in the University of Liege, 4000 Liege, Belgium (e-mail: louis.denis@uliege.be).

J. Dular and M. Wozniak are with CERN, 1211 Meyrin, Switzerland.

V. Nuttens is with Ion Beam Applications, 1348 Louvain-la-Neuve, Belgium.

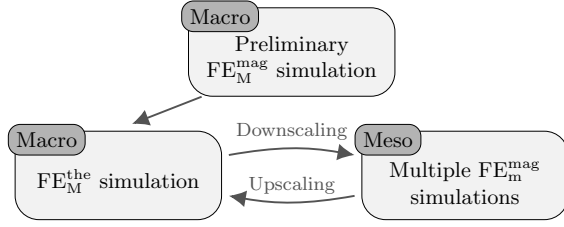


Fig. 1. Flowchart of the FE-based MSA, coupling the macroscopic thermal FE simulation of the LTS coil with mesoscopic magnetodynamic FE simulations.

ating the MSA using a GRU-based NN surrogate of the mesoscopic-scale FE model. It is structured as follows. Section II introduces the generalized FE-based MSA and its GRU-based counterpart. Section III discusses the NN surrogate, its architecture, its automated dataset generation and training procedure. Both the FE-based MSA and the novel GRU-based MSA are applied to ramp-up simulations of the existing IBA S2C2 synchrotron [32] in Section IV, demonstrating the significant speedup achieved with the NN surrogate.

II. MULTI-SCALE APPROACH

The basic scheme of the FE-based multi-scale approach is detailed in [9], and its generic workflow is shown in Fig. 1. Subscripts $\cdot_M$ and $\cdot_m$ denote macro- and mesoscopic-scale fields, respectively. In the LTS magnet, conductor turns are assumed to be connected in series and with negligible transverse dimensions w.r.t. the magnet radius. Given a prescribed current source $I_M(t)$, the macroscopic magnetic flux density $\mathbf{b}_M(\mathbf{x}_M, t)$ in the magnet is obtained from a preliminary magnetodynamic simulation FE_M^{mag} (e.g. with the a - v formulation as in [9]), based on the engineering current density imposed in the LTS coil.

The second step of the MSA involves using a macroscopic thermal simulation FE_M^{the} to solve the heat equation for the temperature field T_M within the cold mass of the magnet:

$$\rho_T c_p \partial_t T_M + \text{div} \cdot (-\kappa \cdot \mathbf{grad} T_M) = q_s, \quad (1)$$

where q_s is the heat source (in W/m^3), while ρ_T , c_p and κ are temperature-dependent thermal material properties that have been homogenized in the LTS coil (see [9, Appendix A] for their description). Notably, including the epoxy insulation in the homogenization increases the effective specific heat of the coil and leads to a strong anisotropy in its effective thermal conductivity. The combination of (1) with spatial boundary conditions (BCs) typically requires 3D simulations. Cooling is modelled with convective BCs, as described in [9].

In normal conductors (e.g. copper bars, aluminum coil formers), the heat source corresponds to conventional Joule losses $q_s = \rho_M \|\mathbf{j}_M\|^2$, with ρ_M the electrical resistivity and $\mathbf{j}_M$ the current density from the macroscopic simulation. In composite superconductors (i.e. in the LTS magnet), the loss density can be decomposed into its three main contributions:

$$q_s = q_{\text{fil}} + q_{\text{coupl}} + q_{\text{eddy}}, \quad (2)$$

with q_{fil} , q_{coupl} and q_{eddy} denoting filament hysteresis, inter-filament coupling and eddy losses [10], respectively. While some of these components can be estimated analytically

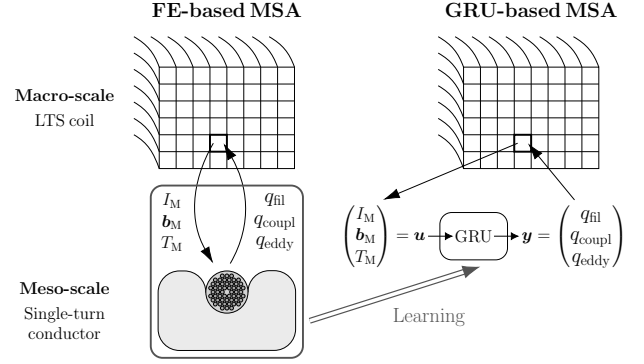


Fig. 2. FE-based MSA: fields are downscaled from the macroscopic LTS coil model to the mesoscopic conductor model and AC losses are upscaled back (left). GRU-based MSA: the mesoscopic model FE_m^{mag} is replaced with a GRU-based NN surrogate (right).

[1], [2], [33], [9], the FE-based MSA relies on multiple mesoscopic-scale models FE_m^{mag} of a single conductor turn for reliable loss prediction. In particular, the MSA in [9] is generalized here by replacing its meso-scale single-filament model with a more comprehensive model of the composite superconductor, which enables the accurate description of inter-filament coupling currents. The mesoscopic FE model determines the averaged loss components in (2) by solving for the local magnetic field $\mathbf{h}_m$ and its corresponding current density $\mathbf{j}_m$, given transport current $I_M(t)$ and applied external flux density $\mathbf{b}_M(t)$ sequences from the macroscopic scale. Due to the high thermal conductivity of copper ensuring rapid thermal equilibrium, temperature is not solved at the mesoscopic scale and is assumed uniform $T_m = T_M$ within each turn. Still, mesoscopic losses are dependent on T_M via the local critical current density $j_c(\mathbf{b}_m, T_M)$.

Consequently, the FE-based MSA relies on a two-way coupling between the macroscopic thermal simulation FE_M^{the} and multiple mesoscopic loss simulations FE_m^{mag} . This is illustrated in Fig. 2 for a wire-in-channel (WIC) conductor, as used in the IBA S2C2 magnet [9]. Rather than associating one mesoscopic model FE_m^{mag} with each macroscopic integration point, as in the FE^2 method [7], the coil cross-section is subdivided into zones and the FE-based MSA in Fig. 2 assigns one mesoscopic model to each macroscopic zone. This improves efficiency as it greatly reduces the number of mesoscopic simulations required [9]. In this case, the inputs $\mathbf{u} = (I_M, \mathbf{b}_M, T_M)^T$ of each single-turn model correspond to field sequences spatially averaged over the corresponding zone.

From the macroscopic perspective, the mesoscopic FE model acts as a black box returning the loss density contributions, that can be expressed at a given time $t^n = n\Delta t_M$ (with Δt_M the time step of FE_M^{the}) as

$$\mathbf{y}(t^n) = \mathbf{y}(\underbrace{\mathbf{u}(t \leq t^n)}_{\text{history}}, \text{conductor properties, geometry}), \quad (3)$$

where $\mathbf{y} = (q_{\text{fil}}, q_{\text{coupl}}, q_{\text{eddy}})^T$. Since the conductor features are fixed for a given magnet, (3) shows that a history-dependent surrogate could replace the single-turn FE model for faster loss prediction. A GRU-based NN surrogate is considered for this purpose and its integration in the MSA is represented

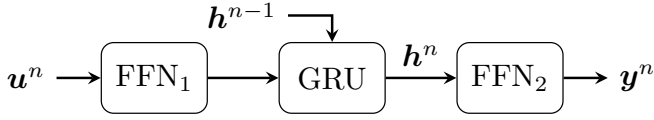


Fig. 3. Structure of the GRU-based RNN. Two feedforward networks FFNs are placed in front and after the GRU, respectively.

in Fig. 2. The GRU-based MSA is thus similar to the FE-based MSA, with the mesoscopic FE model (bottom left in Fig. 1) replaced by the surrogate. Both FE-based and GRU-based MSAs are implemented in GetDP [34] with a Python interface for communication between scales.

Limitations of the FE-based and GRU-based MSA include:

- The subdivision into zones may introduce spatial aliasing in the macroscopic thermal simulation. This can be mitigated by conducting a dedicated convergence study as performed in [9, Section VII.C], as 36 zones already showed satisfactory convergence of AC losses during the S2C2 ramp-up. However, the simulation of local defects in the LTS conductor, possibly leading to quench phenomena, remains out-of-scope of the proposed MSA.
- Similarly, the variability of material and geometrical parameters along conductor length is neglected: critical current scaling law, copper matrix RRR, superconductor filling ratio and twist pitch are considered fixed.

III. NEURAL NETWORK SURROGATE

A. Neural Network Architecture

GRUs are designed to capture temporal dependencies by treating data sequentially, with the history at time t_n encoded in the hidden state vector h^n [24]. The simplified GRU-based surrogate architecture, adapted from [11], is shown in Fig. 3. At time t_n , the input u^n is first processed by a feedforward network (FFN₁), whose output is passed to the GRU along with its previous hidden state h^{n-1} . The GRU (one hidden layer of size 100, initialized to zero) produces the updated h^n , which is then passed through a second feedforward network (FFN₂) to generate the final predictions y^n . FFN₁ and FFN₂ each have one hidden layer of size 32, with LeakyReLU and Sigmoid activation functions, respectively. By design, the three loss components are predicted as separate output channels, while they share a common history via the GRU hidden state.

Enriching the input vector u with additional features increased accuracy. The optimal setting includes eight scaled macroscopic features, denoted $\tilde{u}$,

$$u^n = \left(\tilde{I}_M^n, \tilde{b}_{M,r}^n, \tilde{b}_{M,z}^n, \tilde{T}_M^{n-1}, \tilde{t}^n, \partial_t \tilde{I}_M^n, \partial_t \tilde{b}_{M,r}^n, \partial_t \tilde{b}_{M,z}^n \right)^T, \quad (4)$$

where $b_{M,r}$ and $b_{M,z}$ are the radial and vertical components of the macroscopic flux density. In particular, passing the macroscopic temperature at the previous time step T_M^{n-1} was found to increase the efficiency of the GRU-based MSA.

For a current ramp-up of the LTS magnet, input features $u_i \in [u_i^{\min}, u_i^{\max}]$ are linearly scaled to $\tilde{u}_i \in [-1, 1]$ for $b_{M,r}, b_{M,z}, \partial_t b_{M,r}, \partial_t b_{M,z}$, and to $\tilde{u}_i \in [0, 1]$ for $I_M, T_M, t, \partial_t I_M$, where $u_i^{\min}$ and $u_i^{\max}$ are the minimal and maximal values in the dataset. To maintain accuracy across

the full range of losses, each output feature y_i is scaled as $\tilde{y}_i = (y_i / y_i^{\max})^{1/3}$. Combined with the Sigmoid final activation of the NN, this scaling provides a physical safeguard: predictions cannot exceed the loss values observed during training.

B. Automated Training Procedure

As the input space is high-dimensional, a dedicated three-step automated training procedure is proposed in Fig. 4, such that a surrogate model can be trained for each particular LTS magnet to be studied. The first step requires human input. It consists in describing the macroscopic magnet model (geometry, material properties and FE solver), and the corresponding mesoscopic single conductor turn model. A sampling method (e.g. subdividing the coil cross-section in 100 zones) must be defined to retrieve representative field sequences.

The second step focuses on dataset generation. During magnet ramp-up, the control variable is the imposed current I_M . Accordingly, 80 different current sequences are generated as random walks, with the constraint that current can only increase and must reach its nominal value after ramp-up: $I_M(T_{\text{up}}) = I_{\text{nom}}$. One example sequence is shown in Fig. 4. The total ramp-up time T_{up} is varied to cover different time scales. For the S2C2 magnet, it is drawn randomly in $[3600, 36000]$ (in s). Dataset generation based on random-walk loading has been widely used for data-driven or multi-scale modelling of history-dependent nonlinear mechanical systems [11], [36], [37], [38], [39]. Given the similarity of the studied configuration, the same strategy is adopted here to ensure broad coverage of possible input sequences and to avoid bias towards specific ramping patterns, improving the surrogate's interpolation capacity. For each current profile, a FE-based MSA simulation is then run (cf. left part in Fig 2). Each run produces a set of representative u - y sequences (e.g. 1 per zone, thus 100 per run, as in Section III), diversified by the local magnetic field and temperature variations, leading to distinguishable losses within the LTS coil [9, Fig. 19]. Together, these sequences form the dataset.

The final step is the training of the NN surrogate. To this end, the generated dataset is randomly split into training (85%) and validation (15%) sets, with u - y sequences from the same FE-based MSA simulation assigned to the same set. An independent test set based on physical current ramps is evaluated in Section IV. Training is performed with PyTorch, the Adam optimizer and the mean squared error (MSE) loss function. The optimal hyperparameters of the surrogate, cf. Section III-A, were determined through a grid search [35] over the following ranges: number of hidden layers (1-5), hidden state sizes (10-250), and FFN layer sizes (8-256).

IV. APPLICATION, RESULTS AND DISCUSSION

The existing S2C2 LTS magnet [32] is used as a case study to compare the GRU-based MSA with the previously validated but time consuming FE-based MSA [9]. The 3D macroscopic thermal model and the preliminary axisymmetric magnetic model, including the ferromagnetic yoke, are detailed in [9, Sec. VII.a]. Notably, the LTS coil is conduction-cooled with

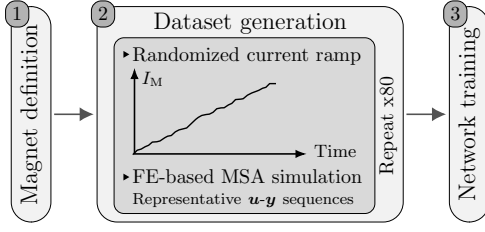


Fig. 4. Three steps of the offline automated training procedure.

TABLE I
MAIN PARAMETERS OF THE WIC CONDUCTOR IN THE S2C2 COIL.

Description	Value
Critical current density $j_c(b, T)$ -fit	[41], [42], [43]
$j_c(5 \text{ T}, 4.2 \text{ K})$	2783 A/mm ² [9]
Power-law n -index	50
NbTi Filament diameter	156 μm
Superconductor filling ratio in WIC	0.148
Central strand twist pitch length	0.1 m
Copper (strand matrix and channel) RRR	80
Nominal current I_{nom}	656 A
Copper channel dimensions	$\approx 4 \times 2.5 \text{ mm}^2$

four cryocoolers. Their position and cooling capacity is detailed in [9, Sec. VII.a]. In the FE-based MSA, the WIC-geometry is assumed periodic and is solved using the CATI method [10] implemented in FiQuS [40], which replaces the full 3D twisted conductor model with two coupled 2D finite-element problems. The WIC model is similar to the one described in [10, Sec. 5.2] and sketched in Fig. 2, with its main parameters adapted to the epoxy-impregnated S2C2 conductor and gathered in Table I.

The NN surrogate is trained following the procedure described in Section III-B. For dataset generation, 80 FE-based MSA simulations are performed with the LTS coil subdivided into 100 zones, providing 100 representative u - y sequences per simulation and a total of 8000 sequences. Each sequence consists of 500 time steps, resulting in 4 million samples. Input and output features are scaled as described in Section III.A. While a more detailed description of the generated dataset is left out-of-scope, the corresponding ranges are reported in Table II. Training and validation MSE curves for different training set sizes, averaged over five random training/validation dataset splits, are shown in Fig. 5. These results demonstrate that insufficient training data leads to overfitting, as indicated by early saturation of the validation MSE. No significant reduction of the validation MSE is observed beyond 5200 sequences (65% of the dataset), which suggests that the considered dataset size is appropriate.

Once trained, the GRU-based MSA associates one NN surrogate with each zone, considering the same subdivision into 100 zones as the FE-based MSA. As shown in Fig. 6, the GRU-based MSA reproduces FE-based MSA loss predictions when tested on practical engineering ramp-ups of different durations, including the practical current IBA S2C2 ramp-up profile. The current profiles in Fig. 6 are not part of the training dataset, highlighting the good interpolation capacity of the NN surrogate. The surrogate extracts the relevant information from the 49k degrees of freedom of the FE WIC model into

TABLE II
INPUT (LEFT) AND OUTPUT (RIGHT) FEATURE RANGES BEFORE SCALING.

Variable	Range	Variable	Range
I_M (A)	[0, 656]	q_{fil} (W/m ³)	[0, 323]
$b_{M,r}$ (T)	[-3.35, 3.35]	q_{coupl} (W/m ³)	[0, 128]
$b_{M,z}$ (T)	[-5.25, 5.25]	q_{eddy} (W/m ³)	[0, 0.167]
T_M (K)	[3.64, 8.07]		
t (h)	[0, 10]		
$\partial_t I_M$ (A/s)	[0, 0.363]		
$\partial_t b_{M,r}$ (mT/s)	[-4.86, 4.86]		
$\partial_t b_{M,z}$ (mT/s)	[-4.53, 4.53]		

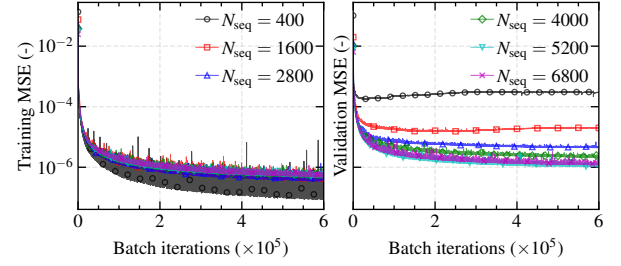


Fig. 5. Training (left) and validation (right) MSE curves for different training-set sizes N_{seq} , against number of batch (fixed size: 64) iterations. Results have been averaged over five independent runs to account for stochastic variability.

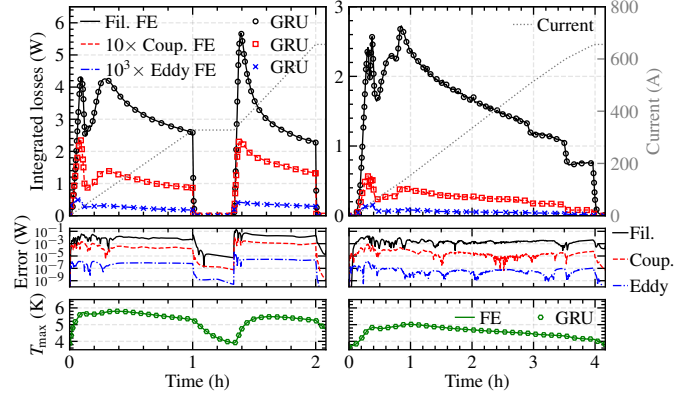


Fig. 6. Volume-integrated filament hysteresis, coupling and eddy losses in the LTS coil, computed with FE-based and GRU-based MSAs (top), together with the corresponding absolute error (center) and maximal coil temperature (bottom). Example 2 h ramp-up composed of two linear ramps with an intermediate pause (left). Practical 4 h ramp-up profile of the S2C2 [9] (right).

100 hidden states, without explicitly resolving the detailed current density j_m distribution. Notably, loss contributions of different physical nature, see (2), and spanning several orders of magnitude are accurately captured. This is quantified by computing the coefficient of determination R^2 for each loss contribution $Q_{\text{GRU}}(t)$ predicted by the GRU-based MSA w.r.t. the FE-based reference $Q_{\text{FE}}(t)$ (mean value: $\bar{Q}_{\text{FE}}$), defined as

$$R^2 = 1 - \frac{\int_0^{T_{\text{up}}} (Q_{\text{FE}}(t) - Q_{\text{GRU}}(t))^2 dt}{\int_0^{T_{\text{up}}} (Q_{\text{FE}}(t) - \bar{Q}_{\text{FE}})^2 dt}. \quad (5)$$

The $1 - R^2$ values gathered in Table III confirm the excellent agreement between the GRU-based MSA and its FE-based counterpart, and demonstrate the ability of the NN surrogate to produce reliable predictions in realistic operation conditions. This agreement is further assessed through the pointwise relative error ε_{rel} of local GRU predictions against the mesoscopic

TABLE III

1 — R^2 COEFFICIENT OF VOLUME-INTEGRATED LOSSES, WITH MEDIAN AND 95TH-PERCENTILE RELATIVE ERRORS OF LOCAL PREDICTIONS, FOR THE TWO ILLUSTRATED CURRENT RAMP-UP PROFILES.

1 — R^2	Filament	Coupling	Eddy
Example 2 h ramp	2.27×10^{-4}	2.93×10^{-4}	4.71×10^{-4}
Practical 4 h ramp	2.57×10^{-5}	7.37×10^{-5}	3.68×10^{-5}
median(ϵ_{rel})	0.45%	0.48%	0.015%
$P_{95}(\epsilon_{\text{rel}})$	4.1%	2.2%	0.25%

FE reference, summarized by the median and 95th-percentile (P_{95}) values reported in Table III. These low values indicate the absence of large localized prediction errors. As shown in Fig. 6, the coil maximum temperature evolution is also well captured, with corresponding 1 — R^2 values of 1.23×10^{-4} and 1.6×10^{-5} for the 2 h and 4 h ramp-ups, respectively.

The maximal temperature increase predicted by the WIC-FE-based MSA is $\Delta T_{\text{max}} = 1.331$ K during the 4 h ramp-up. This high value is driven by the large filament size and agrees with the value obtained with the simplified single-filament FE-based MSA in [9] ($\Delta T_{\text{max}} = 1.336$ K), showing that inter-filament coupling losses are negligible during normal operation. This strong agreement also indirectly verifies the present WIC-FE-based MSA, since the SF-FE-based MSA was experimentally validated in [9, Sec. VII.E]. In the same reference, an analytical power-law model for ramping hysteresis losses was proposed and shown to underestimate the peak losses (and resulting temperature rise) during the 4 h ramp-up, while the FE-based MSA captured additional effects such as locally rotating fields more accurately [9, Sec. VII.F]. Since the NN surrogate is trained to reproduce the mesoscopic FE results, it therefore inherits this improved accuracy over simpler analytical models.

The computational performance of the different approaches for the 4 h S2C2 ramp-up simulation is summarized in Table IV. A single FE-based MSA simulation requires 1500 CPU-h. In contrast, after training, the GRU-based MSA simulation (with one surrogate per macroscopic zone, or 100 in total) runs in less than 2 h on a single CPU, achieving an 800-fold speedup. Its memory requirement (in terms of the maximum resident set size RSS) is also significantly lower. The remarkable efficiency of the GRU-based MSA enables the evaluation of one distinct mesoscopic-scale surrogate at each of the 41k integration points in the FE_M^{the} model, leading to a FE^2 -type method. This is a significant step towards full-scale simulations of LTS magnets, since the FE-based MSA with as many mesoscopic models would require $41\text{k} \times 15 \approx 600\text{k}$ CPU-h. The 1 — R^2 coefficients between the full 41k-points GRU-based MSA and the 100-zone GRU-based MSA remain below 5×10^{-4} , confirming that the zone-discretization introduced in [9] does not lead to significant aliasing errors at the macroscopic FE level.

Table IV also shows the significant cost of the training procedure. While the training of the NN itself takes less than one hour on a single GPU, the 80 FE-based MSA simulations required for dataset generation correspond to about 120k CPU-h. This cost cannot be overlooked, and the proposed training

TABLE IV
COMPUTATIONAL PERFORMANCE FIGURES.

MSA	FE-based	GRU-based	GRU-based
Nb. mesoscopic models	100	100	41k
Wall time (s)	54055	6056	7278
Nb. CPUs	100	1	1
CPU hours	1500	1.7	2.02
Maximum RSS (GB)	155	0.664	4.1
Offline dataset generation	-	$80 \times 1500 = 120\text{k CPU-h}$	

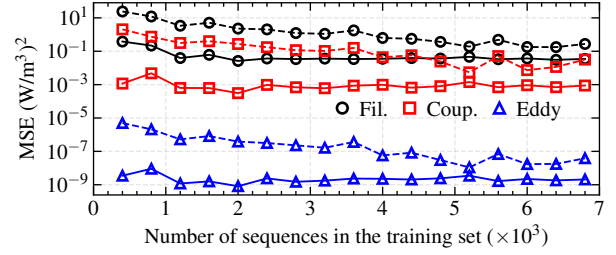


Fig. 7. Training (solid lines) and validation (dashed lines) MSE of the best model (selected at minimal validation MSE) for each loss contribution, as a function of training set size. Results have been averaged over five independent runs to account for stochastic variability.

strategy is justified only when a large number of simulations of a given magnet are needed. Among other applications, this approach is suited to the optimization of the magnet ramp-up profile, coil layout, cable arrangement, and magnetic circuit, given a fixed cable design.

The importance of generating sufficiently many input-output sequences is further illustrated in Fig. 7, where each physical loss component was scaled back to its physical range. With a small training set, the NN surrogate overfits: the MSE is significantly lower when evaluated on the training set. As the training size increases, the gap between training and validation MSE decreases, until enough data is provided for training. Notably, the validation MSE for all loss contributions decreases by two orders of magnitude when increasing the training set size from 400 up to 5200 sequences. Beyond this size, no significant improvement is observed, which is consistent with the results provided in Fig. 5.

V. CONCLUSION

Multi-scale magneto-thermal FE simulations of LTS magnets have been accelerated by replacing costly mesoscopic FE models with a neural network surrogate. An automated framework for training the surrogate has been developed and applied to the 3D simulation of the IBA S2C2 magnet describing a current ramp-up. Once trained on detailed simulation data, the surrogate delivers an 800-fold speedup of the 3D multi-scale model while showing excellent agreement in predicting AC loss contributions of different physical origins. This approach provides a reliable tool for optimizing the operation of existing magnets, the design of next-generation magnets, and real-time simulation of large-scale LTS systems.

Future work could introduce cable material and geometric parameters as inputs to the NN surrogate. This would allow to model their variability along the conductor length, or to evaluate different cable designs without retraining the surrogate.

ACKNOWLEDGMENTS

The authors thank F. Henrotte, F. Purnode and D. Colignon (University of Liège) for valuable discussions on the implementation of the methods. 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.

REFERENCES

- [1] A. Campbell, "A general treatment of losses in multifilamentary superconductors," *Cryogenics*, vol. 22, no. 1, pp. 3-16, 1982.
- [2] M. N. Wilson, "Superconducting Magnets," Oxford, U.K.: Clarendon, 1983.
- [3] J. Dular, A. Verweij and M. Wozniak, "A Finite Element a-h-Formulation for the Reduced Order Hysteretic Magnetization Model for Composite Superconductors," *IEEE Trans. Appl. Supercond.*, vol. 35, no. 5, Aug. 2025, Art no. 8200205.
- [4] N. Riva *et al.*, "H- ϕ Formulation in Sparselizard Combined With Domain Decomposition Methods for Modeling Superconducting Tapes, Stacks, and Twisted Wires," *IEEE Trans. Appl. Supercond.*, vol. 33, no. 5, Art. no. 4900405, 2023.
- [5] C. Xue *et al.*, "Holistic numerical simulation of a quenching process on a real-size multifilamentary superconducting coil," *Nature Communications*, vol. 15, no. 1, p. 10454, 2024.
- [6] M. El Feddi, Z. Ren, A. Razek, and A. Bossavit, "Homogenization technique for Maxwell equations in periodic structures," *IEEE Trans. Magn.*, vol. 33, no. 2, pp. 1382-1385, Mar. 1997.
- [7] F. Feyel, "A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua," *Comput. Methods Appl. Mech. Eng.*, vol. 192, no. 28-30, pp. 3233-3244, 2003.
- [8] A. Abdulle, "The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs," *GAKUTO Int. Ser. Math. Sci. Appl.*, vol. 31, p. 135-184, 2009.
- [9] L. Denis, V. Nottens, B. Vanderheyden, and C. Geuzaine, "AC Loss Computation in Large-Scale Low-Temperature Superconducting Magnets: Multiscale and Semianalytical Procedures," *IEEE Trans. Appl. Supercond.*, vol. 35, no. 2, 2025, Art no. 5900319.
- [10] J. Dular *et al.*, "Coupled axial and transverse currents method for finite element modelling of periodic superconductors," *Supercond. Sci. Technol.*, vol. 37, 2024.
- [11] L. Wu, N. G. Kilinger, 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," *Comput. Meth. Appl. Mech. Eng.*, vol. 369, 2020.
- [12] A. Bergqvist, "Magnetic vector hysteresis model with dry friction-like pinning," *Physica B: Condensed Matter*, vol. 233, no. 4, pp. 342-347, 1997.
- [13] F. Henrotte, A. Nicolet, and K. Hameyer, "An energy-based vector hysteresis model for ferromagnetic materials," *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 25, no. 1, pp. 71-80, 2006.
- [14] J. Dular, A. Verweij, and M. Wozniak, "Reduced Order Hysteretic Magnetization Model for Composite Superconductors," *Supercond. Sci. Technol.*, vol. 38, no. 3, p. 035017, 2025.
- [15] A. Glock, J. Dular, A. Verweij, and M. Wozniak, "Reduced Order Hysteretic Flux Model for Transport Current Homogenization in Composite Superconductors," *IEEE Trans. Magn.*, 2025.
- [16] G. Tomassetti, "Direct and surrogate optimization in applied superconductivity: state of the art, perspectives and challenges," *Supercond. Sci. Technol.*, vol. 38, no. 7, p. 073001, 2025.
- [17] M. Yazdani-Asrami *et al.*, "Roadmap on artificial intelligence and big data techniques for superconductivity," *Supercond. Sci. Technol.*, vol. 36, no. 4, p. 043501, 2023.
- [18] A. Sadeghi, W. Song, and M. Yazdani-Asrami, "Intelligent surrogate model of a high-temperature superconducting cable for reliable energy delivery: short-circuit fault performance," *Supercond. Sci. Technol.*, vol. 37, no. 8, p. 085025, 2024.
- [19] A. Arsenault and M. Calvi, "Feed-forward neural network surrogate model for rapid simulations of a high-temperature superconducting bulk undulator," *Supercond. Sci. Technol.*, vol. 38, no. 8, p. 085018, 2025.
- [20] L. Giannini and D. P. Boso, "Machine Learning-Based Surrogate Modeling for Enhancing the Magnet Design of Future Tokamaks," *IEEE Trans. Appl. Supercond.*, vol. 35, no. 5, pp. 1-5, Art no. 4202805, Aug. 2025.
- [21] D. Wu, D. Sotnikov, M. Haajari, and T. Salmi, "Minimizing Conductor Consumption in High-Field HTS Solenoid Design Using Adaptive ANN-Based Optimization Algorithm," *IEEE Trans. Appl. Supercond.*, vol. 35, no. 5, pp. 1-5, Art no. 4603505, Aug. 2025.
- [22] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686-707, 2019.
- [23] S. Alipour Bonab, W. Song, and M. Yazdani-Asrami, "Physics-informed neural network model for transient thermal analysis of superconductors," *Supercond. Sci. Technol.*, vol. 38, no. 8, p. 08LT01, 2025.
- [24] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv:1406.1078*, 2014.
- [25] J. Leclerc, L. M. Hell, C. Lorin, and P. J. Masson, "Artificial neural networks for AC losses prediction in superconducting round filaments," *Supercond. Sci. Technol.*, vol. 29, no. 6, p. 065008, 2016.
- [26] M. Yazdani-Asrami *et al.*, "Prediction of nonsinusoidal AC loss of superconducting tapes using artificial intelligence-based models," *IEEE access*, vol. 8, pp. 207287-207297, 2020.
- [27] L. Zhou *et al.*, "AC Loss Calculation of High Temperature Superconducting Coils Based on a Surrogate Model," *IEEE Trans. Appl. Supercond.*, vol. 35, no. 5, 2025, Art no. 5902205.
- [28] Y. Huang, M. Yazdani-Asrami, and W. Song, "A Deep Learning-Based Ultra-Fast Surrogate Model for AC Loss Estimation in Superconductors Using COMSOL," *IEEE access*, vol. 13, pp. 153845-153854, 2025.
- [29] S. Alipour Bonab, W. Song, Z. Jiang, and M. Yazdani-Asrami, "Physics-constrained deep learning framework for accurate AC loss prediction of high-temperature superconductors: Impact of data scarcity and extrapolation beyond the range of experimental measurements," *Supercond. Sci. Technol.*, vol. 38, no. 12, p. 12LT01, 2025.
- [30] S. Alipour Bonab, F. Berg, W. Song, A. Colle, and M. Yazdani-Asrami, "Advanced deep-learning model for temporal-dependent prediction of dynamic behavior of AC losses in superconducting propulsion motors for hydrogen-powered cryo-electric aircraft," *Commun. Eng.*, 2025.
- [31] F. Purnode, F. Henrotte, G. Louppe, and 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, p. e3226, 2024.
- [32] W. Kleeven *et al.*, "The IBA superconducting synchrotron project S2C2," in *Proc. Cyclotrons*, 2013, pp. 115-119.
- [33] W. J. Carr Jr., *AC loss and macroscopic theory of superconductors*. 2nd ed. London, UK: Taylor and Francis, 2001.
- [34] P. Dular, C. Geuzaine, F. Henrotte, and W. Legros, "A general environment for the treatment of discrete problems and its application to the finite element method," *IEEE Trans. Magn.*, vol. 34, no. 5, pp. 3395-3398, 1998.
- [35] C. M. Bishop, *Pattern Recognition and Machine Learning*. 1st ed. New York, NY, USA: Springer, 2006.
- [36] Y. He and S. J. Semnani, "Machine learning based modeling of path-dependent materials for finite element analysis," *Comput. and Geotechnics*, vol. 156, p. 105254, 2023.
- [37] E. Ghane, M. Fagerström, and M. Mirkhalaf, "Recurrent neural networks and transfer learning for predicting elasto-plasticity in woven composites," *Eur. J. Mech.-A/Solids*, vol. 107, p. 105378, 2024.
- [38] J. N. Heidenreich and D. Mohr, "Extended minimal state cells (EMSC): Self-consistent recurrent neural networks for rate- and temperature dependent plasticity," *Int. J. Plasticity*, vol. 188, p. 104305, 2025.
- [39] M. O. Bozkurt and V. L. Tagarielli, "Data-driven homogenisation of viscoelastic porous elastomers: Feedforward versus knowledge-based neural networks," *Int. J. Mech. Sci.*, vol. 286, p. 109824, 2025.
- [40] A. Vitranò *et al.*, "An Open-Source Finite Element Quench Simulation Tool for Superconducting Magnets," *IEEE Trans. Appl. Supercond.*, vol. 33, no. 5, 2023, Art no. 4702006.
- [41] L. Bottura, "A practical fit for the critical surface of NbTi," *IEEE Trans. Appl. Supercond.*, vol. 10, no. 1, pp. 1054-1057, 2000.
- [42] C. Spencer, P. Sanger, and M. Young, "The temperature and magnetic field dependence of superconducting critical current densities of multifilamentary Nb3Sn and NbTi composite wires," *IEEE Trans. Magn.*, vol. 15, no. 1, pp. 76-79, 1979.
- [43] M. S. Lubell, "Empirical scaling formulas for critical current and critical field for commercial NbTi," *IEEE Trans. Magn.*, vol. 19, no. 3, pp. 754-757, 1983.