

---

# Hybridizing Physics and Neural ODEs for Predicting Plasma Inductance Dynamics in Tokamak Fusion Reactors

---

Allen M. Wang<sup>1,2</sup>, Darren T. Garnier<sup>1</sup>, and Cristina Rea<sup>1</sup>

<sup>1</sup>MIT Plasma Science and Fusion Center, Cambridge, MA, USA

<sup>2</sup>MIT Laboratory for Information and Decision Systems, Cambridge, MA, USA

## Abstract

While fusion reactors known as tokamaks hold promise as a firm energy source, advances in plasma control, and handling of events where control of plasmas is lost, are needed for them to be economical. A significant bottleneck towards applying more advanced control algorithms is the need for better plasma simulation, where both physics-based and data-driven approaches currently fall short. The former is bottle-necked by both computational cost and the difficulty of modelling plasmas, and the latter is bottle-necked by the relative paucity of data. To address this issue, this work applies the neural ordinary differential equations (ODE) framework to the problem of predicting a subset of plasma dynamics, namely the coupled plasma current and internal inductance dynamics. As the neural ODE framework allows for the natural inclusion of physics-based inductive biases, we train both physics-based and neural network models on data from the Alcator C-Mod fusion reactor and find that a model that combines physics-based equations with a neural ODE performs better than both existing physics-motivated ODEs and a pure neural ODE model.

## 1 Introduction

Recent advances in nuclear fusion have offered hope that it may arrive in time to help combat climate change by serving as a firm energy source to complement intermittent renewables. Within the fusion landscape, the tokamak is often considered the leading candidate for a fusion pilot plant, promising long-duration pulses with an existing physics basis offering a clear path to high energy gain Wurzel and Hsu [2022], Sorbom et al. [2015]. However, tokamaks are prone to sudden, highly energetic losses of plasma confinement induced by plasma instabilities. These events are known as “plasma disruptions”. While plasma disruptions don’t pose a threat to public safety, they do pose a significant threat to the economic viability of tokamaks Maris et al. [2023]. This threat motivates the development of reliable high-performance real-time plasma control systems that can predict the onset of instabilities and “soft-land” the plasma by rapidly de-energizing it to a manageable state.

One of the major challenges to realizing more reliable and high-performance tokamak plasma control is the difficulty of modelling plasma dynamics. While recent work successfully demonstrated the application of deep reinforcement learning to magnetic control of tokamak plasmas Degraeve et al. [2022], the physics relevant to magnetic control, i.e. Ideal Magnetohydrodynamics (MHD), is the most well-simulated, and, furthermore, classical control techniques are already effective at tackling it.

To successfully control other aspects of plasma dynamics, advances in modelling are needed. In this regard, both purely physics-based and data-driven approaches currently fall short. On the purely physics-based side, the highest fidelity plasma simulations available today require millions of CPU hours to arrive at a steady state solution, making them unusable for control development, and even

then require assumptions on boundary conditions such as the pressure gradient at the edge of the plasma Rodriguez-Fernandez et al. [2022]. This gap has motivated work on purely data-driven dynamics modelling with residual and recurrent neural networks, however, the relative paucity of data, distributional drift due to changes in hardware, and irregular measurement time-bases are major challenges Char et al. [2023], Abbate et al. [2021].

To tackle this problem, the neural ordinary differential equations (Neural ODEs) framework offers a promising path to improving the sample efficiency and robustness of data-driven prediction of plasma dynamics via the introduction of physics-based inductive biases. In this work, we explore the application of the neural ODE framework to a subset of the plasma dynamics problem. Namely, we model the coupled dynamics of plasma current,  $I_p$ , and plasma internal inductance  $L_i$  to experimental data from the Alcator C-Mod fusion reactor. These coupled dynamics are of particular interest for “soft landing” as decreasing  $I_p$  is one of the primary objectives of soft-landing, but doing so generally increases  $L_i$ , which is correlated with reduced plasma stability. Traditionally, partial differential equations (PDEs) would need to be evolved forward in time to predict the coupled dynamics of these two variables, but doing so is computationally intractable for the purposes of real-time control where decisions need to be made in milliseconds. Prior work Romero et al. [2010] derived a simple ODE model to predict these coupled dynamics. While parts of the derived model are high-confidence physics, parts of it involve intuition-based physics assumptions. In this work, we demonstrate that replacing the physics assumptions with a neural ODE, but keeping the high-confidence physics, yields a model that outperforms both the original model in Romero et al. [2010] and a full neural ODE.

## 2 Background

### 2.1 The Current Diffusion Equation

The principled method to predict the coupled dynamics of plasma current and internal inductance is to evolve a partial differential equation (PDE) known as the current diffusion equation (CDE) Felici [2011]. The CDE, however, is coupled to the plasma density and temperature spatial profiles, which themselves require complex models of plasma turbulence and transport to evolve forward in time. Thus, many existing simulations that simulate the CDE are highly computationally expensive Pereverzev and Yushmanov [2002], Grierson et al. [2018]. To address these challenges, the RAPTOR code was developed using physics-motivated approximations and has been demonstrated to run in real-time with a particular set of settings Felici et al. [2011]. However, the code doesn’t scale well to highly parallel workflows as it has many manually configured settings, has numerical difficulties as a PDE solver, and is implemented in MATLAB which introduces license constraints.

### 2.2 The Romero Model

Motivated by the desire to develop an extremely fast and simple ODE model for internal inductance dynamics, Romero et al. [2010] derived a three state ODE system to model the coupled dynamics of  $L_i$  and  $I_p$ :

$$I_p \frac{dL_i}{dt} = -2V_{ind} - 2V \tag{1a}$$

$$L_i \frac{dI_p}{dt} = 2V_{ind} + V \tag{1b}$$

$$\frac{dV}{dt} \approx -\frac{V}{\tau} - \frac{k}{\tau} V_{ind} \tag{1c}$$

where  $V_{ind}$  is the inductive loop voltage, a control variable that is easily controlled in real-time,  $k, \tau$  are free parameters to be fit to data, and  $V$  is a summary statistic of spatial variables. While the first two equations are based on solid physics, the third equation is an ad-hoc approximation, which motivates the application of ML alternatives.

## 3 Data

We use a dataset of 489 shots, plasma-producing pulses of a fusion experiment, from the Alcator C-Mod fusion experiment. As internal inductance is not a directly measurable quantity, we used

the estimated values from EFIT, a standard software tool in tokamaks used to infer the magnetic properties of a tokamak from sensor measurements Lao et al. [1990]. Of particular interest is that the measurement timebase is non-uniform; the time steps in our dataset have a mean of 19.416ms with a standard deviation of 3ms. While architectures such as recurrent neural networks (RNNs) do not naturally support non-uniform time bases, the Neural ODE framework easily handles this non-uniform timebase. We also performed a normalization of the state and control variables and introduced a constant factor where appropriate to the equations in 1 to account for this normalization.

## 4 Methodology

### 4.1 Neural ODEs with Control Inputs

The classic neural ODE is given by:

$$\dot{\mathbf{x}} = f_{\theta}(\mathbf{x}) \quad (2)$$

In our problem, we have the single control input  $V_{ind}$  and thus we augment the standard Neural ODE with a control vector  $\mathbf{u}$  to arrive at a system with the form:

$$\dot{\mathbf{x}} = f_{\theta}(\mathbf{x}, \mathbf{u}) \quad (3)$$

All of our work was done in the Jax ecosystem, using Equinox to define models and Diffrax for differentiable ODE solvers Kidger and Garcia [2021], Kidger [2021]. The control signal was interpolated in between samples with cubic Hermite splines with backwards differences, a method known for preserving causality Morrill et al. [2021].

### 4.2 Multiple Shooting

It has been observed in the literature that when training Neural ODEs with long time horizons, it is important to utilize “multiple shooting” to ensure stable and effective training Massaroli et al. [2021]. In this work, we apply multiple shooting with a group size of 10 and zero continuity weight, as we found it was not necessary.

### 4.3 Loss Function

During training, each “shot” has a set of measurements  $[\hat{\mathbf{x}}_0, \dots, \hat{\mathbf{x}}_N]$  where  $\hat{\mathbf{x}}_i$  is sampled at time  $t_i$ , and where the sampling time intervals are irregular. We adopt the approach of utilizing an integral loss function to appropriately take into account this irregular time spacing. First, we define an “instantaneous loss” as such:

$$l(t) = \text{Huber}_{\delta=0.1}(\text{RelativeError}(\hat{\mathbf{x}}, \mathbf{x}_t)) \quad (4)$$

Where we apply a Huber loss function with an outlier parameter of 0.1 to the relative error to reduce the effect of outliers Huber [1992]. As we are interested in only predicting the coupled  $L_i, I_p$  dynamics, we excluded the  $V$  variable from this loss function. The loss function that is minimized is the time-integrated value of this instantaneous loss, which we approximate via the trapezoidal rule:

$$\mathcal{L} = \int_0^T l(t) dt \approx \sum_{i=0}^{T-1} \frac{(t_{i+1} - t_i)}{2} [l(t_{i+1}) + l(t_i)] \quad (5)$$

where  $T$  is the total number of time steps of the episode.

### 4.4 Models and Training

In addition to the existing Romero model, eq. 1, we train two more models. The first is a model that replaces equation 1c with a neural ODE mapping state and control to  $\frac{dV}{dt}$ , which we will refer to as “RomeroNNV”:

$$I_p \frac{dL_i}{dt} = -2V_{ind} - 2V \quad (6a)$$

$$L_i \frac{dI_p}{dt} = 2V_{ind} + V \quad (6b)$$

$$\frac{dV}{dt} = f_{V,\theta}(\mathbf{x}, \mathbf{u}) \quad (6c)$$

The second is a MLP that replaces all of the equations, a model we will refer to as “MlpODE”. Both neural networks are fully connected MLPs with a width of 2, depth of 32, and softplus activation. Given the relatively small size of the dataset, the entire training set was included in each batch. We used the AdamW optimizer implemented in Optax with an exponential decay learning rate schedule starting at  $5 \times 10^{-3}$ , a decay rate of 0.995 per epoch, and a terminal value of  $1 \times 10^{-4}$ . The training was terminated if the mean validation loss over the past 1000 epochs had increased.

## 5 Results

While the training loss for the MlpODE continuously decreased and attained the lowest value out of the three models, the validation loss curves show it started over-fitting after several hundred epochs, a behavior not observed with the other two models, which have embedded physics structures. By both validation loss and test set accuracy metrics, shown in Table 1, the RomeroNNV model performed the best with the MlpODE performing the worst on the test set. Figure 1 shows an example of the three models’ predictions against reactor data and control signals.

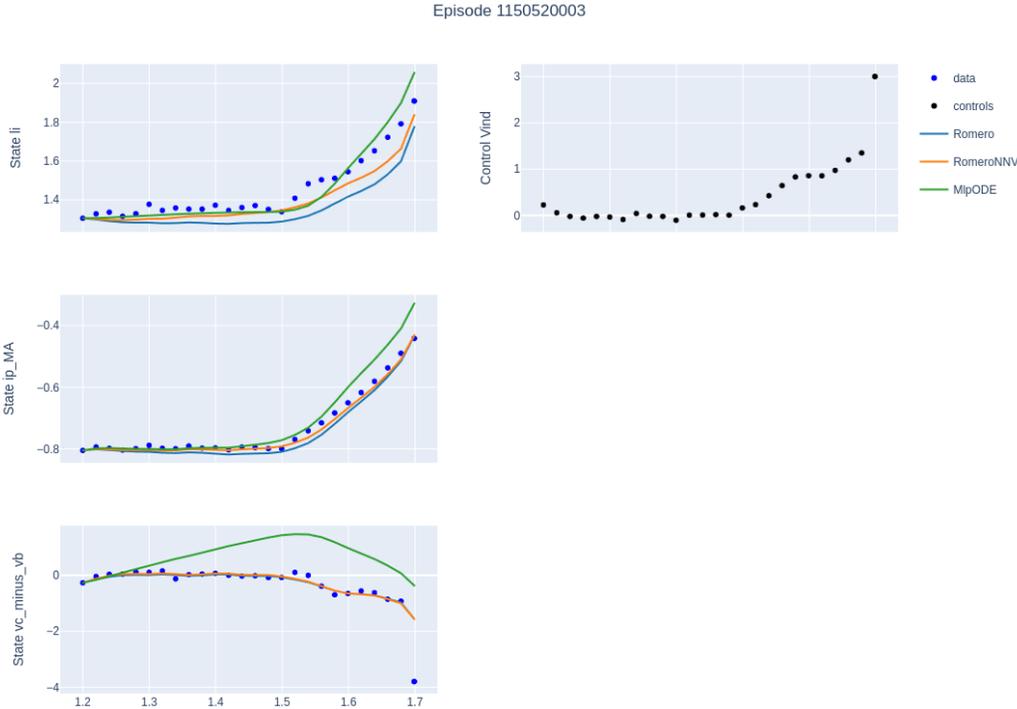


Figure 1: Example of model predictions against reactor data for Alcator C-Mod shot number 1150520003. Note that internal inductance is plotted with the standard tokamak normalization Miyamoto [2005], plasma current is in units of MA, and the  $V$  variable is instead labeled as “vc\_minus\_vb”

Table 1: Test set means and standard deviations of percent error at the end of each episode.

	Romero	RomeroNNV	MlpODE
li	$4.59 \pm 4.37\%$	<b><math>3.50 \pm 3.70\%</math></b>	$5.99 \pm 7.23\%$
Ip	$2.57 \pm 2.34\%$	<b><math>1.97 \pm 1.90\%</math></b>	$8.10 \pm 9.69\%$

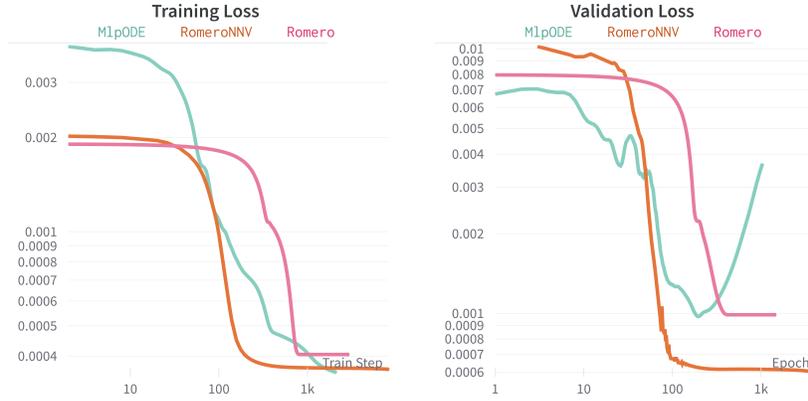


Figure 2: Training and validation losses for the three models.

### 5.1 Double Descent

The increasing validation loss of the MlpODE model towards the end of the training run motivated longer runs of  $3 \times 10^4$  epochs to check for “double descent” behavior, which did appear Nakkiran et al. [2021]. The run shown in Figure 2 shows the MlpODE undergoing multiple rounds of overfitting followed by generalization. The RomeroNNV model proved to exhibit much more stable validation loss behavior, although it did appear to hit a plateau in both training and validation loss for thousands of epochs before descending further.

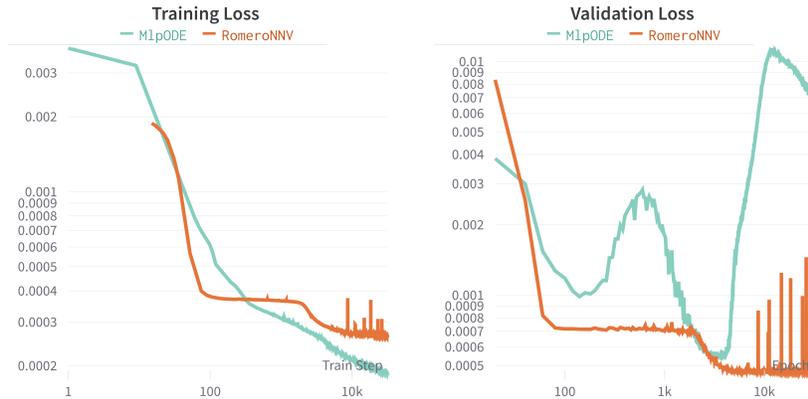


Figure 3: Training and validation loss from a longer training run of the two neural network models. Note that signs of double descent from shorter training runs motivated a more aggressive learning rate schedule starting at  $7.5 \times 10^{-3}$ , ending at  $10^{-3}$ , with a decay rate of 0.9 per every 500 steps.

## 6 Conclusion

We demonstrated the application of the neural ODE framework to predicting plasma dynamics and found that a hybrid dynamics model using both high-confidence physics paired with a neural network, which models the low confidence physics, outperforms the only known existing physics-motivated model, and a pure neural ODE. Future work should investigate adopting a similar modelling approach to a broader range of plasma dynamics.

## 7 Acknowledgements

This work was funded in part by Commonwealth Fusion Systems (CFS). The authors would also like to acknowledge Paula Harder for serving as a workshop mentor and providing feedback on this paper.

## References

- J. Abbate, R. Conlin, and E. Kolemen. Data-driven profile prediction for diii-d. *Nuclear Fusion*, 61(4):046027, 2021.
- I. Char, J. Abbate, L. Bardóczi, M. Boyer, Y. Chung, R. Conlin, K. Erickson, V. Mehta, N. Richner, E. Kolemen, et al. Offline model-based reinforcement learning for tokamak control. In *Learning for Dynamics and Control Conference*, pages 1357–1372. PMLR, 2023.
- J. Degraeve, F. Felici, J. Buchli, M. Neunert, B. Tracey, F. Carpanese, T. Ewalds, R. Hafner, A. Abdolmaleki, D. de Las Casas, et al. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*, 602(7897):414–419, 2022.
- F. Felici. Real-time control of tokamak plasmas: from control of physics to physics-based control. Technical report, EPFL, 2011.
- F. Felici, O. Sauter, S. Coda, B. Duval, T. Goodman, J. Moret, J. Paley, T. Team, et al. Real-time physics-model-based simulation of the current density profile in tokamak plasmas. *Nuclear Fusion*, 51(8):083052, 2011.
- B. Grierson, X. Yuan, M. Gorelenkova, S. Kaye, N. Logan, O. Meneghini, S. Haskey, J. Buchanan, M. Fitzgerald, S. Smith, et al. Orchestrating transp simulations for interpretative and predictive tokamak modeling with omfit. *Fusion Science and Technology*, 74(1-2):101–115, 2018.
- P. J. Huber. Robust estimation of a location parameter. In *Breakthroughs in statistics: Methodology and distribution*, pages 492–518. Springer, 1992.
- P. Kidger. *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021.
- P. Kidger and C. Garcia. Equinox: neural networks in JAX via callable PyTrees and filtered transformations. *Differentiable Programming workshop at Neural Information Processing Systems 2021*, 2021.
- L. Lao, J. Ferron, R. Groebner, W. Howl, H. S. John, E. Strait, and T. Taylor. Equilibrium analysis of current profiles in tokamaks. *Nuclear Fusion*, 30(6):1035, 1990.
- A. D. Maris, A. Wang, C. Rea, R. Granetz, and E. Marmor. The impact of disruptions on the economics of a tokamak power plant. *Fusion Science and Technology*, pages 1–17, 2023.
- S. Massaroli, M. Poli, S. Sonoda, T. Suzuki, J. Park, A. Yamashita, and H. Asama. Differentiable multiple shooting layers. *Advances in Neural Information Processing Systems*, 34:16532–16544, 2021.
- K. Miyamoto. *Plasma physics and controlled nuclear fusion*, volume 38. Springer Science & Business Media, 2005.
- J. Morrill, P. Kidger, L. Yang, and T. Lyons. Neural controlled differential equations for online prediction tasks. *arXiv preprint arXiv:2106.11028*, 2021.
- P. Nakkiran, G. Kaplun, Y. Bansal, T. Yang, B. Barak, and I. Sutskever. Deep double descent: Where bigger models and more data hurt. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124003, 2021.
- G. V. Pereverzev and P. Yushmanov. Astra. automated system for transport analysis in a tokamak. 2002.
- P. Rodriguez-Fernandez, N. Howard, and J. Candy. Nonlinear gyrokinetic predictions of sparc burning plasma profiles enabled by surrogate modeling. *Nuclear Fusion*, 62(7):076036, 2022.

- J. Romero, J.-E. Contributors, et al. Plasma internal inductance dynamics in a tokamak. *Nuclear Fusion*, 50(11):115002, 2010.
- B. Sorbom, J. Ball, T. Palmer, F. Mangiarotti, J. Sierchio, P. Bonoli, C. Kasten, D. Sutherland, H. Barnard, C. Haakonsen, et al. Arc: A compact, high-field, fusion nuclear science facility and demonstration power plant with demountable magnets. *Fusion Engineering and Design*, 100: 378–405, 2015.
- S. E. Wurzel and S. C. Hsu. Progress toward fusion energy breakeven and gain as measured against the lawson criterion. *Physics of Plasmas*, 29(6), 2022.