

IMPROVEMENTS IN EFFICIENCY OF GYROKINETIC SIMULATION RUNS WITH CONVOLUTIONAL NEURAL NETWORK MODELS ANALYZING NONLINEAR SATURATION PROCESSES

E. NARITA

National Institutes for Quantum and Radiological Science and Technology

Naka, Japan

Email: narita.emi@qst.go.jp

M. HONDA

Graduate School of Engineering, Kyoto University

Kyoto, Japan

S. MAEYAMA, T.-H. WATANABE

Department of Physics, Nagoya University

Nagoya, Japan

Abstract

Gyrokinetic codes are the essential tools to predict and understand turbulent transport in magnetically confined fusion plasmas. They calculate the time evolution of the perturbed distribution function in the five-dimensional phase space, and they are performed on supercomputers. It often takes more than a few days to finish their runs and an enormous amount of calculation data is generated per run. When such codes are used to estimate turbulent fluxes, one checks a turbulent saturation state, where the amplitudes of the turbulent fluxes remain almost constant. Since the calculation data produced in the linearly and nonlinearly growing phases sitting prior to the saturation phase are usually not used for turbulent transport analyses, the shorter the growing phase is, the more efficiently the simulations can be carried out. However, there was no way to forecast when the saturation phase starts, i.e. the saturation time, before the simulation is executed or at an early stage of the simulation. In the paper, a new method that forecasts the saturation time is introduced by adopting a machine-learning approach.

When the perturbed distribution function f in the five-dimensional phase space is represented as a combination of two-dimensional images, e.g. $f(k_x, k_y)$ in the wavenumber space, it is visualized that the pattern of the images changes in time showing the saturation process in the linearly and nonlinearly growing phases. In the paper, the images have been generated by the gyrokinetic code GKV [1] to extract the features of the patterns from such images, a convolutional neural network (CNN) model has been constructed, which predicts the simulation time corresponding to the image. The predicted simulation time can be converted into the saturation time. The first CNN model was trained based on the GKV calculation for the JT-60U plasma, where turbulence is driven by both the ion temperature gradient (ITG) mode and the trapped electron mode (TEM). The previous study showed that the calculation gave turbulent fluxes consistent with the experimental ones [2]. About 3,000 images of the amplitude of f in the (k_x, k_y) space were generated from the calculation, and they were split into training, validation and test data. Using the training and validation data, transfer learning and fine tuning have been performed with a pre-trained state-of-the-art CNN model EfficientNet [3] to predict the simulation time from the image. The trained model can predict the simulation time of the test data with a high coefficient of determination, $R^2=0.9949$.

To apply the method to various nonlinear calculations, three CNN models have been prepared, trained by the data of GKV simulations for Cyclone base case (CBC)-like parameter sets. In addition to the simulation with the original CBC case, where turbulence is driven by both ITG mode and TEM, the two cases with equilibrium density and temperature gradients artificially changed to selectively develop ITG mode and TEM, individually have been prepared for training. Focusing on the fact that the nonlinear growth in the nonlinear simulation depends on the dominant instability, the numerically low-cost linear calculation is first performed to identify the dominant instability for the case in which a nonlinear calculation is going to be performed. Then, one can choose the best model out of three to better predict the simulation time. When the three CNN models are applied to the aforementioned JT-60U case, it is found that the CNN trained by the data of the original CBC parameters has the best prediction accuracy. The practical use of the models is as follows: Several runs are launched with different initial conditions that might affect the saturation time. After running the simulations for a while, say six hours after the start, an $f(k_x, k_y)$ image is processed for each case, which is fed into the CNN model. The case predicted to finish the fastest is kept running and the others are terminated. This procedure makes it possible to get the result as fast as possible with minimum computational resources.

[1] T.-H. Watanabe and H. Sugama, Nucl. Fusion **46**, 24 (2006).

[2] M. Nakata et al., Nucl. Fusion **56** 086010 (2016).

[3] M. Tan and Q.V. Le, arXiv: 1905.11946 (2019).

Acknowledgments

This work was supported by MEXT as ‘‘Program for Promoting Researches on the Supercomputer Fugaku’’ (Exploration of burning plasma confinement physics, JPMXP1020200103) and used computational resources of ITO provided by Kyushu University (Project IDs: hp200127 and hp210178).