Nonlinear turbulent transport Co-simulation with heteroscedastic Gaussian processes

Tomonari Nakayama[1], Motoki Nakata[2,3], Shinya Maeyama[4] and Mitsuru Honda[5]

[1] *Tottori University, Tottori, Japan*

[2] *Faculty of Arts and Sciences, Komazawa University, Tokyo, Japan*

[3] *RIKEN iTHEMS, Saitama, Japan*

[4] *National Institute for Fusion Science, Toki, Japan*

[5] *Graduate School of Engineering, Kyoto University, Kyoto, Japan*

Introduction

Achieving the fusion energy with magnetically confined plasmas requires fast and quantitative predictions of turbulent transport and pressure profiles. Global gyrokinetic simulations based on first-principles approaches, such as CGYRO[1], have been extensively studied. These simulations can solve not only self–consistent profile formation but also the multi–species/scale turbulence[2] and the effects of nonlocal turbulent transport phenomena such as ballistic propagation[3]. However, their enormous computational cost makes them impractical for extensive parameter scans for operation scenario optimization and the real–time control, because the global gyrokinetic simulations require huge computational cost. On the other hand, the integrated codes, such as TOPICS[4], that combine some transport surrogate models have been developed for extensive parameter scans for operation scenario optimization and the real–time control. However, the surrogate models used in these integrated codes do not sufficiently capture nonlinear turbulence physics, such as zonal-flow dynamics.

There remains a gap between high–fidelity global gyrokinetic simulations and computationally efficient integrated simulation approaches. To resolve the trade–off between computational cost and physical fidelity, we have developed an Alterable Gyrokinetics-Integrated Transport cOsimulation (AGITO) code[5] which combines a one-dimensional (1D) transport solver with five-dimensional (5D) local gyrokinetic simulations. By using a reduced transport model[6], AGITO code allows for analysis at a lower numerical cost than global gyrokinetic codes. But it still has a high numerical cost for a large number of simulation scans for operation scenario optimization. Therefore, as a first step toward fast prediction of AGITO simulations, we assess the predictive capability of an active-learning-based Gaussian Process Regression (GPR) model using simulation data.

Global transport co–simulation: AGITO

The AGITO code is constructed from 1D global transport solver TRESS[7] and radially distributed 5D local gyrokinetic simulations GKV[8] by using multiple program multiple data (MPMD) parallelization method. GKV requires input parameters, for example the local density gradient R/L_n, temperature gradient R/L_T, and safety factor q. The 1D transport solver TRESS calculates these parameters from the global pressure profile $p(\rho, t)$, where ρ and t denote the normalized radius and time respectively. GKV evaluates the local turbulent diffusivity $\chi_{\mathrm{trb}}/\chi^{\mathrm{GB}}$ normalized by gyro-Bohm heat diffusivity from the received parameters, and the local turbulent diffusivity at each radial position is sent from GKV to TRESS. The turbulent diffusivity is then considered as a system–wide continuous turbulent diffusivity profile via Akima interpolation, which is a piecewise cubic interpolation.

The numelical model used in TRESS is as follows:

$$\frac{3}{2V'}\frac{\partial}{\partial t}V'^{5/3}p_s = \frac{1}{V'}\frac{\partial}{\partial \rho}V'\left(\langle|\nabla\rho|\rangle^2(\chi_{\mathrm{trb,s}} + \chi_{\mathrm{NC,s}})n_s\frac{\partial T_{\mathrm{s}}}{\partial \rho}\right) + S_s. \tag{1}$$

Here, V', χ_{NC}, n and T correspond to the derivative of the volume bounded by a flux surface, the neoclassical heat diffusivity, density and temperature, respectively. The subscript “s” means the particle species. An ad–hoc energy source term resulting from external heating is denoted as S. Turbulent heat diffusivity was obtained from a local gyrokinetic simulation, while neoclassical heat diffusivity was evaluated using the Matrix Inversion method based on the moment approach. TRESS iteratively computes the profile until a steady–state.

The turbulent diffusivity $\chi_{\mathrm{trb,i}}$ is evaluated from nonlinear GKV, which is based on the Gyrokinetic–Poisson equations in the Fourier wavenumber representation. The Gyrokinetic–Poisson equation incurs an enormous computational cost because it is a nonlinear problem in 5D space. Thus, AGITO uses a simplified model which reproduces the results of the nonlinear simulation using only the quantities from linear gyrokinetic simulations such as instability growth rate and the linear response function of zonal flow. In the simplified model, instability growth rate and the linear response function of zonal flow are treated as the quantities characterizing turbulence $\mathcal{L}$ and zonal flow τ_{ZF}. The simplified model is defined as follows:

$$\frac{\chi_{\mathrm{trb,i}}}{\chi_{\mathrm{i}}^{\mathrm{GB}}}(\mathcal{L}, \tau_{\mathrm{ZF}}) = \frac{\Theta_1\mathcal{L}^{\Theta_2}\exp(\Theta_3\tau_{\mathrm{ZF}}^{\Theta_4})}{1 + \Theta_5\mathcal{L}^{\Theta_6}\exp(\Theta_7\tau_{\mathrm{ZF}}^{\Theta_4})[\mathcal{H}(\tau_{\mathrm{ZF}})]^{\Theta_8}}, \tag{2}$$

where $(\Theta_1, \cdots, \Theta_8)$ and $\mathcal{H}$ are fitting parameter and zonal flow function, respectively (see Ref. [6]). Since $\mathcal{L}$ and zonal flow τ_{ZF} are obtained by linear gyrokinetic simulation, the heat diffusivity can be determined using the simplified model without the need for the nonlinear gyrokinetic simulation. This model allows the computational cost to be reduced to about 1/200 of the nonlinear gyrokinetic simulation.

Verification for Gaussian Processes Regression model

Gaussian Process Regression (GPR) is a probabilistic approach for estimating an unknown function $f(\boldsymbol{x})$ that maps input variables $\boldsymbol{x}$ to an output variable y. GPR can provide predictions without performing computationally expensive simulations such as GKV. Moreover, the uncertainty of the prediction can be quantified through the predictive variance. A GPR model is characterized by a mean function and a covariance function, and a Gaussian kernel is adopted as the covariance function in the present study. Conventional GPR generally assumes a uniform observation noise level for all training data. In this study, GPR is first applied to transport and profile data obtained from AGITO, and its capability for predicting turbulent transport is evaluated.

The integration of GPR into AGITO is performed in an active-learning framework. During the initial stage (approximately the first ten transport time steps), independent GPR models are constructed for each radial position using temperature gradients and temperatures as input variables and the turbulent transport evaluated by GKV as training outputs. After the initial training, turbulent transport is predicted by GPR. If the predictive variance is below a prescribed threshold, the GPR prediction is adopted as the turbulent diffusivity. Otherwise, high-fidelity turbulent transport is evaluated using GKV. In the latter case, the newly obtained data are added to the training dataset and the GPR model is retrained. This procedure reduces the number of expensive GKV evaluations.

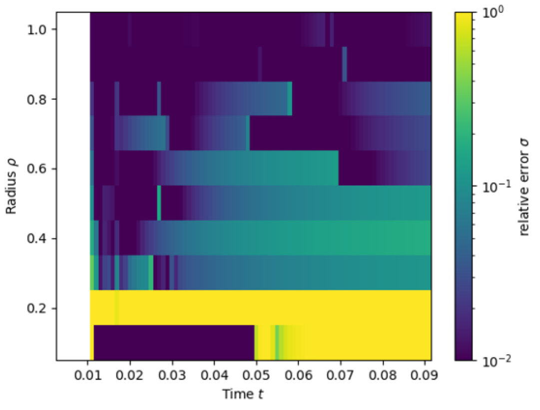

Figure 1: Spatiotemporal map of the local relative error in AGITO predictions obtained with GPR

Figure 1 shows the prediction error map obtained with the GPR-integrated AGITO simulation. Reasonably accurate predictions are obtained in the region of $\rho \geq 0.3$, whereas the model fails to predict the transport accurately at $\rho = 0.1$ and 0.2. This degradation originates from overfitting to the initial training data, which are primarily located below the critical-gradient threshold where turbulent transport remains nearly zero even when the temperature gradient increases. Because the transport response changes rapidly around the critical gradient, the assumption of uniform observation noise in conventional GPR is insufficient for achieving accurate predictions. Therefore, an extension of the GPR model is required so that the influence of subcritical data is reduced while maintaining high predictive accuracy in regions with finite turbulent transport.

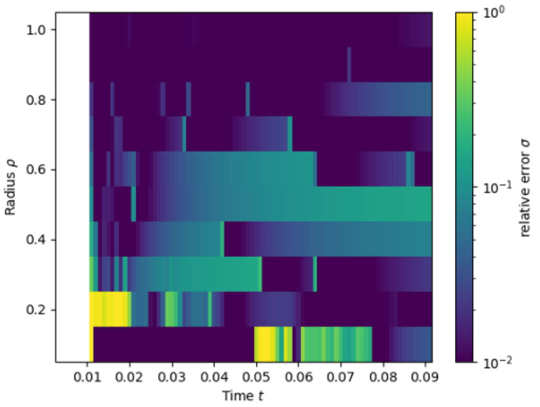

Figure 2: Spatiotemporal map of the local relative error in AGITO predictions obtained with HGPR

To address this issue, a Heteroscedastic Gaussian Process Regression (HGPR) model[9], which allows data-dependent observation noise, is introduced. In HGPR, the observation noise directly affects the contribution of individual data points to model construction. The observation noise is parameterized using the relative zonal-flow intensity, which varies sharply near the critical-gradient threshold. As a result, the influence of data in subcritical regions dominated by zonal flows is reduced, whereas data in turbulence-dominated regions above the critical gradient contribute more strongly to the model. Figure 2 shows the prediction results obtained with HGPR. The introduction of HGPR significantly improves the prediction accuracy at $\rho = 0.1$ and 0.2, including the critical-gradient region.

References

[1] J. Candy, et al., J. Comput. Phys. 324 73–93 (2016)

[2] D. Esteve, et al., Nucl. Fusion 58.3 (2018)

[3] X. Garbet and R. E. Waltz, Phys. Plasmas 5 2836–45 (1998)

[4] N. Hayashi and JT–60 Team, Phys. Plasmas 17 056112 (2010)

[5] T. Nakayama, et al., Plasma Phys. Control. Fusion 67 (2025)

[6] T. Nakayama, et al., Sci. Rep. 13 2319 (2023)

[7] M. Honda and M. Nakata, ITPA Transport and Confinement Topical Group, (2016)

[8] T-H.Watanabe and H. Sugama, Nucl. Fusion 46 24–32 (2006)

[9] P. Goldberg, et al., Advances in neural information processing systems 10 (1997).