Neural-Network-based Probabilistic Surrogate Modeling for Rapid Bayesian Inference in Thomson Scattering Diagnostics

Liyun Zhang[1], **Mamada Naoya**[1], **Kentaro Sakai**[2], **Takeo Hoshi**[2], **Toru Aonishi**[1]

[1] *Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 5-1-5, Japan*
[2] *National Institute for Fusion Science, Toki, Gifu 509-5292, Japan*

Email: liyun.zhang@edu.k.u-tokyo.ac.jp

Thomson scattering (TS) is a primary diagnostic for measuring electron temperature (T_e) and density (n_e) in fusion plasmas. While Bayesian inference via Markov Chain Monte Carlo (MCMC) provides effective parameter estimation, its real-time application is hindered by the heavy computational burden of physical forward models and extreme sensitivity to statistical noise in low-photon-yield regimes [1].

In this work, we develop a neural-network-based probabilistic surrogate framework, defined as $M = EM + N$, to facilitate rapid and reliable MCMC inference. The generative model utilizes a dual-head neural network architecture to simultaneously estimate the expected physical emission spectrum (EM, serving as the mean μ) and the measurement noise variance (N, serving as σ^2). To address the inherent training instabilities and ensure physical validity under a Gaussian Negative Log-Likelihood (GNLL) loss, we implement a log-variance formulation. By configuring the network to predict the logarithmic variance ($log\ \sigma^2$) rather than the absolute variance, the architecture intrinsically guarantees the strict positivity of the inferred statistical uncertainty. This architectural constraint prevents numerical explosions and variance collapse during gradient descent, maintaining the physical consistency of the model without the need for complex regularizations [2, 3].

Evaluations against the established Monte Carlo simulation approach [4] indicate that the $M = EM + N$ framework exhibits enhanced fitting robustness in noise-dominated environments. Compared to traditional physical models, this surrogate approach significantly accelerates the MCMC inference process, effectively overcoming the computational bottleneck for real-time analysis. Furthermore, the methodology demonstrates promising potential in mitigating parameter covariance, providing a high-fidelity "intelligent measurement" foundation for future predictive digital twin systems in fusion energy research [5].

This work was supported by JST (Moonshot R&D Program) Japan Grant Number JPMJMS24A3.

References

[1] Sakai, K., Tomita, K., Hoshi, T., Nakano, A., Goto, M., Nagaoka, K., & Yasuhara, R. (2026). Conceptual design of Thomson scattering system with high wavelength resolution in magnetically confined plasmas for electron phase-space measurements. Plasma Physics and Controlled Fusion, 68(1), 015034.
[2] Kendall, A., & Gal, Y. (2017). What uncertainties do we need in bayesian deep learning for computer vision?. Advances in neural information processing systems, 30.
[3] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, 686-707.
[4] Sakai, K., Tomita, K., Hoshi, T., & Yasuhara, R. (2026). Monte Carlo simulation method for incoherent Thomson scattering spectra from arbitrary electron distribution functions. Journal of Instrumentation, 21(01), C01002.
[5] Hammel, B. A., Hammel, B. D., Scott, H. A., & Peterson, J. L. (2024). Machine learning assisted bayesian inference of mix and hot-spot conditions in NIF implosions. High Energy Density Physics, 50, 101077.