Data-driven dynamics model of Mirnov-coil spectrograms for L-H onset timing towards tokamak experiment design

Hyeongjun Noh[1], Chweeho Heo[1], Xiaotian Gao[2*] and Yong-Su Na[1**]

[1] *Seoul National University (SNU), Seoul, Korea*

[2] *Zhongguancun Academy, Beijing, China*

Recently, data-driven methods have been applied to tokamak researches such as disruption prediction, tearing-mode prediction, plasma-state classification, plasma control, and surrogate modelling. While these approaches have been effective, many models are formulated for predefined labels or specific plasma states, limiting their reuse when the downstream physics question changes.

Instead of predicting a specific plasma-state label, our model predicts future signals from Mirnov-coil (MC) diagnostics. The L-H transition is then used as a downstream readout to test whether the predicted MC retains sufficient information to infer the H-mode onset.

KSTAR data and Mirnov-coil representation

The model is trained on KSTAR [1] experimental discharges using time sequences of operational control variables and MC spectrograms. The input variables comprise the toroidal magnetic field and plasma current (B_t, I_p), external heating parameters, including EC heating powers and resonant (r,z) positions for channels 2–5 and NB1/NB2 heating powers for beamlines A–C, fuelling (GAS1–3 and PVD), plasma-shape quantities (a, R_0, dR_{sep}, diverted/limited configuration, κ, δ_u, and δ_l), and Mirnov-coil signals. The time grid is set to $\Delta t = 50\,\mathrm{ms}$ to match the time resolution of EFIT.

The MC signals are processed through cross-spectral analysis [2], yielding a two-channel representation consisting of cross-power and cross-phase spectrograms, $\mathbf{m}_t \in \mathbb{R}^{F\times T\times 2}$. Here, $F = 128$ is the number of frequency bins, $T = 48$ is the number of time bins within each Δt interval, and the two channels correspond to cross-power and cross-phase.

Early plasma ramp-up phases and discharges outside the selected pulse-duration range ($1.7\,\mathrm{s} < t_{\mathrm{pulse}} < 20\,\mathrm{s}$) are excluded. The dataset consists of 978 KSTAR discharges from the 2020–2022 campaigns, divided into training, validation, and test subsets with an 8:1:1 ratio.

Rather than directly predicting dense, high-dimensional MC spectrograms that evolve rapidly near L-H transitions and MHD activity, we reformulate future MC-spectrogram prediction as a discrete-token prediction problem. A Vector Quantised-Variational AutoEncoder (VQ-VAE) [3] is trained to map each MC chunk, through its encoder, into a compact sequence of discrete la-

tent indices, or tokens: $\mathbf{m}_t \in \mathbb{R}^{F\times T\times 2} \xrightarrow{\text{VQ-VAE}} \mathbf{z}_t \in \{1,\ldots,V_o\}^d$, where $V_o = 512$ is the codebook size and $d = 96$ is the number of tokens per MC chunk. The decoder reconstructs the spectrogram from the corresponding codebook entries. In this representation, the Transformer [4] does not predict each spectrogram pixel directly. Instead, it predicts the next set of discrete MC tokens, and the VQ-VAE decoder maps the predicted tokens back to the original spectrogram space.

This tokenisation has three practical advantages for the present application. First, it substantially compresses the MC spectrogram representation, reducing the sequence length and computational burden for autoregressive prediction. Second, it converts future spectrogram prediction into a classification problem over a learned codebook, which is naturally compatible with next-token prediction. Third, the decoder separates the reconstruction of fine spectrogram structure from the temporal prediction of latent MC evolution. In the trained tokenizer, the MC spectrogram representation is compressed by a factor of approximately 258, while maintaining high reconstruction fidelity, with reconstruction PSNR values of 31.6 dB for cross-power and 24.1 dB for cross-phase.

Scalar control variables are also discretised. The plasma-control token $\mathbf{c}_t$ contains actuator and shape quantities for the discharge scenario. In the autoregressive prediction task, the model receives recent MC-token history and the corresponding prescribed control sequence, and predicts future MC tokens.

PanoMHD: Control-conditioned autoregressive prediction

Analogous to the way a large language model (LLM) predicts the next text token from a text context, Panoramic MagnetoHydroDynamics (PanoMHD) [5] predicts the next MC tokens from a plasma context composed of recent MC history and prescribed operational control variables. The model uses a GPT-2-style decoder-only Transformer architecture [6], trained from scratch on KSTAR plasma-token sequences.

For a context length L, the model is trained to represent the conditional distribution $p_\theta\left(\mathbf{z}_t \mid \mathbf{z}_{t-L:t-1}, \mathbf{c}^{\text{pre}}_{t-L+2:t}\right)$, where $\mathbf{z}_{t-L:t-1}$ denotes the recent MC-token history and $\mathbf{c}^{\text{pre}}$ denotes the prescribed control variables. For a rollout starting at time t_0, the predicted tokens are recursively generated as

$$\hat{\mathbf{z}}_{t_0+k} \sim p_\theta\left(\mathbf{z}_{t_0+k} \mid \tilde{\mathbf{z}}_{t_0+k-L:t_0+k-1}, \mathbf{c}^{\text{pre}}_{t_0+k-L+2:t_0+k}\right),\ \tilde{\mathbf{z}}_s = \begin{cases} \mathbf{z}_s, & s < t_0, \\ \hat{\mathbf{z}}_s, & s \geq t_0, \end{cases}$$

where $k = 0, 1, \ldots$ and $\tilde{\mathbf{z}}$ denotes the rollout context, which contains measured MC tokens before t_0 and predicted MC tokens from t_0 onward. In this work, we focus on whether the predicted

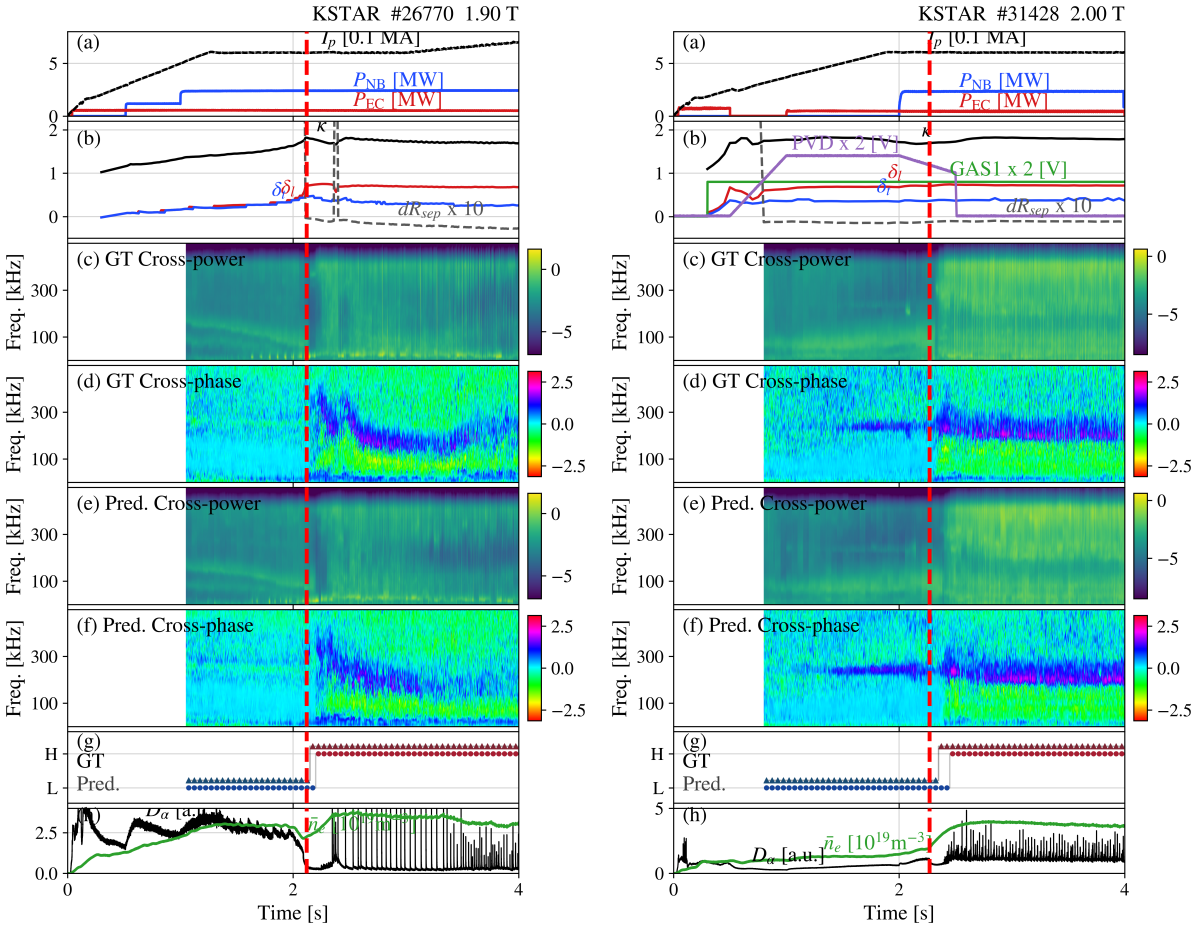

Figure 1: Representative held-out L-H transition replays for two KSTAR discharges. Shown are (a), (b) KSTAR control variables, (c)–(f) measured and PanoMHD-predicted MC spectrograms, (g) OASIS-derived L/H readouts from measured (GT) and predicted (Pred.) MC, and (h) independent D_α and $\bar{n}_e$ traces. The red vertical line marks the transition time inferred from (h).

MC $\hat{\mathbf{m}}$ preserves information needed to infer the L-H transition.

OASIS readout for L-H transition and evaluation

OASIS [7] is a KSTAR-validated L/H mode classifier that operates on MC spectrograms. The same classifier is applied to measured MC spectrograms and to MC spectrograms predicted by PanoMHD. For measured and predicted MC spectrogram sequences, the corresponding OASIS-derived L/H sequences are written as $OASIS(\mathbf{m}_t)$,$OASIS(\hat{\mathbf{m}}_t)$.

The transition onset time is then obtained from the corresponding OASIS-derived L/H sequence. The transition-time error is defined as $\Delta t_{LH}^{OASIS} = t_{LH}^{OASIS(\hat{\mathbf{m}})} - t_{LH}^{OASIS(\mathbf{m})}$.

Figure 1 compares measured and autoregressively predicted MC spectrograms for held-out KSTAR shots. Applying the same OASIS classifier to measured and predicted MC gives similar L/H mode sequences. The independent D_α signal, not used by PanoMHD or OASIS, exhibits a baseline reduction near the same time, while $\bar{n}_e$ increases. These trends are consistent with

H-mode access and provide physical corroboration of the MC/OASIS-based transition readout.

Figure 2 shows the onset-time agreement evaluated over held-out KSTAR discharges containing L-H transitions. The mean transition-time error is $\langle \Delta t_{LH}^{OASIS} \rangle = +0.019\,\mathrm{s}$, and the mean absolute error (MAE) is 0.064 s. The MAE is close to the model time resolution of $\Delta t = 50\,\mathrm{ms}$, suggesting that the predicted spectrograms retain L-H-onset information at a temporal accuracy comparable to a single prediction time step.

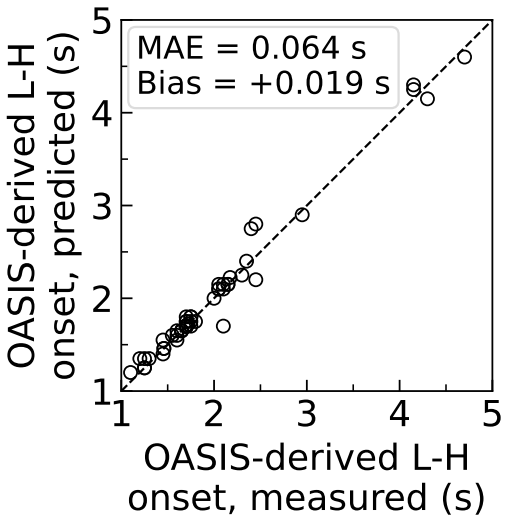

Figure 2: OASIS-derived L-H onset-time comparison. The reference onset time is obtained by applying OASIS to measured MC spectrograms, while the predicted onset time is obtained by applying the same OASIS classifier to predicted MC spectrograms.

Conclusion

The results suggest that predicted MC spectrograms can provide a reusable diagnostic representation for downstream plasma-state readouts. Here, the readout is limited to the L-H transition, for which OASIS-derived onset times from predicted MC spectrograms agree with those from measured MC spectrograms.

This validation is retrospective on held-out KSTAR discharges, and validation on newly executed discharges remains necessary before using such predictions for experiment-design studies. The model is not a mechanistic description of the L-H trigger. Prospective validation and extension to ELMs, tearing modes, and other MHD-instability readouts remain future work.

References

[1] G.S. Lee *et al.* Nuclear Fusion, **40**, 575 (2000)

[2] J. S. Kim *et al.* Plasma Physics and Controlled Fusion, **41**, 1399 (1999)

[3] A. van den Oord *et al.* NeurIPS, **30** (2017)

[4] A. Vaswani *et al.* NeurIPS, **30** (2017)

[5] H. Noh *et al.* arXiv:2603.02672 (2026)

[6] A. Radford *et al.* OpenAI (2019)

[7] H. Noh *et al.*, 51st EPS Conf. on Plasma Physics (EPS 2025), Vilnius, Lithuania, Europhysics Conf. Abstracts, **51A**, P1.029 (2025).