

CONFERENCE PRE-PRINT**NEURAL NETWORK REDUCED MODELS FOR PLASMA TURBULENCE**

Z.S. QU, K.P. LI, J.C. HUANG, Y.W. CHO, X. GARBET, R. VARENNES, C.G. WAN, R.C. ZHANG, A. SHAA
C.G. GUET, K. LIM, D. NIYATO, Y.S. ONG

Nanyang Technological University

Singapore

Email: zhisong.qu@ntu.edu.sg

V. GRANDGIRARD

IRFM, CEA Cadarache

Saint-Paul-lès-Durance, France

Abstract

Plasma turbulence plays a central role in determining energy and particle transport in magnetically confined fusion systems. However, direct numerical simulations of turbulence remain computationally expensive at reactor-relevant scales. This paper presents two neural-network-based reduced models for capturing key features of drift-wave turbulence and its interaction with zonal flows. First, a physics-informed neural network framework is applied to the Hasegawa–Wakatani equations to identify closure coefficients for large-eddy simulations. By training on high-resolution data from the TOKAM2D code, the reduced Extended HW-Closure (EHW-C) model reproduces spectral properties and transport fluxes across near-hydrodynamic and near-adiabatic regimes with over 90% reduction in computational cost. Second, a stochastic surrogate model is developed for the turbulence–zonal flow predator–prey dynamics. Neural networks are used to infer the drift and diffusion terms of coupled stochastic differential equations. The surrogate reproduces both oscillatory predator–prey dynamics and stochastic phenomena observed in first-principles simulations. Together, these models demonstrate how neural networks can provide efficient, physically grounded surrogates of plasma turbulence.

1. INTRODUCTION

Understanding plasma turbulence and its self-organisation is a longstanding challenge in fusion research [1]. Plasma turbulence and its nonlinear interactions with zonal flows [2] are recognised as the dominant mechanisms governing anomalous transport in magnetically confined plasmas. Although direct numerical simulations (DNS) of gyrokinetic and/or fluid models have provided significant insights [3], their high computational cost limits parameter scans, transport-time-scale simulations, and applications in integrated modelling suites. This motivates the development of reduced models that can retain essential turbulent dynamics while substantially lowering computational requirements.

Recent advances in scientific machine learning, particularly physics-informed neural networks (PINNs) and data-driven surrogate modelling [4], provide promising tools for this task. Physics-aware AI and wider machine-learning toolkits are already exerting a measurable impact across fusion science. Deep reinforcement-learning controllers have been demonstrated on real tokamaks to shape and maintain plasmas far more flexibly than hand-tuned regulators, most notably in TCV [5]. Physics-aware, real-time compatible predictors and controllers have been proposed for disruption avoidance and tearing-instability suppression on large devices, demonstrating practical, safety-critical applications of AI in tokamak operation [6, 7]. Complementary efforts use ML to accelerate core physics codes and build high-fidelity surrogates, reducing runtimes to milliseconds for certain transport and heating calculations, thus enabling fast parameter scans, uncertainty quantification, and near-real-time decision support for experiments. [8, 9, 10].

Using the Hasegawa–Wakatani (HW) equations [11] as an example, this manuscript develops two neural-network-based reduced modelling approaches to ease the computational cost of DNS, and to extract the relationship between the key physics players. First, a closure model for large-eddy simulations is constructed by training PINNs on high-resolution simulations of the HW system, resulting in the Extended HW-Closure (EHW-C) equations [12]. Second, a stochastic neural surrogate model is introduced for turbulence–zonal flow dynamics,

where stochastic differential equations with neural drift and diffusion terms are constrained by physics-based assumptions [13]. Together, these approaches aim to advance turbulence modelling toward predictive, efficient, and physically grounded frameworks relevant for fusion reactor scenarios.

2. THE HASEGAWA-WAKATANI EQUATIONS

The Hasegawa-Wakatani (HW) equations are arguably the simplest paradigm to study drift-wave turbulence and its interaction with zonal flows. The equations couple the perturbed density of the plasma, n , and the perturbed electrostatic potential, ϕ . In a two-dimensional (2D) system, the coordinates x and y are the radial coordinate and the poloidal direction coordinate, respectively, both with periodic boundary conditions. These equations are given by

$$\begin{aligned}\frac{\partial \omega}{\partial t} + \{\phi, \omega\} &= \alpha(\phi - n) - D\nabla^4 \omega - \nu \langle \omega \rangle_y, \\ \frac{\partial n}{\partial t} + \{\phi, n\} &= \alpha(\phi - n) - \kappa \frac{\partial \phi}{\partial y} - D\nabla^4 n,\end{aligned}\tag{1}$$

where $\omega = \nabla^2 \phi$ is the vorticity, ∇^2 the two-dimensional Laplacian operator in 2D, and $\langle \dots \rangle_y$ stands for averaging over the y direction. The term $\{\phi, n\} = \frac{\partial \phi}{\partial x} \frac{\partial n}{\partial y} - \frac{\partial n}{\partial x} \frac{\partial \phi}{\partial y}$ is the standard Poisson bracket operator. The hyper-diffusion coefficient D is introduced for numerical stability. The drive of the system is specified by $\kappa = -\frac{\partial \ln n_0}{\partial x}$, the density gradient scale-length, where n_0 is the plasma density of the reference equilibrium. The adiabatic parameter α is proportional to the plasma conductivity. In the collisionless case $\alpha \rightarrow \infty$, the density follows the Boltzmann relation $n = \phi$, and Eq. (1) becomes the Hasegawa-Mima equation [14]. In the highly collisional case $\alpha \rightarrow 0$, the HW equations resemble the incompressible 2D Navier-Stokes equation. In addition, a zonal flow damping term $-\nu \langle \omega \rangle_y$ can be added to mimic the collisional damping of zonal flow, which is crucial in the predator-prey dynamics, as discussed later. We will build our LES closure model based on the HW equations in Section 3.

The HW system only has weak zonal flows due to the damping term $\alpha(\phi - n)$ on the right-hand side of Eq. (1). The modified Hasegawa-Wakatani (mHW) equations change this term to $\alpha(\tilde{\phi} - \tilde{n})$, only acting on the non-zonal component, where

$$\tilde{\phi} = \phi - \langle \phi \rangle_y, \quad \tilde{n} = n - \langle n \rangle_y.\tag{2}$$

We will use the mHW equations to build our turbulence-zonal-flow NN model in Section 4.

The simulations of the HW and mHW equations are performed by the TOKAM2D code [15]. TOKAM2D uses spectral methods for spatial dimensions, and fourth-order Runge-Kutta for time advection. It is GPU compatible: in our case, a simulation takes a few minutes to one hour on a single A100 card, depending on the system size and the number of time steps.

3. LARGE EDDY SIMULATIONS OF DRIFT WAVE TURBULENCE

We introduce the Extended Hasegawa-Wakatani with Closure (EHW-C) model, shown as

$$\frac{\partial \omega}{\partial t} + \{\phi, \omega\} = \alpha(\phi - n) + D_{nn} \nabla^2 n + D_{n\omega} \nabla^2 \omega - \mu_n \nabla^4 n,\tag{3}$$

$$\frac{\partial n}{\partial t} + \{\phi, n\} = \alpha(\phi - n) - \kappa \frac{\partial \phi}{\partial y} + D_{\omega\omega} \nabla^2 \omega + D_{\omega n} \nabla^2 n - \mu_\omega \nabla^4 \omega.\tag{4}$$

Compared to Eq (1), the second to fourth terms on the R.H.S. of Eqs. (3) and (4) are the closure terms derived from the Direct Interaction Approximation (DIA) theory [16]. Specifically, D_{nn} and $D_{\omega\omega}$ are the diffusion coefficients for density and vorticity fluctuations, respectively. $D_{n\omega}$ and $D_{\omega n}$ are the cross terms, representing stresses between the density by the vorticity fields. μ_n and μ_ω are the hyper-diffusion terms. Although analytical expressions for these coefficients have been derived in previous works [17, 18], they require exact information about the nonlinear interaction from short wavelength fluctuations. Therefore, directly measuring these coefficients and applying them to low-resolution simulations is challenging.

In this work, we utilise PINNs to determine these coefficients. By embedding the governing equations directly into the loss function alongside boundary and initial conditions, PINNs seamlessly integrate fundamental physical laws into the network architecture, thereby recasting traditional PDE formulations as equivalent constrained optimisation tasks. For inverse problems, the unknown parameters within the PDEs can be treated as additional trainable variables. In this way, PINNs can allow a recovery of latent physics and solutions in a single end-to-end learning loop.

For this purpose, the TOKAM2D code is employed to generate the training data by Hasegawa-Wakatani equations and to perform simulations of EHW-C model. The high resolution simulations use 512×512 grids, corresponding to cutoff wave-number $k_{cutoff} \sim 15$. In contrast, the low resolution simulations and the simulations for EHW-C model utilise 64×64 grids, for which the cutoff wave-number $k_{cutoff} \sim 2.5$. Other parameters are chosen to be $L_x = L_y = 51.5$, $\kappa = 1$, $D = 0.0001$, and $\nu = 0$, while C varies in different simulations. To construct low-resolution training inputs, we adopt a down-sampling strategy that maps the high-resolution data to 64×64 grids by extracting every eighth point along each spatial dimension from the simulation fields (plasma density n , potential ϕ , and vorticity ω). The down-sampled data serves as input for training the PINNs framework to determine closure coefficients. We systematically evaluate a set of candidate models known for their effectiveness in capturing spatiotemporal dynamics, and chose the convolutional long short-term memory networks (ConvLSTM) [19]. ConvLSTM integrates convolutional operations within the gating mechanisms of LSTM units, enabling efficient modelling of both spatial and temporal correlations.

To verify and validate the closure coefficients identified by the neural network, we adopt a physics-based evaluation strategy where the identified coefficients are directly fed into the numerical solver (TOKAM2D). More specifically, we use DNS to solve the EHW-C model with the inferred parameters on a low-resolution grid, and the resulting simulation is taken as the sole criterion for assessment, to ensure that the learned coefficients are not only numerically reasonable but also physically meaningful. The quality of the identified parameters is assessed based on their ability to reproduce key macroscopic and statistical features of the turbulent system. The simulation of the EHW-C model on a low-resolution grid is compared with high-resolution reference solutions. Only when the reproduced fields and spectra align well with the ground truth, the identified coefficients are considered valid and reliable.

The representative results are summarised in Fig. 1. For the near-hydrodynamic regime ($C = 0.2$), the flux evolution shows that the low-resolution simulation without closure severely underestimates the particle flux, remaining nearly constant at a low value throughout the simulation time. In contrast, the EHW-C model successfully recovers the fluctuating behaviour observed in the high-resolution reference, capturing both the magnitude and temporal variations of the particle flux. The spectral analysis reveals that EHW-C model accurately reconstructs the spectral slopes across all three quantities ($|n_{k_y}|^2$, $|\phi_{k_y}|^2$, and $|n_\omega|^2$). Most notably, the EHW-C model correctly captures the steeper decay in the potential spectrum, which is a signature characteristic of this regime. This spectral consistency reinforces the physical validity of the EHW-C model in capturing the correct energy transfer mechanisms of the near-hydrodynamic regime.

In the near-adiabatic limit ($C = 5.0$), the situation becomes more subtle. Unlike the $C = 0.2$ case, the spatial spectra show relatively similar behaviour across all three cases, making it difficult to distinguish the closure effects through spectral analysis alone. Additionally, the temporal flux evolution exhibits smaller differences among the three cases, compared to the $C = 0.2$ case, with all simulations showing comparable flux saturation levels. However, the frequency spectrum reveals crucial differences that are not apparent in the spatial domain. As shown in Fig. 1 (b), the frequency spectrum demonstrates that EHW-C model successfully captures the proper temporal dynamics of the turbulent fluctuations. While the low-resolution simulation fails to reproduce the correct frequency distribution, particularly at higher frequencies.

Across both regimes, the EHW-C model achieves the same level of accuracy on a significantly coarser spatial grid (64×64) compared to the high-resolution DNS benchmark (512×512). This represents a 64-fold reduction in grid points, leading to over 90% reduction in computational time while preserving essential turbulent characteristics. The consistent performance across different physical regimes demonstrates the robustness and generalisation capability of our physics-informed approach.

4. STOCHASTIC SURROGATE MODEL FOR THE TURBULENCE-ZONAL FLOW DYNAMICS

In this section, we will extract the form and parameters of a predator-prey model describing the interaction between turbulence and zonal flow directly from TOKAM2D simulations of the mHW equations. From the simulation data, we take the volume averaged turbulence energy E and the zonal flow energy U as proxies for the predator and the prey, respectively. They are given by

$$E = \frac{1}{V} \int \frac{1}{2} (|\nabla \tilde{\phi}|^2 + \tilde{n}^2) dV, \quad (22)$$

$$U = \frac{1}{V} \int \frac{1}{2} |\langle \partial_x \phi \rangle_y|^2 dV, \quad (23)$$

where V is the area of the simulation domain.

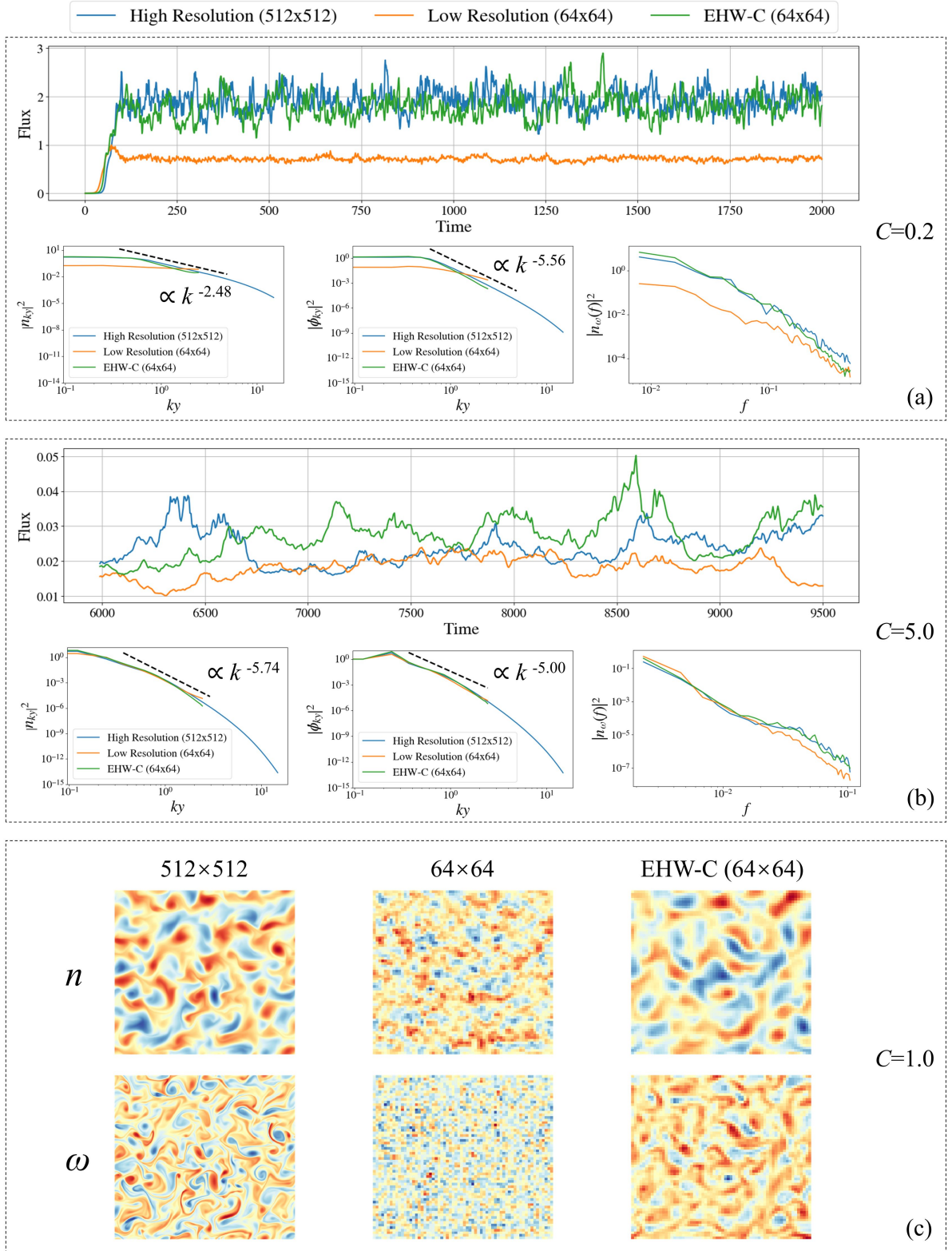


FIG. 1. Comparison of simulations under different resolutions. Subplot (a) and (b) compares the specific quantities at the limit cases, four key turbulent features are shown: the time evolution of particle flux, the density spectrum $|n_{ky}|^2$, the potential spectrum $|\phi_{ky}|^2$, and the frequency spectrum $|n_{\omega}(f)|^2$. Each plot compares the results from high resolution DNS (512×512), low resolution DNS (64×64), and the proposed EHW-C model at the same low resolution (64×64). Subplot (c) compares the 2D plots of the case of $C = 1$.

The simplest version of predator-prey model takes the form of the Lotka–Volterra equations, given by

$$\begin{aligned}\frac{\partial E}{\partial t} &= \alpha E - \beta EU, \\ \frac{\partial U}{\partial t} &= -\gamma U + \sigma EU,\end{aligned}\tag{5}$$

where E and U are the energies associated with turbulence and zonal flow shear, respectively. The coefficients α , β , γ and σ are positive constants. Here, E grows exponentially with growth rate α due to linear drift-wave instability, representing the “prey”. The zonal flow shearing effect suppresses the turbulence via the cross term $-\beta EU$, representing the “predator”. Meanwhile, the zonal flows are driven via Reynolds stress σEU , and linearly damped due to collisional effects $-\gamma U$. This model and its variant were successfully applied to capture the dynamic between the turbulence energy and the zonal flow shear energy [2].

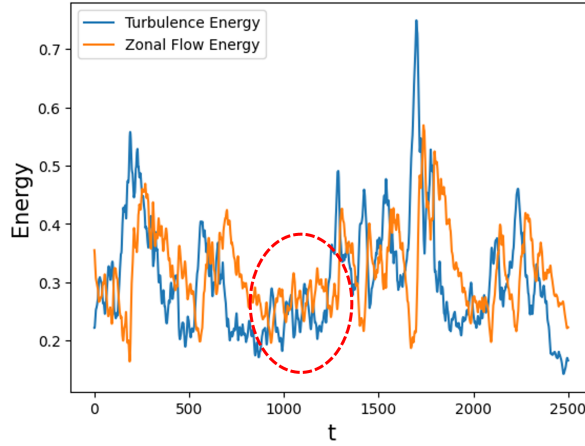


FIG. 2. Turbulence and zonal flow energy extracted from a simulation of the modified Hasegawa-Wakatani equations using the TOKAM2D code. A stagnation phenomenon is indicated by the red dash circle.

Despite the qualitative success of the predator-prey model in describing turbulence-zonal flow interaction, direct extraction of the model from data is a difficult task. An example is given in FIG. 2 showing the stochastic time evolution of the turbulence energy and zonal flow energy. Even though the signature of Lotka–Volterra is clear, as the zonal flow energy lags behind the turbulence energy, the amplitude and period of oscillation change from peak to peak. The reason behind this is that the dynamics between turbulence and zonal flows are complicated, as turbulence itself is inherently chaotic and nonlinear. To capture features in such a system by a physics model, one attempt is to generalise coefficients α , β , γ and σ to functions of E and U , and to introduce terms that capture small fluctuations in data, in our case, stochastic terms. Indeed, stochastic predator-prey model [20] is suitable for capturing the dynamics of turbulence and zonal flow [21]. In this manuscript, we postulate the interaction between turbulence and zonal flow to take the form of stochastic differential equations (SDEs) given by

$$dE = g_{11}(E, U)dt + g_{21}(E, U)dw_1,\tag{6}$$

$$dU = g_{12}(E, U)dt + g_{22}(E, U)dw_2,\tag{7}$$

where g_{11} and g_{12} are “drift” terms while g_{21} and g_{22} are “diffusion” terms related to stochasticity that captures small fluctuations. The terms dw_1 and dw_2 are Brownian motions (can be correlated or independent). Neural network is a suitable tool for representing the unknown g functions and has been known to possess the potential to learn nonlinear relationships.

The parameters of the mHW models are chosen to be $\kappa = 1.05$, $\alpha = 2$, $D = 0.0001$ and $\nu = 0.01$. The box size is $L_x = L_y = 51.5\rho_0$, with a spatial resolution of 512×512 . The time step size is set to $2^{-9} \frac{L}{C_0}$ and the time interval between each output step $\Delta t = 0.25$. We take 50 TOKAM2D simulations with different random initial perturbations in n and ϕ , such that each run will produce a completely different pair of trajectories $(E(t), U(t))$. The linear stage and early nonlinear stage are discarded. We then collect the data from $t = 500$ to $t = 3000$ with a time interval of output $\Delta t = 0.25$, in total 10000 steps per simulation.

We first discretise the entire 2D (E, U) domain into equally spaced boxes. All states falling within the same box are considered to share the same preceding state (E, U) , which is represented by the centre of the corresponding box, illustrated in the left figure of FIG. 3. The right figure shows the distribution of states (E, U) in each

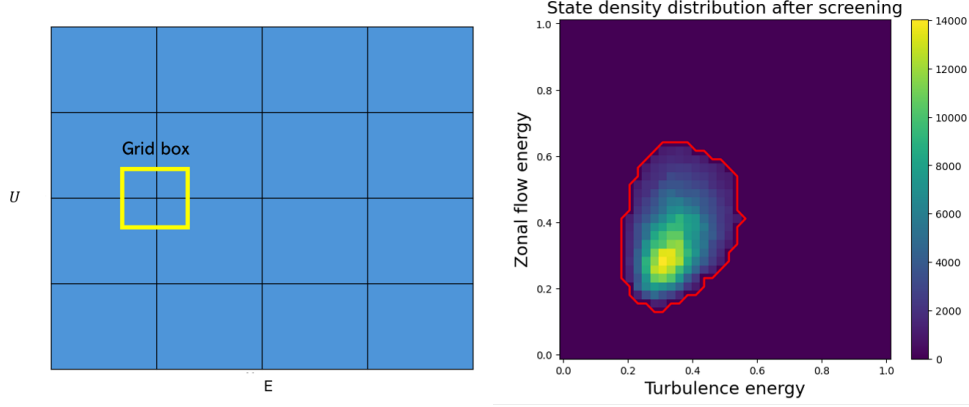


FIG. 3. Left: Illustration of grid buckets: the (E, U) space are discretised into equally spaced grid points. The yellow box illustrates a grid buckets centred on one of these grid points, within which all data are assumed to originate from the same grid box. Right: The number of pairs (E, U) in each grid box. The red contour shows the boundary where the number of data in the grid box exceeds 500.

grid box (histogram). We only make use of the boxes where the number of states exceeds a threshold of 500. After sorting data into buckets, we can compute $\Delta E(E(t), U(t)) = E(t + \Delta t) - E(t)$ by taking the difference between the neighbouring states in time, and similarly for ΔU . Due to the stochasticity, $\Delta E(E, U)$ is not unique, but rather has a Gaussian distribution around a mean value μ with a standard deviation of σ , proportional to g_{11} and g_{21} , respectively. We compute μ and σ from all the data in the same box and use the result to train our neural network. The full theory, procedure and details of training can be found in Huang *et al.* [13].

Motivated by physics, instead of letting g_{11} and g_{12} be arbitrary functions of (E, U) , we constrain their forms to follow

$$\begin{aligned} g_{11}(E, U, \theta_1) &= E f_{11}(U, \theta_1) + A E^2, \\ g_{12}(E, U, \theta_2) &= -B U + E U f_{12}(U, \theta_2), \end{aligned} \quad (8)$$

where $f_{11}(U, \theta_1)$ and $f_{12}(U, \theta_2)$ are two neural networks taking only the zonal flow energy U as input with their network parameters θ_1 and θ_2 , respectively. The constants A and B are learnable parameters, where the coefficient A represents the nonlinear damping of turbulence, while B corresponds to the linear damping rate of the zonal flow. The additional U in front of f_{12} is to ensure $g_{12} \rightarrow 0$ when $U \rightarrow 0$. The diffusion terms g_{21} and g_{22} are assumed to be proportional to E when $E \rightarrow 0$. Adding these physical constraints yields a better result compared to a non-constrained model.

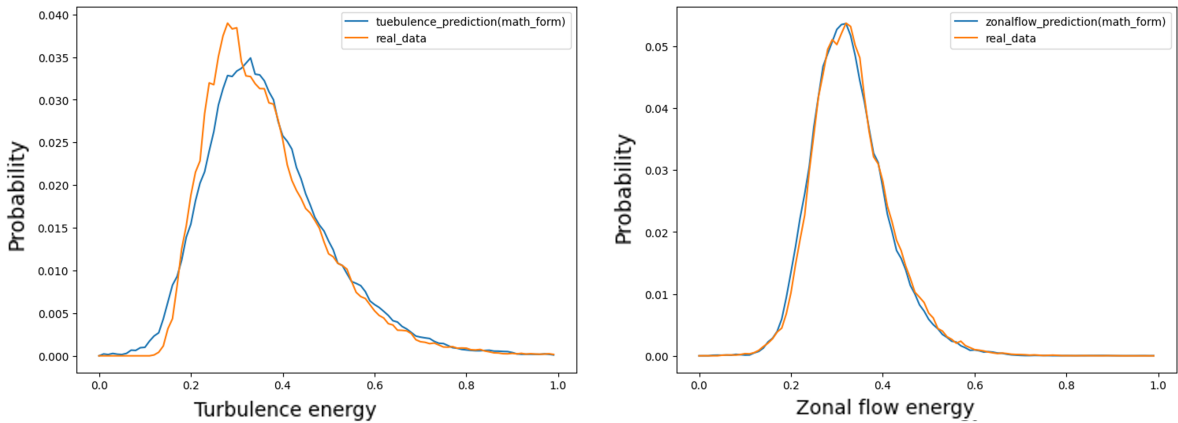


FIG. 4. Predicted states distribution from the learned NN model with physics constraints, compared to the distribution from TOKAM2D data.

We train the neural networks using the data μ and σ of each box. After that, we fit the output of f_{11} and f_{12} from neural networks into quadratic functions for a better extrapolation into the region with insufficient data. Then we generate 50 simulations of 10000 time steps based on the fitted quadratic functions. The state distribution shows a good match with the data, as shown in FIG. 4. In addition, by observing a single trajectory as illustrated

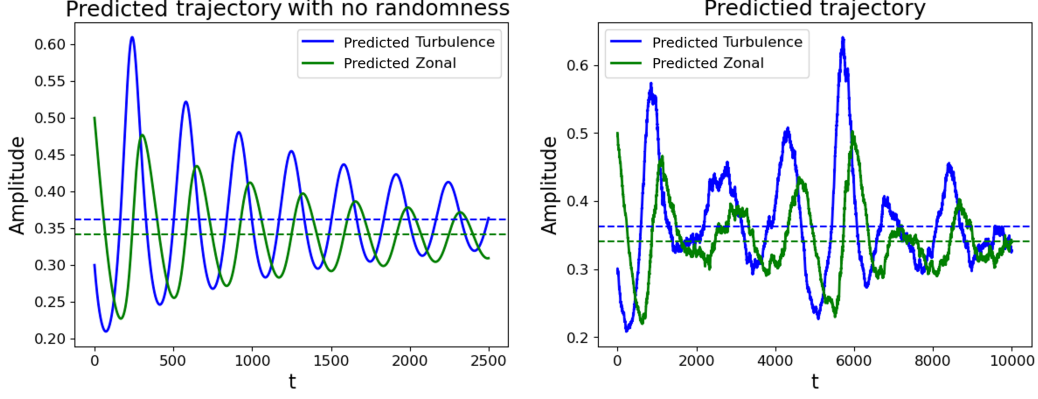


FIG. 5. Simulations by integrating the learned SDE with physics constraints when the stochasticity term is disabled (left) and enabled (right). The dashed lines indicates the averaged level for E and U in each simulation.

in FIG. 5, the predator-prey oscillation damps in the absence of randomness, and the stagnation phenomenon is clearly visible when randomness is introduced.

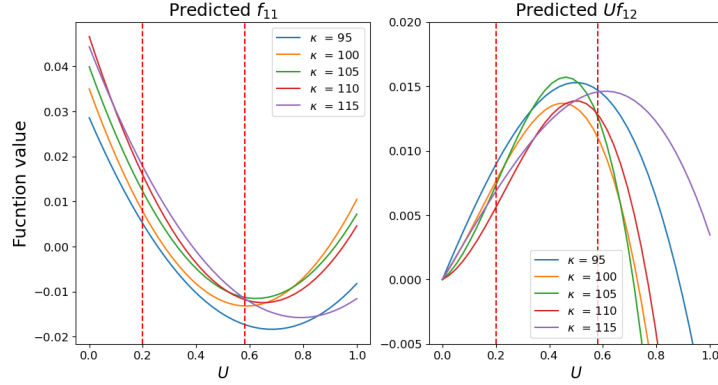


FIG. 6. The learned function f_{11} and f_{12} with respect to κ ($K = 100\kappa$), where vertical red dot line enclose the region with sufficient data.

Taking the same process, we performed a parameter scan over the density gradient coefficient κ from 0.95 to 1.15 with a step size of 0.05, and extracted the unknown functions, shown in FIG. 6. The function f_{11} is the effective growth rate of turbulence considering the presence of zonal flow. Turbulence is effectively suppressed by zonal flow since f_{11} decreases as U increases, until the growth rate becomes negative for $U > 0.35$. We find that for large U , the value of f_{11} is saturated around -0.01 , indicating limit in efficiency to suppress turbulence. Furthermore, while the shape of f_{11} remains largely unchanged across different values of κ in the range $[0.95, 1.15]$, its overall value shifts up according to the increased linear growth rate. To some extent, it is consistent with a physical facts that the strength of the turbulence–zonal flow interaction depends on the intrinsic properties of the turbulence and zonal flow rather than on the linear growth rate [22]. This presents the first quantitative confirmation of the theoretical prediction that the efficiency of zonal flow shearing effects on turbulence decreases as the zonal flow energy increases. The physical meaning of Uf_{12} is the efficiency of turbulence in producing Reynold stress, which then drive the zonal flow. Similarly, the efficiency decreases with stronger zonal flow. It is observed that f_{11} and Uf_{12} exhibit a near mirror symmetry, flipping upside down, which is consistent with energy conservation principles. However, the amplitude of Uf_{12} is noticeably smaller than that of f_{11} .

5. CONCLUSION

This work introduced two neural network–based reduced models for plasma turbulence, addressing both subgrid closure and turbulence–zonal flow dynamics. For drift-wave turbulence, the Extended HW-Closure model identified by physics-informed neural networks successfully reproduced transport fluxes and spectral features of high-resolution simulations on coarse grids, achieving more than an order-of-magnitude reduction in computational

cost. For turbulence–zonal flow interactions, a stochastic surrogate model learned from modified HW simulations captured predator–prey oscillations, saturation mechanisms, and stochastic variability, providing a compact but physically interpretable representation of the dynamics.

Together, these results highlight the potential of machine learning in constructing reduced turbulence models that combine efficiency with physical fidelity. Future work will extend these methods to more complicated fluid and/or gyrokinetic systems and explore integration into transport solvers, with the ultimate goal of enabling reliable and efficient predictions of confinement and performance in future fusion devices.

ACKNOWLEDGEMENTS

The authors would like to thank to Dr. Liu Jiao, Dr. Chin Chun Ooi and Dr. Jian Cheng Wong for their technical assistance and insightful suggestions. The authors would also like to acknowledge the Singapore Alliance with France for Fusion Energy (SAFE) team for providing access to the TOKAM2D code, which was essential for the numerical simulations carried out in this study, and would also like to acknowledge the organisers of the Festival de Théorie 2025 for providing a discussion environment that greatly benefited this work. This work is partly funded by National Research Foundation Singapore (NRF) core funding “Fusion Science for Clean Energy” and Ministry of Education (MOE) AcRF Tier 1 grants RS02/23 and RG156/23. The computational work for this article was partially performed on resources of the National Supercomputing Centre (NSCC), Singapore.

REFERENCES

- [1] W. Horton, *Reviews of Modern Physics* **71**, 735 (1999).
- [2] P. H. Diamond, S.-I. Itoh, K. Itoh, T. S. Hahm, *Plasma Physics and Controlled Fusion* **47**, R35 (2005).
- [3] X. Garbet, Y. Idomura, L. Villard, T. H. Watanabe, *Nuclear Fusion* **50** (2010).
- [4] M. Raissi, P. Perdikaris, G. Karniadakis, *Journal of Computational Physics* **378**, 686 (2019).
- [5] J. Degraeve, *et al.*, *Nature* **602**, 414 (2022).
- [6] J. Kates-Harbeck, A. Svyatkovskiy, W. Tang, *Nature* **568**, 526 (2019). Publisher: Springer US.
- [7] J. Seo, *et al.*, *Nature* **626**, 746 (2024).
- [8] O. Meneghini, *et al.*, *Nuclear Fusion* **57** (2017). Publisher: IOP Publishing.
- [9] K. L. Van De Plassche, *et al.*, *Physics of Plasmas* **27** (2020). ArXiv: 1911.05617 Publisher: American Institute of Physics Inc.
- [10] C. Wan, *et al.*, *Nuclear Fusion* **65**, 054001 (2025).
- [11] A. Hasegawa, M. Wakatani, *Physical Review Letters* **50**, 682 (1983).
- [12] K. Li, *et al.*, Reconstructing high-fidelity plasma turbulence with data-driven tuning of diffusion in low resolution grids arxiv:2509.11576 (2025).
- [13] J. C. Huang, *et al.*, Extracting a stochastic model for predator-prey dynamic of turbulence and zonal flows with limited data arxiv:2508.10408 (2025).
- [14] A. Hasegawa, K. Mima, *Phys. Rev. Lett.* **39**, 205 (1977).
- [15] P. Ghendrih, *et al.*, *Journal of Physics: Conference Series* **1125**, 012011 (2018).
- [16] R. Kraichnan, *Journal of Fluid Mechanics* **5**, 497 (1959).
- [17] F.Y. Gang, P.H. Diamond, J.A. Crotinger, A.E. Koniges, *Physics of Fluids B* **3**, 955 (1991).
- [18] Ö.D. Gürçan, P.H. Diamond, T.S. Hahm, *Physics of Plasmas* **13**, 052306 (2006).
- [19] X. Shi, *et al.*, *Advances in neural information processing systems* **28** (2015).
- [20] P. Fuller, E.-j. Kim, R. Hollerbach, B. Hnat, *Physics of Plasmas* **31**, 092506 (2024).
- [21] E.-j. Kim, R. Hollerbach, *Phys. Rev. Res.* **2**, 023077 (2020).
- [22] P. H. Diamond, Y.-M. Liang, B. A. Carreras, P. W. Terry, *Physical Review Letters* **72**, 2565 (1994).