# Supplementary Information
# Weak-topological neural operators for certificate-guided physical prediction

This document provides the benchmark protocols, certificate definitions, finite-selector derivation and reproducibility details that support the main Article.

## 1 Benchmark and section overview

WTNO is evaluated as a finite-section selection framework. A section is a validated pairing of a state predictor and a weak/topological certificate. The reported manuscript sections are:

- topology-changing 2D phase-field selection against feedback-FNO references;
- official PDEBench Burgers robust Huber selection with mollified BV/entropy certificates;
- internal 3D phase-field derived-surface and sparse-support stability diagnosis;
- public PDEBench 3D Turb one-step full64 validation against matched U-Net and strong-FNO controls.

The four settings instantiate a shared certificate-guided section-selection principle with benchmark-specific weak or geometric objects.

## 2 Phase-field protocol

Inputs contain initial phase and source or forcing channels; targets contain final phase and physical-response fields. Evaluation reports relative state error, rho-channel error, physical-response error, merge-region physical error, topology, varifold and merge-event certificates. Product-CWTS pairs phase/geometric selectors with physical-response selectors, and the trust-region variant starts from strong feedback-FNO references. A section is promoted only when validation state and certificate criteria remain inside the admissible region.

Source-data rows for this experiment are indexed under figure=Fig2 in the accompanying source-data table.

## 3 Burgers protocol

The Burgers track uses the official PDEBench viscous Burgers data. The pointwise state is evaluated by relative $u$ error. Weak geometry is evaluated through mollified BV density $|u_x|$ and entropy density $\nu|u_x|^2$, with shape and CDF diagnostics at the selected smoothing scale. Robust Huber training defines the selected section family. Reported summaries include event conditioned relative error, CVaR90, total variation and weak-measure CDF distance. Unselected terminal states are retained as drift controls.

Source-data rows for this experiment are indexed under figure=Fig3 in the accompanying source-data table.

# 4 Internal 3D phase-field protocol

The internal 3D track evaluates diffuse-interface volumetric states. Surface geometry is computed as a derived shell or surface measure from the predicted rho field. Deeper physical sections are checked for sparse-rho support collapse using a diagnostic rho-collapse threshold of 0.20. Support-weighted rho losses and channel-balanced losses are treated as certificate-gated variants. The Extended Data matched-split distribution reports all matched splits used for the stability interpretation.

Source-data rows for this experiment are indexed under figure=Fig4 in the accompanying source-data table. Extended Data rows are provided in a separate matched-split source-data table.

# 5 Public PDEBench 3D Turb protocol

The public 3D track uses the official PDEBench 3D Turb HDF5 tensor with Vx, Vy, Vz, density and pressure channels on 600 samples, 21 time points and $64^3$ spatial grids. The reported task predicts the final state from the penultimate time point. Public comparisons use matched input-output definitions, matched splits and matched metric computation across residual-convolutional, U-Net and strong-FNO sections.

The high-data public comparison uses 256/64/64 train/validation/test samples, 32 epochs and the same full64 one-step task for all compared section families. The matched high-data sections are residual-convolutional width24/depth5, U-Net width24/depth5 and strong-FNO width32/modes12. Reported metrics include rel-all, velocity, density, pressure, vorticity, enstrophy, gradient-density, gradient-pressure and kinetic-mass diagnostics. The public 3D conclusion is bounded to this one-step protocol.

Source-data rows for this experiment are indexed under figure=Fig5 in the accompanying source-data table.

# 6 Weak-topological selection theory

This section expands the finite-selector statement used in the main manuscript. It is a validation-selection result for a finite family of trained sections, not a claim that every neural architecture converges under stochastic optimisation.

## 6.1 Weak-topological state space

Let $Q$ be a compact space-time cylinder. A topology-changing prediction is represented by a weak-topological state

$$\mathfrak{z} = (\mu, V, \Pi) \in \mathcal{M}(Q; \mathbb{R}^m) \times \mathcal{M}_+(Q \times \mathbb{P}^{d-1}) \times \mathcal{M}_+(\Delta_R),$$

where $\mu$ is the physical state measure, $V$ is an interface varifold and $\Pi$ is a topological-persistence measure. On total-variation-bounded sets, each measure component is metrised by a countable separating family of test functions,

$$d_Y(\nu_1,\nu_2) = \sum_{j\ge 1} 2^{-j} \min\{1, |\langle \nu_1-\nu_2, \psi_j\rangle|\}.$$

The WTS metric is the product metric over the components of $\mathfrak{Z}$. Because $Q$, $Q \times \mathbb{P}^{d-1}$ and $\Delta_R$ are compact and the measure balls are total-variation bounded, these weak metrics metrise weak-star convergence on the bounded sets used by the selector.

## 6.2 Certificate instances

Each experimental regime instantiates the abstract certificate distance $d_\Gamma$ with a weak or geometric object:

- Phase-field topology-changing dynamics use topology, varifold and merge-event certificates derived from the diffuse phase field.
- Burgers dynamics use mollified BV density $|u_x|$ and entropy density $\nu|u_x|^2$, compared as weak measures through shape and CDF diagnostics.
- Internal 3D phase-field dynamics use derived shell and surface measures from the predicted $\rho$ field, together with a sparse-support diagnostic.
- Public 3D turbulence uses vorticity, enstrophy, scalar-gradient and kinetic-mass diagnostics computed from the predicted full-grid state.

These are not separate theories. They are separate choices of certificate map $\Gamma$ and certificate distance $d_\Gamma$ inside the same finite selection rule.

## 6.3 Finite admissible selector

Let $\mathcal{S} = \{s_1, \dots, s_M\}$ be the finite section family being compared under a fixed benchmark protocol. For each section, define a state risk $R_{\text{state}}(s)$ and a certificate risk $d_\Gamma(s)$. The population admissible set is

$$\mathcal{S}_\tau^\star = \{s \in \mathcal{S} : d_\Gamma(s) \le \tau_\Gamma,\ R_{\text{state}}(s) \le \tau_s\}.$$

The empirical admissible set replaces these risks by validation estimates. WTNO then selects

$$\widehat{s} = \arg\min_{s\in\widehat{\mathcal{S}}_\tau} \widehat{R}_{\text{state}}(s).$$

The rule is deliberately finite. It selects among trained sections recorded in the source-data package and does not claim to solve a continuous architecture search problem.

## 6.4 Oracle inequality

Assume:

- the target solution map is continuous into the WTS metric on a compact input family;
- the section family approximates the relevant state and certificate moments up to $\epsilon_{\text{app}}$;
- validation estimates of state and certificate risks are uniformly concentrated over $M$ sections;

- the admissible thresholds are fixed independently of the test set.

Then, with high probability, the empirical selector satisfies

$$R(\hat{s}) \leq \inf_{s \in \mathcal{S}_\tau^\star} R(s) + C\sqrt{\frac{\log M}{|V_{\mathrm{val}}|}} + \epsilon_{\mathrm{app}}.$$

Sketch of proof. The weak metric representation makes the state and certificate moments continuous on bounded compact sets. Finite neural heads or finite trained sections approximate these moment maps up to $\epsilon_{\mathrm{app}}$. Uniform concentration over $M$ candidate sections controls the difference between empirical and population state and certificate risks. A union bound gives the $\sqrt{\log M/|V_{\mathrm{val}}|}$ term. On the high-probability event where all validation estimates are close to their population values, the empirical minimiser over the admissible set cannot exceed the best population-admissible section by more than the concentration and approximation terms.

### 6.5 Event-conditioned limitation

If a certificate or metric is defined only on an event subset, the effective validation size is the number of validation events, not the total validation size. The same argument gives a concentration term of the form $\sqrt{\log M/N_{\mathrm{event}}}$. If $N_{\mathrm{event}} = 0$, no validation selector can distinguish event-specific sections uniformly. This is why WTNO uses confidence-gated event promotion and why non-event validation alone is not sufficient for topology-changing or rare-tail claims.

## 7 Figure and compilation reproducibility

The accompanying reproducibility package contains the source-data tables, visual assets and the following figure-generation scripts:

- make_fig1_framework.py
- make_fig1_phasefield.py
- make_fig2_burgers.py
- make_fig3_internal3d.py
- make_extended_internal3d_failure_distribution.py
- make_fig5_public3d.py

The phase-field, Burgers and internal-3D script names retain historical numbering for compatibility, but their outputs correspond to Figs. 2, 3 and 4. The mechanism script generates Fig. 1, and the public-3D script generates Fig. 5. The main LaTeX source, BibTeX database and embedded figure files are provided in the editable manuscript-source ZIP.