EMBEDDED NEURAL BOUNDARY: LEARNING NEAR-WALL IMPULSE FOR FAST FLOW PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Learned correctors for coarse-grid flow solvers are typically *volumetric*: a large CNN corrects every cell and is trained by unrolling through a differentiable solver. Classical simulation concentrates its effort within a few cells of the wall; we embed the learned component there, as a *neural boundary*. For coarse embedded-boundary flow past a sphere, an impulse confined, before the pressure projection, to a three-cell near-wall band removes $\sim 60\%$ of the near-wall velocity error, and most of what remains is static per operating condition. We shape the corrector accordingly: a small MLP predicts eight POD band coefficients per frame from the near-wall state and inflow condition, and a static bias field absorbs the rest. Its network adds 1% to the solver's wall-clock on a 24-thread CPU, against 190% for a U-Net. In closed loop it has 1.3–$2.1\times$ lower near-wall error than the best of five training variants of a volumetric solver-in-the-loop CNN with $177\times$ its trainable parameters ($21\times$ counting fixed tables), and, with its impulse counted as body force, drag within 2–4% of the fine reference; the static field raises the margin to $2.6\times$ and recovers 87–96% of a replayed full-field oracle's improvement. The same construction holds on 13 two-dimensional geometries and a spheroid at incidence for held-out Reynolds numbers.

1 INTRODUCTION

Fluid simulation underpins engineering design, climate prediction and graphics (Rhie & Chow, 1983; Taylor et al., 2012; Kochkov et al., 2024; Kim et al., 2019; Xie et al., 2018). Learned correctors make coarser grids usable: a classical solver integrates the resolved dynamics while a network supplies what the discretisation misses (Ghosal, 1996; Bar-Sinai et al., 2019; Um et al., 2020; Kochkov et al., 2021). The dominant design pays twice for this. The corrector is *volumetric*: a plain CNN, ResNet, dilated ResNet or local neural closure (Um et al., 2020; Kochkov et al., 2021; Stachenfeld et al., 2022; Sirignano & MacArt, 2022) writes a correction into every cell, so a full-volume network runs beside every solver step. And it is trained *through the loop*: a network fitted to single steps meets, in closed loop, the states its own corrections create, and errors compound (Tompson et al., 2017; Kim et al., 2019), so it is unrolled inside a differentiable solver (Um et al., 2020; Kochkov et al., 2021; List et al., 2022; Sirignano & MacArt, 2022) or given surrogates for that exposure: input noise, pushforward losses, rollout retraining, unrolling without solver gradients (Sanchez-Gonzalez et al., 2020; Brandstetter et al., 2023; Ross et al., 2011; List et al., 2025); plain one-step supervision, the *non-interacting* regime, Um et al. (2020) find weakest. Both expenses descend from a decision made before the error has been located: to correct *everywhere*.

Classical simulation knows where to spend: boundary-layer theory (Schlichting & Gersten, 2017) places separation and the wall force in a thin layer at the solid, and embedded-boundary solvers keep the bulk on a plain grid and pay for accuracy only in the cells the solid cuts: variational and cut-cell coupling (Batty et al., 2007; Ng et al., 2009), boundary re-meshing (Azevedo et al., 2016) and virtual-element cells (Beirão da Veiga et al., 2013; Tao et al., 2022). Learned correctors have not followed; we put the *neural* component there: an *embedded neural boundary*.

For coarse embedded-boundary flow past bluff bodies an impulse confined to a thin near-wall band, where the cut-cell boundary misplaces the separating shear layers, removes most of the near-wall error, and most of what it leaves in the wake is *static*; both are almost constant per operating condition, so the learned stage need only supply that dependence (Section 3).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

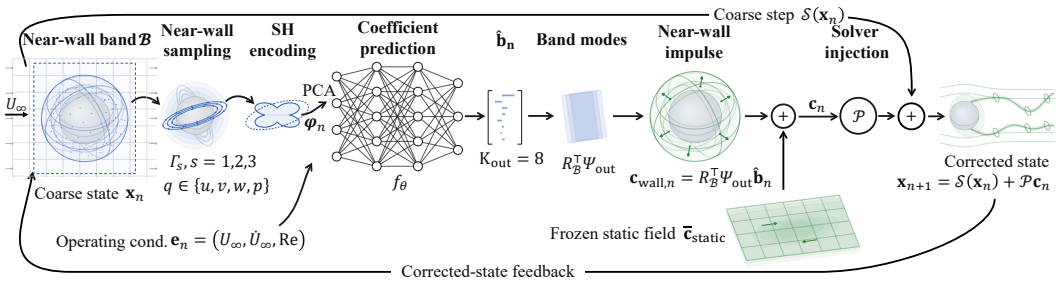


Figure 1: **Inference.** Shell samples of the coarse state, encoded and whitened to ϕ_n, feed the frozen f_θ with $\mathbf{e}_n$; the band modes spread its eight coefficients into a near-wall impulse, the static field is added, and the solver injects $\mathbf{c}_n$ through $\mathcal{P}$ (1).

We therefore shape the corrector like the error (Section 2): band-restricted defects give a 32-mode POD basis, the near-wall state is compressed with spherical harmonics on three shells, and a two-layer memoryless MLP predicts the leading eight coefficients from those features and the inflow condition; every online operation scales with the body surface, and a windowed average of reference defects supplies the static field. In our experiments these priors supplied what unrolled training is usually needed for, and one-step training sufficed: band localisation keeps mistakes out of the far field and conditioning absorbs the operating-condition shift; the solver's projection slot, shared by every method we compare, keeps any injected increment divergence-free and free of net momentum.

Our contributions are (i) an *error anatomy* for coarse embedded-boundary flow, how much of the coarse error a near-wall impulse can correct, exposed by nested near-wall metrics and a consistent force error; (ii) an *embedded neural boundary shaped like that anatomy*: a frozen memoryless MLP emitting inflow-conditioned band impulses, plus a windowed static bias field; and (iii) a *complete closed-loop system* that, against a 1.16M-parameter volumetric solver-in-the-loop U-Net (Um et al., 2020) in five training variants, from one-step to unrolled without solver gradients (Figure 3), has 1.3–2.1× lower near-wall error with 177× fewer trainable parameters and a learned stage costing 1.2 ms instead of 25 ms per frame on an H100 (6 versus 880 ms on a 24-thread CPU; full pipeline in Table 2); with the static field the margin reaches 2.6×, the uncorrected near-wall error falls 3–6× and 87–96% of a replayed full-field oracle's improvement is recovered, while controls trace the gap to the near-wall actuation prior rather than to the backbone. The same construction, trained and validated separately on each of 13 two-dimensional geometries and on a 6:1 spheroid at incidence, improves every body 2–5× at Reynolds numbers held out within the training range, and fails on an inflow programme outside it.

2 Method

Section 2.1 fixes the setting, the notation and the goal. Section 2.2 describes what runs at every coarse step and why the correction is eight coefficients of a near-wall impulse rather than a field (Figure 1). Section 2.3 describes how the band modes and the training targets are prepared from reference data (Figure 2), and Section 2.4 how the corrector is trained on them. Section 2.5 adds the static bias field and explains how the two stages are composed.

2.1 Setting and notation

Our goal is to run a low-resolution solver and obtain the dynamics of a high-resolution one. For a fixed body, a network is shown the correct high-resolution flow at a few operating conditions; afterwards the coarse solver, with the network in its loop and *no* further access to high-resolution data, should reproduce those dynamics, also at operating conditions it was not trained on. The operating condition is the inflow state $\mathbf{e}_n = (U_\infty, \dot{U}_\infty, \mathrm{Re})$, joined by the angle of attack on the spheroid suite. We pursue this by correcting the coarse flow field at every step: after each coarse step an *impulse*, a velocity increment, is added to the coarse state so that the corrected state lands where the fine solution, filtered to the coarse grid, would be. Concretely, the coarse solver is a cut-cell embedded-boundary code on a uniform grid of spacing Δ ($D/\Delta = 8$ cells per body diameter), and $\mathcal{S}$ advances its state $\mathbf{x}_n$ by one frame $\delta t = 0.1\, D/U_\infty$; the target $\mathbf{x}_n^*$ is the fine simulation ($D/\Delta = 32$, Appendix I) filtered to that grid. The impulse is a cell-centred velocity field $\mathbf{c}_n$ and

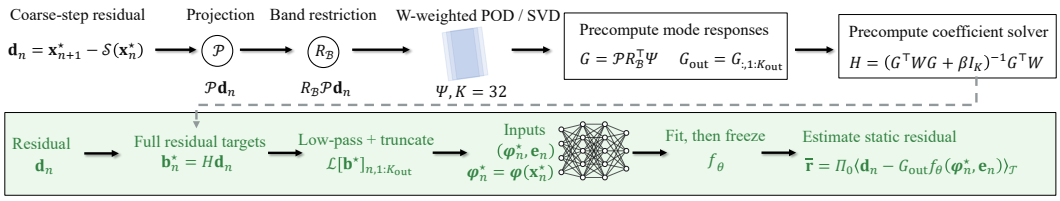


Figure 2: **Training.** The defect $\mathbf{d}_n$ is pushed through $\mathcal{P}$, band-restricted and reduced by a W-weighted POD to $K = 32$ modes, giving G, H of (4). Targets $\mathbf{b}^*_n$ are low-passed, truncated to 8 and fitted one-step (5); the static field is the frozen network's residual mean (6).

enters through one slot, the solver's own end-of-step face interpolation and pressure projection $\mathcal{P}$, which is linear and keeps the state discretely divergence-free:

$$\mathbf{x}_{n+1} = \mathcal{S}(\mathbf{x}_n) + \mathcal{P}\,\mathbf{c}_n, \qquad \mathbf{d}_n = \mathbf{x}^*_{n+1} - \mathcal{S}(\mathbf{x}^*_n). \tag{1}$$

$\mathbf{d}_n$, the one-step *defect*, is the impulse that would carry a coarse step taken from the reference exactly onto the reference; injecting $\mathbf{c}_n = \mathbf{d}_n$ is the *full-field oracle* (`vol`). Oracles are computed along the reference and replayed open-loop, so they serve as reference points rather than as bounds. Our impulses instead come from a network that sees only the coarse state and are supported only in a thin *band* $\mathcal{B}$ at the wall, the shell of about three cells off the body ($r \le 0.905$ for the sphere of radius 0.5), 10^{-3} of the domain's volume, where the shear layers separate and the force is generated; $R_\mathcal{B}$ restricts a field to the band and $R_\mathcal{B}^\top$ embeds a band field in the domain. The locality is that of $\mathbf{c}_n$ before projection: $\mathcal{P}$ is global, so the velocity response $\mathcal{P}\mathbf{c}_n$ extends beyond the band. Three further symbols recur. W is the diagonal matrix of fluid-volume fractions, with $\|\mathbf{v}\|^2_W = \mathbf{v}^\top W\mathbf{v}$ the fluid-weighted norm in which all fits and errors are measured; $\langle\cdot\rangle_\mathcal{T}$ is the time mean over a window $\mathcal{T}$; and Π_0 is the projector that subtracts from each velocity component its fluid-weighted mean.

2.2 Inference

This section describes the steps at inference time (Figure 1, Algorithm 2). In one closed-loop step the solver writes $\mathbf{x}_n$, the corrector returns $\mathbf{c}_n$, and the solver injects $\mathcal{P}\mathbf{c}_n$ (1). A grid-valued $\mathbf{c}_n$ has 3×128^3 unknowns per step; even an impulse confined to the band, $\mathbf{s} \in \mathbb{R}^{3|\mathcal{B}|}$, has $\sim$4,300. We reduce the output instead: $\mathbf{c}_n$ is supported in $\mathcal{B}$ and drawn from $K_{\text{out}} = 8$ fixed band modes $\Psi_{\text{out}} = [\psi_1, \ldots, \psi_{K_{\text{out}}}]$ (16 in 2D; Section 2.3). The injected correction is $G_{\text{out}}\hat{\mathbf{b}}_n$ with $G_{\text{out}} = \mathcal{P}R_\mathcal{B}^\top\Psi_{\text{out}}$, the solver's divergence-free response to the band impulse. Since $\mathcal{P}$ enforces the inflow and no-penetration conditions, the injected increment carries no net momentum whatever is fed to it, for our corrector and for every baseline alike; the pre-projection momentum $\hat{\mathbf{b}}^\top\mathbf{m}$, with $\mathbf{m}_k = \Delta^3\,\mathbf{1}^\top W\psi_k$, is a diagnostic here (its sign, towards higher drag, checks the fitted targets) and, booked to the body, the corrected system's physical drag (Appendix F). The input is reduced likewise, to shell features $\boldsymbol{\phi}$ of the near-wall state and the condition $\mathbf{e}_n$:

$$\hat{\mathbf{b}}_n = f_\theta\big(\boldsymbol{\phi}(\mathbf{x}_n), \mathbf{e}_n\big), \qquad \phi_{s,q,\ell m}(\mathbf{x}) = \oint_{\Gamma_s} q\,Y_{\ell m}\,d\Omega, \qquad \mathbf{c}_{\text{wall},n} = R_\mathcal{B}^\top\Psi_{\text{out}}\hat{\mathbf{b}}_n. \tag{2}$$

The features $\phi_{s,q,\ell m}$ are the coefficients of the four fields $q \in \{u, v, w, p\}$ in the real spherical harmonics $Y_{\ell m}$ on three shells Γ_s at radius $R + s\Delta$, $s \in \{1, 2, 3\}$: q is interpolated to a 16×32 Gauss–Legendre quadrature of each shell and integrated against $Y_{\ell m}$, $\ell \le 6$, giving 49 coefficients per field and shell, 588 in all. The network's first operation is a fixed PCA whitening fitted on the training frames,

$$\mathbf{z}_n = \operatorname{diag}(\boldsymbol{\sigma})^{-1}\,V\,\big(\boldsymbol{\phi}(\mathbf{x}_n) - \boldsymbol{\mu}\big), \qquad V \in \mathbb{R}^{24\times 588}, \tag{3}$$

with $\boldsymbol{\mu}$ the training mean, the rows of V the 24 leading principal directions (99.5% of the variance) and $\boldsymbol{\sigma}$ their standard deviations; $\mathbf{e}_n$ is standardised likewise. Band, shells and mask follow the signed distance to the body (Fourier modes on offset loops in 2D), so the construction carries over unchanged to each new body, on which basis, corrector and static field are retrained. The impulse $\mathbf{c}_{\text{wall},n}$ is the learned part of the correction. A second, non-learned part, a per-condition static bias field $\bar{\mathbf{c}}_{\text{static}}$ estimated beforehand (Section 2.5), is added to it every frame, so the correction applied at inference is $\mathbf{c}_n = \bar{\mathbf{c}}_{\text{static}} + \mathbf{c}_{\text{wall},n}$ (Figure 1, Algorithm 2); with $\bar{\mathbf{c}}_{\text{static}} = \mathbf{0}$ the loop runs the corrector alone (`mlp`).

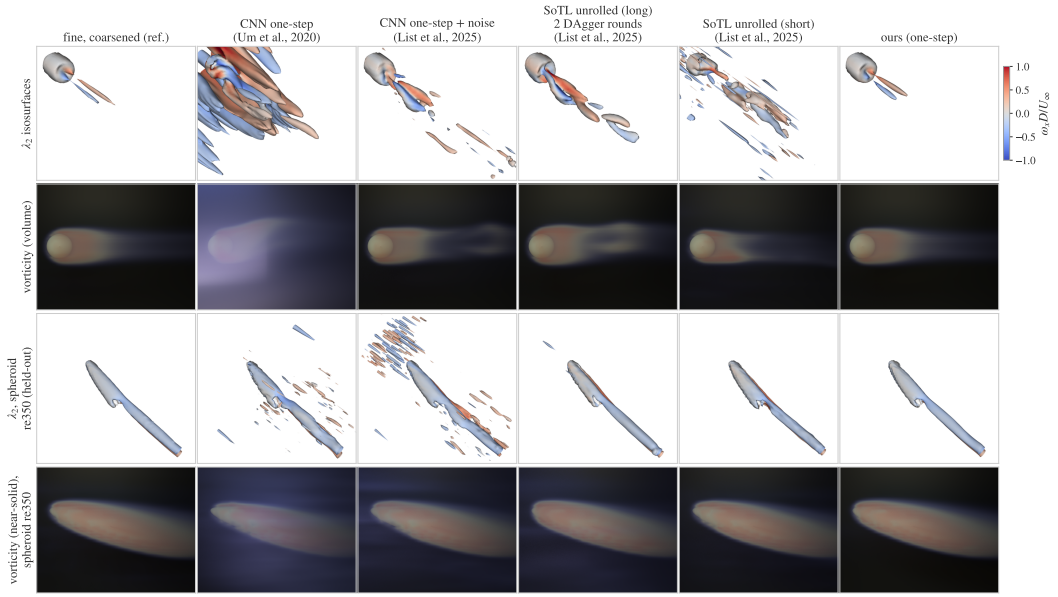


Figure 3: Wake structure deep in the closed loop (λ_2 isosurfaces, $|\boldsymbol{\omega}|$ renders): sphere at held-out $\mathrm{Re} = 250$ (rows 1–2), 6:1 spheroid at held-out $\mathrm{Re} = 350$ (rows 3–4). Columns: filtered reference; volumetric CNN (Um et al., 2020) trained one-step, with noise, unrolled (long, truncated); ours.

2.3 Training data preparation

This section describes how the training data are prepared (Figure 2, Algorithm 1). Along each training trajectory the coarse step is taken from the filtered reference and the defect $\mathbf{d}_n$ of (1) recorded. We restrict the impulse to a W-weighted POD of the band-restricted projected defects $R_{\mathcal{B}}\mathcal{P}\mathbf{d}_n$ over the training frames, $\Psi = [\psi_1, \ldots, \psi_K]$ ($K = 32$, $\geq 99.4\%$ of their energy on every suite, unit-W-norm columns). Because $\mathcal{P}$ is linear, the solver's response G to each mode is computed once, here; the response to any coefficient vector is then $G\mathbf{b}$ without a further projection, and the best coefficients per frame follow from a $K \times K$ solve

$$G = \mathcal{P}R_{\mathcal{B}}^{\top}\Psi, \qquad \mathbf{b}_n^* = \arg\min_{\mathbf{b}} \|G\mathbf{b} - \mathbf{d}_n\|_W^2 + \beta\|\mathbf{b}\|^2 = H\mathbf{d}_n, \quad H = (G^{\top}WG + \beta I_K)^{-1}G^{\top}W \tag{4}$$

($\beta = 10^{-6}$). Injecting $\mathbf{c}_n = R_{\mathcal{B}}^{\top}\Psi\mathbf{b}_n^*$ is the *POD-restricted band oracle* (`opt`), optimal within $\mathrm{span}\,\Psi$; it reaches 77% of the one-step reduction of the free whole-band solve (Appendix C).

2.4 Training

At inference time the reference is unavailable, so the corrector (2) must predict $\mathbf{b}_n^*$ from the coarse state. Its targets are the coefficient series $\mathbf{b}_n^*$ of each training trajectory, processed in two ways. A zero-phase low-pass $\mathcal{L}$ with cut-off $f_c = U_\infty/D$ removes the frame-to-frame jitter that the cut-cell geometry imprints on $\mathbf{b}_n^*$ while keeping the shedding dynamics ($\mathrm{St} \approx 0.13$–0.2). Only the leading $K_{\mathrm{out}} = 8$ coefficients are kept ($\Psi_{\mathrm{out}} = \Psi_{:,1:K_{\mathrm{out}}}$, 97.8% of the band-POD energy), the rest contributing little. f_θ is then fitted *one-step*, on features of reference states $\mathbf{x}_n^*$ rather than of its own closed-loop states (no unrolling), the non-interacting regime that Um et al. (2020) found weakest:

$$\min_{\theta} \sum_{n \in \mathrm{train}} \left\| f_\theta\big(\boldsymbol{\phi}(\mathbf{x}_n^*), \mathbf{e}_n\big) - \mathcal{L}[\mathbf{b}^*]_{n,1:K_{\mathrm{out}}} \right\|^2. \tag{5}$$

2.5 Static bias field and composition

The static bias field is the second stage (`volstatic`) and is not learned: it is the momentum-free average of the one-step defects of a reference trajectory over a window $\mathcal{T}$ of 40–80 time units, $\bar{\mathbf{d}} = \Pi_0\langle\mathbf{d}_n\rangle_{\mathcal{T}}$, injected unchanged at every frame. On a training condition, $\mathcal{T}$ is the stretch of that trajectory just before the scored window. At a held-out Reynolds number between two training

Algorithm 1 TRAINING: basis, corrector, static field

Input: $\{(\mathbf{x}_n^*, \mathbf{e}_n)\}_{n\in\text{train}}, \mathcal{S}, \mathcal{P}, W, R_{\mathcal{B}};\ K, K_{\text{out}}, \beta, \mathcal{L}, \mathcal{T}$
Output: $\mathcal{A} = (\Psi_{\text{out}}, \theta, \bar{\mathbf{c}}_{\text{static}})$

$\mathbf{d}_n \leftarrow \mathbf{x}^*_{n+1} - \mathcal{S}(\mathbf{x}^*_n)$ ▷ Eq. (1)
$\Psi \leftarrow \text{POD}(\{R_{\mathcal{B}}\mathcal{P}\mathbf{d}_n\}_n,\ W,\ K)$
$G \leftarrow \mathcal{P}R_{\mathcal{B}}^\top\Psi,\ H \leftarrow (G^\top W G + \beta I_K)^{-1}G^\top W$ ▷ Eq. (4)
$\mathbf{b}^*_n \leftarrow H\mathbf{d}_n$ ▷ opt
$\boldsymbol{\phi}_n \leftarrow \text{SHELLFEATURES}(\mathbf{x}^*_n)$ ▷ Eq. (2)
$\theta \leftarrow \text{FIT}(\{\boldsymbol{\phi}_n, \mathbf{e}_n\},\ \mathcal{L}[\mathbf{b}^*]_{1:K_{\text{out}}})$ ▷ Eq. (5)
$\Psi_{\text{out}} \leftarrow \Psi_{:,1:K_{\text{out}}},\ G_{\text{out}} \leftarrow \mathcal{P}R_{\mathcal{B}}^\top\Psi_{\text{out}}$ ▷ mlp
$\bar{\mathbf{c}}_{\text{static}} \leftarrow \Pi_0\big\langle \mathbf{d}_n - G_{\text{out}} f_\theta(\boldsymbol{\phi}_n, \mathbf{e}_n)\big\rangle_{\mathcal{T}}$ ▷ Eq. (6)
return $\mathcal{A}$

Algorithm 2 INFERENCE: closed loop

Input: $\mathcal{A}, \mathcal{S}, \mathcal{P}, \mathbf{e}(t), \mathbf{x}_0, N$
Output: $\{\mathbf{x}_n\}_{n=1}^N$

for $n = 0, \ldots, N-1$ **do**
  $\boldsymbol{\phi}_n \leftarrow \text{SHELLFEATURES}(\mathbf{x}_n)$
  $\hat{\mathbf{b}}_n \leftarrow f_\theta(\boldsymbol{\phi}_n, \mathbf{e}_n)$ ▷ Eq. (2)
  $\mathbf{c}_n \leftarrow \bar{\mathbf{c}}_{\text{static}} + R_{\mathcal{B}}^\top\Psi_{\text{out}}\hat{\mathbf{b}}_n$
  $\mathbf{x}_{n+1} \leftarrow \mathcal{S}(\mathbf{x}_n) + \mathcal{P}\,\mathbf{c}_n$ ▷ Eq. (1)
end for
return $\{\mathbf{x}_n\}$

conditions we average the fields of the two neighbours. Where no training condition lies on either side (the decelerating inflow) the corrector runs alone (Appendix B).

When the two stages are combined, $\mathbf{c}_n = \bar{\mathbf{c}}_{\text{static}} + \mathbf{c}_{\text{wall},n}$, the static field should not repeat what the corrector already supplies. We therefore freeze the trained corrector first and take the static field to be the mean of what it leaves uncorrected on the reference frames of the window:

$$\bar{\mathbf{c}}_{\text{static}} = \bar{\mathbf{r}} = \Pi_0\big\langle \mathbf{d}_n - G_{\text{out}} f_\theta(\boldsymbol{\phi}(\mathbf{x}^*_n), \mathbf{e}_n)\big\rangle_{\mathcal{T}}, \qquad \theta \text{ frozen.} \tag{6}$$

The corrector's per-frame output is untouched; only the constant field depends on the window. Adding the two stages estimated independently (`mlpvol`) is an ablation (Appendix B).

3 RESULTS

3.1 SETUP AND EVALUATION

Sphere suite. Fine references are level-9 octree simulations of flow past a sphere, a bluff-body benchmark (Johnson & Patel, 1999) (Appendix I), filtered to the coarse 128^3 grid. The corrector is trained on 16 trajectories spanning $\text{Re} = 100$–500 with steady, chirped, sinusoidal and decelerated inflow. Closed-loop runs start from the filtered reference at $t_0 = 135$, are scored over $t \in [135, 150]$ (rollouts to $t = 235$: Appendix G), and test **re300** (in-distribution, steady), **re300-decel** (inflow decelerates $U{=}1 \to 0.3$ before the window) and the held-out **re350**, **re250**, **re450**, on which every deployed component is interpolated from neighbouring training Re. Repeated runs differ by 1.3%; differences within this noise floor are ties.

2D and spheroid suites. The pipeline is run separately on each of (i) 13 two-dimensional geometries (cylinder (Rajani et al., 2009), square, tandem pairs at $L/D \in \{1.5, 2, 3, 4, 5\}$, NACA0012 at $\alpha \in \{0, 5, 10, 15, 20\}^\circ$, an inclined plate; eight Reynolds numbers each ($\text{Re} = 58.6$–625); basis, corrector and static field trained per geometry on six, with the interior $\text{Re} = 156.3$ and 312.5 held out of everything) and (ii) a 6:1 prolate spheroid whose angle of attack joins the conditioning vector: six steady training trajectories ($\text{Re} \in \{200, 250, 300, 400\}$ at $\alpha = 15^\circ$, $\alpha \in \{10, 20\}^\circ$ at $\text{Re} = 300$), held-out re350 and re300-decel; the training set has *no* dynamic inflow, so re300-decel probes deployment *outside* the trained range. This suite deploys the corrector alone (Appendix A).

Modes and metrics. Besides the correctors and oracles of Section 2 and the CNN variants of Section 3.4 we report `none` (uncorrected solver), `mean` (unconditioned static band mean) and the oracle `volmean` (evaluation-window mean defect); oracles are italicised. Field metrics are the fluid-volume-weighted RMS velocity error against the filtered reference, window-averaged and inflow-scaled, on nested masks **glob** (full domain), **near** ($r < 1.5$) and **surf** (the band $r \le 0.905$); **glob** dilutes near-wall error by three orders of magnitude. Drag is scored by the *consistent* error $\mathrm{d}C_d = \bar{C}_d^{\text{run}} - \bar{C}_d^{\text{ref}}$ against the filtered reference under the coarse force operator (positive: under-correction; Appendix A). It scores the field, not the physical force: the coarse operator under-reads the fine drag by 0.12 on any field, and Appendix F reports the error against the fine drag raw, with the corrector's impulse counted as body force, and calibrated. $|C_l|$, the window-mean transverse force, reads out shedding activity. *Improvement fractions* are shares of the replayed full-field oracle's **surf** reduction: ratios to a reference point, not to a bound.

Table 1: **Main result.** Band error **surf** ($\times 10^{-2}$) and consistent drag error, sphere tests. Blocks: CNN baselines; our band prior on the MLP corrector, alone or with the static field from bracketing training Re (†: own window); the same prior on the U-Net (App. E). Bold: best deployable (excl. †). ‡: uses the decelerated trajectory's own reference window; with the steady Re 300 window instead, the static band mean fails (19.87, Table 15).

	re300 (in-dist.)		re300-decel		re350 (h.-out)		re250 (h.-out)		re450 (h.-out)	
mode	surf↓	dCd→0	surf↓	dCd→0	surf↓	dCd→0	surf↓	dCd→0	surf↓	dCd→0
`none` (no correction)	12.41	+.005	2.35	+.021	13.83	−.010	10.55	+.027	16.64	−.051
Volumetric solver-in-the-loop CNN (Um et al., 2020)										
`sotl1s` (CNN one-step)	6.89	+.014	4.13	+.006	12.45	−.018	23.08	+.030	12.14	−.010
`sotlnoise` (CNN + noise)	5.19	−.008	4.34	+.017	7.86	−.012	4.54	−.026	8.05	−.006
`sotlur` (CNN unrolled, 1 rd.)	12.15	+.018	4.33	+.017	11.75	+.006	8.54	+.029	14.20	−.009
`sotlur2` (CNN unrolled, 2 rds.)	12.18	+.033	5.49	+.007	14.50	+.029	13.38	+.025	14.58	+.039
`sotlsur` (CNN unrolled, short)	10.68	−.009	2.47	+.007	12.62	**+.004**	10.68	**+.012**	13.52	**−.003**
Near-wall band prior (ours), MLP corrector										
`bandstatic` (cond.-matched band mean, no learning)	4.43	**+.003**	1.33‡	**+.002**	6.17	+.015	3.25	**−.012**	8.38	+.027
Ours (near-wall corrector)	4.90	**+.004**	1.15	**+.003**	5.38	+.007	3.38	**−.012**	7.87	+.018
Ours (+ static bias, bracketing training Re)	2.95	−.010	—	—	3.74	−.010	**1.72**	−.029	5.19	−.005
Ours (+ static bias, same trajectory)†	2.29	−.015	0.40	+.005	—	—	—	—	—	—
Near-wall band prior (ours), U-Net backbone (App. E)										
unrolled-short CNN, output projected on 8 band modes (post hoc)	**2.49**	−.027	**0.96**	+.005	**3.21**	−.041	2.27	−.033	**5.05**	−.030
`cnnband` (band-limited U-Net, retrained)	3.04	−.012	1.80	**+.002**	4.40	−.006	3.14	−.014	5.33	−.013
`cnnpod` (U-Net encoder → 8 band coefficients, retrained)	4.32	**+.003**	1.33	**+.002**	3.49	**−.003**	4.88	−.016	7.90	+.024
`vol` *(full-field oracle)*	*1.55*	*−.010*	*0.24*	*+.006*	*2.30*	*−.006*	*1.36*	*−.025*	*3.83*	*+.016*

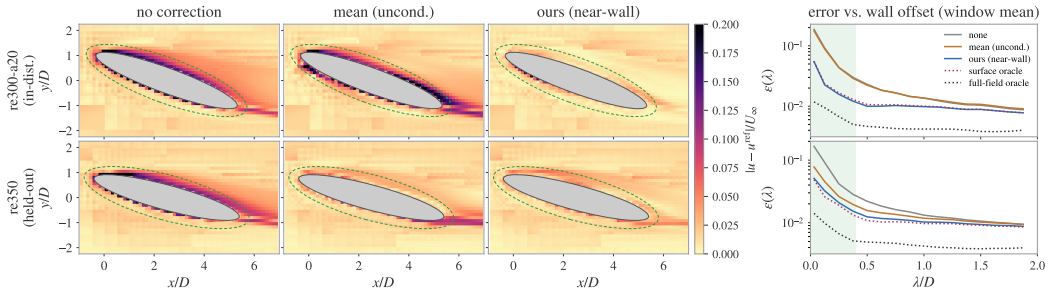


Figure 4: Spheroid suite: mid-plane velocity error $|u - u^{\mathrm{ref}}|/U_\infty$ in the loop (band edge dashed). The uncorrected error blankets the body; the unconditioned band mean fails at $\alpha = 20°$, the α-conditioned corrector clears it. Right: window-mean error vs. wall offset λ (band shaded).

3.2 Main result and error anatomy

Table 1 supports three readings. First, the corrector has lower near-wall error than every CNN variant on all five trajectories, by 1.3–2.1× on re300-decel, re350 and re250 and by 2–6% on re300 and re450; the static field from training conditions widens the gap to 1.6–2.6× on the steady trajectories (empty cells: Appendix B). Second, drag is the corrector's home game: it is generated in the band the corrector acts on, whose output follows the inflow condition, whereas volumetric corrections, oracles included, buy field accuracy with a systematic drag overshoot; on re300 and re300-decel its drag error is below the full-field oracle's, and its impulse (0.13–0.15) matches the 0.12 the coarse operator under-reads (Appendix F). Third, each baseline fails where conditioning predicts: the unconditioned band mean turns drag into thrust under deceleration (Table 15), every CNN variant is worse than no correction there, and the one-step CNN diverges at $\mathrm{Re} = 250$.

Error anatomy. The band corrector alone removes 51–68% of the raw near-wall error, an improvement fraction of 57–78%: that share is *correctable from* the three-cell band. The static field brings the fraction to 87–96% on the steady trajectories, so nearly all the band leaves behind is a condition-dependent *static* field; the remaining 4–13% is per-frame wake dynamics outside the band or beyond the 32 modes. Figure 4 shows the error fields behind the numbers (spheroid suite); the three metrics rank methods alike with very different contrast (re300, deployed: **glob** 1.7×, **near** 3.9×, **surf** 5.4×; wake topology in Appendix H).

Table 2: **Wall-clock cost**, re300 (2×48-core EPYC, 24 threads; one H100; CPU / GPU). Learned stage: on-device network compute. Full pipeline: adds host–device transfer and the solver's injection projection (breakdown: Appendix A).

	128^3 solver	+ corrector	+ U-Net	level-9 adaptive	uniform 256^3	uniform 512^3
learned stage, ms per frame (CPU / GPU)	—	**6 / 1.2**	**880 / 25**	—	—	—
learned stage only, s per time unit (relative)	4.7 (1×)	4.8 / 4.7 (1.01 / 1.00×)	13.5 / 5.0 (2.9 / 1.05×)	24 (5.1×)	42 (8.9×)	∼650 (∼140×)
full pipeline, s per time unit (relative)	—	6.8 / 6.9 (1.44 / 1.47×)	17.3 / 9.1 (3.7 / 1.95×)	—	—	—

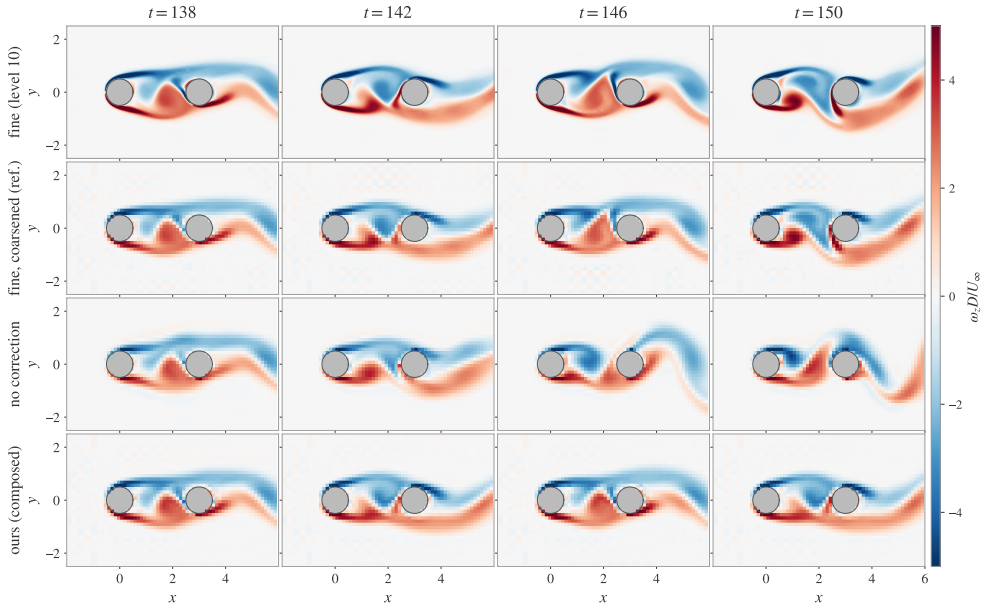


Figure 5: Held-out 2D deployment: tandem cylinders, L/D=3, Re = 312.5; vorticity $\omega_z D/U_\infty$, 3–15 time units into the loop. Rows: fine; filtered reference; uncorrected solver, which weakens the gap vortices and drifts out of phase; composed correction, which stays locked (other groups: Appendix L).

3.3 Static versus dynamic correction

The corrector's output splits into a per-condition mean and a state-dependent *fluctuation* (Appendix D). The fluctuation is small one-step, 2–6% of the output, and the optimal band coefficients are likewise a per-condition constant to 99.5–99.99% of their energy (Table 6); in closed loop, however, it is not negligible: alone it removes 14–21% of the sphere's near-wall error, because feedback amplifies what is invisible one-step, and 5% in 2D. The mean carries the rest. The condition-matched band mean (`bandstatic` in Tables 1 and 3) recovers most of the corrector's improvement with no learning, but it is a *calibration* that needs a fine reference at the deployment condition: under deceleration the steady-Re 300 band mean fails like the unconditioned one (19.87 vs. 1.15 for the corrector, Table 15). Retrained on the conditioning scalars alone, the corrector reproduces the band mean; on the flow features alone it degrades under deceleration (Table 7). The deployed corrector leads the band mean where Re interpolation is uneven (re350, re450) and, in 2D, on strongly shedding bodies (Table 20); the learned stage therefore contributes conditioning first and dynamics second. Drag accuracy and the restored shedding amplitude are likewise properties of the band, matched by the static band mean; the composed configuration's larger drag error is the static field's (Appendix D).

3.4 Baseline: a solver-in-the-loop volumetric CNN

The baseline is a volumetric solver-in-the-loop CNN (Um et al., 2020): a 1.16M-parameter 3D U-Net (Ronneberger et al., 2015) that predicts the full one-step defect from u, p and the broadcast conditioning scalars, trained on 1,280 reference pairs (Appendix A) in five variants: one-step supervision (`sotl1s`), input noise (`sotlnoise`), one and two rounds of DAgger-style unrolling (Ross et al., 2011) (`sotlur`, `sotlur2`) and truncated unrolling (`sotlsur`; Appendix A); unrolling

Table 3: **2D multi-geometry suite.** Averages over 13 geometries: 26 held-out-Re deployments ($\mathrm{Re} \in \{156.3, 312.5\}$) and 13 in-distribution ($\mathrm{Re} = 234.4$). **surf** $\times 10^{-2}$; $\overline{|\mathrm{d}C_d|}$ mean drag-error magnitude. Bold: best deployable per column (per group: Appendix K).

	held-out Re (26 runs)		in-dist. Re (13 runs)	
mode	surf ↓	$\overline{\lvert\mathrm{d}C_d\rvert}$ ↓	surf ↓	$\overline{\lvert\mathrm{d}C_d\rvert}$ ↓
`none` (no correction)	13.82	0.067	13.37	0.053
`mean` (uncond. static)	6.57	0.040	5.46	0.026
`sotl1s` (CNN one-step)	21.44	0.076	19.54	0.055
`sotlnoise` (CNN + noise)	17.91	0.052	12.99	0.049
`sotlur` (CNN unrolled, 1 rd.)	14.02	0.091	9.55	0.072
`sotlur2` (CNN unrolled, 2 rds.)	18.66	0.090	19.37	0.089
Ours (near-wall corrector, `mlp`)	4.91	0.030	4.52	**0.021**
`bandstatic` (cond.-matched band mean)	5.38	**0.024**	5.14	0.022
`volstatic` (static bias alone)	4.90	0.036	4.52	0.026
`mlpvol` (naive stack)	3.78	0.029	3.81	0.028
Ours (+ static bias)	**3.55**	0.034	**3.52**	0.036
`opt` (band oracle)	*4.38*	*0.030*	*4.16*	*0.025*
`vol` (full-field oracle)	*1.84*	*0.036*	*1.78*	*0.027*

Table 4: **Spheroid suite. surf** ($\times 10^{-2}$) and consistent drag error, corrector and best CNN retrained here; re300, a20 training, re350 held-out Re, decel held-out *inflow programme*. Bold: best deployable; full metrics in Table 19.

	re300 (in-dist.)		a20 (in-dist.)		re350 (h.-out Re)		decel (h.-out inflow)	
mode	surf ↓	dCd → 0	surf ↓	dCd → 0	surf ↓	dCd → 0	surf ↓	dCd → 0
`none` (no correction)	7.16	+.064	8.47	+.063	7.91	+.060	**1.29**	+.022
`mean` (uncond. static)	3.16	+.016	8.27	+.040	4.05	+.016	12.02	**−.005**
`sotlsur` (CNN unrolled, short)	2.44	+.018	3.36	+.006	3.27	+.014	10.04	+.072
Ours (near-wall corrector)	**2.28**	**+.004**	**2.62**	**−.001**	**2.96**	**+.001**	9.56	−.009
`opt` (band oracle)	*2.25*	*+.002*	*2.67*	*−.002*	*2.52*	*+.001*	*0.75*	*+.013*
`vol` (full-field oracle)	*0.79*	*+.010*	*0.83*	*−.007*	*0.92*	*+.005*	*0.18*	*+.017*

relabels self-generated states and retrains (gradient-free, List et al. 2025). Although the U-Net explains $\sim$93% of the one-step defect variance, no variant beats the corrector in closed loop, and under deceleration all five are 1.7–4$\times$ worse than no correction (cost: Table 2). At held-out Reynolds numbers the one-step variant diverges, noise only recovers the unconditioned-static level, and unrolling over-actuates or, truncated, keeps 1.7–3.2$\times$ the corrector's near-wall error and halves the shedding amplitude. The gap is set by where the network acts rather than by its backbone (Appendix E): projected post hoc on the corrector's band modes, the short-unroll U-Net has the lowest near-wall error of any deployable mode (Table 1) but 3–6$\times$ the corrector's drag error; retrained band-limited (`cnnband`) or as an encoder for the eight coefficients (`cnnpod`) it matches the corrector with 98–177$\times$ the trainable parameters.

3.5 Thirteen 2D geometries

On the 26 held-out-Re deployments (Table 3), each scored with the model trained on its own geometry, the sphere ranking repeats: the corrector alone cuts the near-wall error from 13.8 to 4.9×10^{-2}, the composed configuration reaches 3.6, and all 13 geometries improve, by 2.0–5.1$\times$ and 2.5–11$\times$ respectively (Appendix K, Figure 5). The per-group CNN, despite $\sim$94% one-step variance explained, is no better than the uncorrected solver on aggregate and loses to the corrector in all 13 groups. Steady inflow leaves naive stacking (`mlpvol`) unpunished.

3.6 A second 3D body and a new conditioning axis

On the steady spheroid conditions (Table 4) the corrector cuts the near-wall error 2.4–3.2$\times$, sits on its band oracle and collapses the coarse solver's $+0.06$ drag bias to $|\mathrm{d}C_d| \leq 0.006$ on the consistent metric. Against the fine drag the raw coarse-operator reading of the corrected run is farther off than the uncorrected one (-0.10 vs. -0.04): the operator under-reads the fine field by 0.10 and

the uncorrected solver's overshoot cancels part of it. With the corrector's impulse counted the error is $+0.014$ at the held-out Re, and calibrated at $\mathrm{Re} = 300$ alone -0.001 against the uncorrected $+0.070$ (Appendix F); the short-unroll CNN is competitive on the band metric but carries 3–$14\times$ the drag error, and the unconditioned band mean, pooled over a grid dominated by $\alpha = 15^\circ$, fails at 10° (Appendix H). The Reynolds axis, spanned by the training grid, interpolates; the inflow axis, absent from it, does not: on re300-decel the correction is worse than no correction while the band oracle stays sound.

4 Related work

Learned components in fluid simulation. Networks have replaced solver steps, synthesised detail and evolved latent dynamics (Ladický et al., 2015; Tompson et al., 2017; Xie et al., 2018; Kim et al., 2019; Thuerey et al., 2020; Vlachas et al., 2022; Kutz, 2017). We keep the solver and correct its coarse approximation, following learned discretisations and hybrid solvers (Bar-Sinai et al., 2019; Kochkov et al., 2021; Sirignano et al., 2020; Dresdner et al., 2023; Kochkov et al., 2024). Because one-step errors accumulate in closed loop (Tompson et al., 2017; Kim et al., 2019), such correctors are unrolled through differentiable solvers (Um et al., 2020; Kochkov et al., 2021; List et al., 2022; Sirignano & MacArt, 2022; Holl et al., 2020); List et al. (2025) find gradient-free unrolling recovers part of it. Fully learned simulators (Sanchez-Gonzalez et al., 2020; Pfaff et al., 2021; Ummenhofer et al., 2020; Li et al., 2021; Stachenfeld et al., 2022) face the same shift and meet it with statistical exposure (Brandstetter et al., 2023; Lippe et al., 2023; Kohl et al., 2024).

Closures, wall models, embedded boundaries. Turbulence closures (Ling et al., 2016; Beck et al., 2019; Duraisamy et al., 2019; Novati et al., 2021) and wall models (Bae & Koumoutsakos, 2022; Lozano-Durán & Bae, 2023) are almost all *instantaneous* regressors (Vadrot et al., 2023), as is ours; memory (Gupta & Lermusiaux, 2021; Boral et al., 2023) we replace by solver feedback. Embedded-boundary solvers (Peskin, 2002; Mittal & Iaccarino, 2005; Popinet, 2003) leave a cut-cell truncation error at $D/\Delta = 8$, which we correct. Cut-cell and virtual-element methods (Batty et al., 2007; Ng et al., 2009; Azevedo et al., 2016; Beirão da Veiga et al., 2013; Tao et al., 2022) shrink it by enriching the cut cells; we borrow their locality and learn the residual. Our POD (Berkooz et al., 1993; Rowley & Dawson, 2017) is of *corrections*, not states (Holden et al., 2019).

5 Conclusion

We presented an embedded neural boundary: a learned corrector for coarse embedded-boundary flow confined to a three-cell band at the solid. It rests on an error anatomy measured before learning: an impulse confined to that band removes $\sim 60\%$ of the near-wall error, and both it and most of the remainder are, per operating condition, almost static. A two-layer memoryless MLP emitting near-wall impulses, trained one-step and frozen, plus a per-condition static field, has lower near-wall error than a $177\times$-larger volumetric solver-in-the-loop CNN in five training variants; the CNN becomes competitive only once confined to the band. One-step training sufficed under the conditions tested; we do not claim this in general, and where a differentiable solver exists the two are complementary.

Limitations. Our method has several limitations. Each body has its own basis and corrector; cross-geometry transfer is untested. The static field needs a fine reference at or near the deployment condition and slightly over-corrects the drag when composed (Section 3.3). The coarse force operator under-reads the fine drag by 15–20% on the 3D bodies, whichever field it is applied to; the corrector's physical drag is recovered only when its impulse is counted as body force or a training-condition calibration is applied, and on 2D bluff bodies with sharp separation the impulse explains only part of that bias (Appendix F). The learned dynamics are 2–6% of the output and pay off only where the wake sheds; on the aerofoils a condition-matched band mean does better. The band prior presumes a boundary-generated error and caps recovery: on strong vortex streets the band oracle stays far below the full-field oracle. The evidence covers laminar flow ($\mathrm{Re} \leq 500$) past fixed bodies at $D/\Delta = 8$ in one solver; turbulence and moving boundaries remain untested.

AI USE STATEMENT

We used LLM-based assistants in two ways. First, to polish the writing and checks of compliance with the conference template. Second, as coding assistants: part of the implementation, evaluation, and plotting code was generated by LLM-based tools under the authors' direction and supervision. They were not used to generate research ideas, hypotheses, theory, proofs, or methods, and every claim and number in the paper originates from the authors' own work. All edited text and all generated code were reviewed by the authors, who take full responsibility for the content of this paper.

ETHICS STATEMENT

This work uses only synthetic flow data and involves no human subjects or personal data.

REPRODUCIBILITY STATEMENT

All closed-loop numbers are deterministic solver runs with every component fixed before the test trajectory is touched. Appendix A gives the full experimental protocol, and Appendices J–M report per-trajectory metrics and flow-field galleries. Code, trained assets and the scripts generating every table and figure will be released upon publication.

REFERENCES

Vinicius C. Azevedo, Christopher Batty, and Manuel M. Oliveira. Preserving geometry and topology for fluid flows with thin obstacles and narrow gaps. *ACM Transactions on Graphics*, 35(4):97:1–97:12, 2016. doi: 10.1145/2897824.2925919.

H. Jane Bae and Petros Koumoutsakos. Scientific multi-agent reinforcement learning for wall-models of turbulent flows. *Nature Communications*, 13(1), March 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-28957-7. URL http://dx.doi.org/10.1038/s41467-022-28957-7.

Yohai Bar-Sinai, Stephan Hoyer, Jason Hickey, and Michael P. Brenner. Learning data-driven discretizations for partial differential equations. *Proceedings of the National Academy of Sciences*, 116(31):15344–15349, July 2019. ISSN 1091-6490. doi: 10.1073/pnas.1814058116. URL http://dx.doi.org/10.1073/pnas.1814058116.

Christopher Batty, Florence Bertails, and Robert Bridson. A fast variational framework for accurate solid-fluid coupling. *ACM Transactions on Graphics*, 26(3):100, 2007. doi: 10.1145/1276377.1276502.

Andrea Beck, David Flad, and Claus-Dieter Munz. Deep neural networks for data-driven les closure models. *Journal of Computational Physics*, 398:108910, December 2019. ISSN 0021-9991. doi: 10.1016/j.jcp.2019.108910. URL http://dx.doi.org/10.1016/j.jcp.2019.108910.

Lourenço Beirão da Veiga, Franco Brezzi, Andrea Cangiani, Gianmarco Manzini, L. Donatella Marini, and Alessandro Russo. Basic principles of virtual element methods. *Mathematical Models and Methods in Applied Sciences*, 23(1):199–214, 2013. doi: 10.1142/S0218202512500492.

G Berkooz, P Holmes, and J L Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25(1):539–575, January 1993. ISSN 1545-4479. doi: 10.1146/annurev.fl.25.010193.002543. URL http://dx.doi.org/10.1146/annurev.fl.25.010193.002543.

Anudhyan Boral, Zhong Yi Wan, Leonardo Zepeda-Núñez, James Lottes, Qing Wang, Yi fan Chen, John Roberts Anderson, and Fei Sha. Neural ideal large eddy simulation: Modeling turbulence with neural stochastic differential equations, 2023. URL https://arxiv.org/abs/2306.01174.

Johannes Brandstetter, Daniel Worrall, and Max Welling. Message passing neural pde solvers, 2023. URL https://arxiv.org/abs/2202.03376.

Gideon Dresdner, Dmitrii Kochkov, Peter Norgaard, Leonardo Zepeda-Núñez, Jamie A. Smith, Michael P. Brenner, and Stephan Hoyer. Learning to correct spectral methods for simulating turbulent flows, 2023. URL https://arxiv.org/abs/2207.00556.

Karthik Duraisamy, Gianluca Iaccarino, and Heng Xiao. Turbulence modeling in the age of data. *Annual Review of Fluid Mechanics*, 51(1):357–377, January 2019. ISSN 1545-4479. doi: 10.1146/annurev-fluid-010518-040547. URL http://dx.doi.org/10.1146/annurev-fluid-010518-040547.

Sandip Ghosal. An analysis of numerical errors in large-eddy simulations of turbulence. *Journal of Computational Physics*, 125(1):187–206, 1996.

Abhinav Gupta and Pierre F. J. Lermusiaux. Neural closure models for dynamical systems, 2021. URL https://arxiv.org/abs/2012.13869.

Daniel Holden, Bang Chi Duong, Sayantan Datta, and Derek Nowrouzezahrai. Subspace neural physics: fast data-driven interactive simulation. In *Proceedings of the 18th annual ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '19, pp. 1–12. ACM, July 2019. doi: 10.1145/3309486.3340245. URL http://dx.doi.org/10.1145/3309486.3340245.

Philipp Holl, Vladlen Koltun, and Nils Thuerey. Learning to control pdes with differentiable physics. *International Conference on Learning Representations (ICLR)*, 2020.

T. A. Johnson and V. C. Patel. Flow past a sphere up to a Reynolds number of 300. *Journal of Fluid Mechanics*, 378:19–70, 1999.

Byungsoo Kim, Vinicius C. Azevedo, Nils Thuerey, Theodore Kim, Markus Gross, and Barbara Solenthaler. Deep fluids: A generative network for parameterized fluid simulations. *Computer Graphics Forum*, 2019. ISSN 1467-8659. doi: 10.1111/cgf.13619.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980 [cs]*, December 2014.

Dmitrii Kochkov, Jamie A. Smith, Ayya Alieva, Qing Wang, Michael P. Brenner, and Stephan Hoyer. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences*, 118(21), May 2021. ISSN 1091-6490. doi: 10.1073/pnas.2101784118. URL http://dx.doi.org/10.1073/pnas.2101784118.

Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Griffin Mooers, Milan Klöwer, James Lottes, Stephan Rasp, Peter Düben, Sam Hatfield, Peter Battaglia, Alvaro Sanchez-Gonzalez, Matthew Willson, Michael P. Brenner, and Stephan Hoyer. Neural general circulation models for weather and climate. *Nature*, 632(8027):1060–1066, July 2024. ISSN 1476-4687. doi: 10.1038/s41586-024-07744-y. URL http://dx.doi.org/10.1038/s41586-024-07744-y.

Georg Kohl, Li-Wei Chen, and Nils Thuerey. Benchmarking autoregressive conditional diffusion models for turbulent flow simulation, 2024. URL https://arxiv.org/abs/2309.01745.

J. Nathan Kutz. Deep learning in fluid dynamics. *Journal of Fluid Mechanics*, 814:1–4, 2017. doi: 10.1017/jfm.2016.803.

Lubor Ladický, SoHyeon Jeong, Barbara Solenthaler, Marc Pollefeys, and Markus Gross. Data-driven fluid simulations using regression forests. *ACM Trans. Graph.*, 34(6):199:1–199:9, October 2015. ISSN 0730-0301. doi: 10.1145/2816795.2818129.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations, 2021. URL https://arxiv.org/abs/2010.08895.

Julia Ling, Andrew Kurzawski, and Jeremy Templeton. Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166, October 2016. ISSN 1469-7645. doi: 10.1017/jfm.2016.615. URL http://dx.doi.org/10.1017/jfm.2016.615.

Phillip Lippe, Bastiaan S. Veeling, Paris Perdikaris, Richard E. Turner, and Johannes Brandstetter. Pde-refiner: Achieving accurate long rollouts with neural pde solvers, 2023. URL https://arxiv.org/abs/2308.05732.

Bjoern List, Li-Wei Chen, Kartik Bali, and Nils Thuerey. Differentiability in unrolled training of neural physics simulators on transient dynamics. *Computer Methods in Applied Mechanics and Engineering*, 433:117441, 2025.

Björn List, Li-Wei Chen, and Nils Thuerey. Learned turbulence modelling with differentiable fluid solvers: Physics-based loss-functions and optimisation horizons, 2022. URL https://arxiv.org/abs/2202.06988.

Adrián Lozano-Durán and H. Jane Bae. Machine learning building-block-flow wall model for large-eddy simulation, 2023. URL https://arxiv.org/abs/2211.07879.

Rajat Mittal and Gianluca Iaccarino. Immersed boundary methods. *Annual Review of Fluid Mechanics*, 37(1):239–261, January 2005. ISSN 1545-4479. doi: 10.1146/annurev.fluid.37.061903.175743. URL http://dx.doi.org/10.1146/annurev.fluid.37.061903.175743.

Yen Ting Ng, Chohong Min, and Frédéric Gibou. An efficient fluid–solid coupling algorithm for single-phase flows. *Journal of Computational Physics*, 228(23):8807–8829, 2009. doi: 10.1016/j.jcp.2009.08.032.

Guido Novati, Hugues Lascombes de Laroussilhe, and Petros Koumoutsakos. Automating turbulence modelling by multi-agent reinforcement learning. *Nature Machine Intelligence*, 3(1):87–96, January 2021. ISSN 2522-5839. doi: 10.1038/s42256-020-00272-0. URL http://dx.doi.org/10.1038/s42256-020-00272-0.

Charles S. Peskin. The immersed boundary method. *Acta Numerica*, 11:479–517, January 2002. ISSN 1474-0508. doi: 10.1017/s0962492902000077. URL http://dx.doi.org/10.1017/s0962492902000077.

Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W. Battaglia. Learning mesh-based simulation with graph networks, 2021. URL https://arxiv.org/abs/2010.03409.

Stéphane Popinet. Gerris: a tree-based adaptive solver for the incompressible euler equations in complex geometries. *Journal of Computational Physics*, 190(2):572–600, Sept 2003. ISSN 0021-9991. doi: 10.1016/s0021-9991(03)00298-5. URL http://dx.doi.org/10.1016/S0021-9991(03)00298-5.

BN Rajani, A Kandasamy, and Sekhar Majumdar. Numerical simulation of laminar flow past a circular cylinder. *Applied Mathematical Modelling*, 33(3):1228–1247, 2009.

C. M. Rhie and W. L. Chow. Numerical study of the turbulent flow past an airfoil with trailing edge separation. *AIAA Journal*, 21(11):1525–1532, 1983. doi: 10.2514/3.8284.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241. Springer, 2015.

Stéphane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 627–635, 2011.

Clarence W. Rowley and Scott T.M. Dawson. Model reduction for flow analysis and control. *Annual Review of Fluid Mechanics*, 49(1):387–417, January 2017. ISSN 1545-4479. doi: 10.1146/annurev-fluid-010816-060042. URL http://dx.doi.org/10.1146/annurev-fluid-010816-060042.

Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter W. Battaglia. Learning to simulate complex physics with graph networks, 2020. URL https://arxiv.org/abs/2002.09405.

Hermann Schlichting and Klaus Gersten. *Boundary-Layer Theory*. Springer, Berlin, Heidelberg, 9 edition, 2017. doi: 10.1007/978-3-662-52919-5.

Justin Sirignano and Jonathan F. MacArt. Deep learning closure models for large-eddy simulation of flows around bluff bodies, 2022. URL https://arxiv.org/abs/2208.03498.

Justin Sirignano, Jonathan F. MacArt, and Jonathan B. Freund. Dpm: A deep learning pde augmentation method with application to large-eddy simulation. *Journal of Computational Physics*, 423:109811, December 2020. ISSN 0021-9991. doi: 10.1016/j.jcp.2020.109811. URL http://dx.doi.org/10.1016/j.jcp.2020.109811.

Kimberly Stachenfeld, Drummond B. Fielding, Dmitrii Kochkov, Miles Cranmer, Tobias Pfaff, Jonathan Godwin, Can Cui, Shirley Ho, Peter Battaglia, and Alvaro Sanchez-Gonzalez. Learned coarse models for efficient turbulence simulation, 2022. URL https://arxiv.org/abs/2112.15275.

Michael Tao, Christopher Batty, Mirela Ben-Chen, Eugene Fiume, and David I. W. Levin. VEMPIC: Particle-in-polyhedron fluid simulation for intricate solid boundaries. *ACM Transactions on Graphics*, 41(4):115:1–115:22, 2022. doi: 10.1145/3528223.3530138.

Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012.

Nils Thuerey, Konstantin Weißenow, Lukas Prantl, and Xiangyu Hu. Deep learning methods for reynolds-averaged navier–stokes simulations of airfoil flows. *AIAA Journal*, 58(1):25–36, 2020.

Jonathan Tompson, Kristofer Schlachter, Pablo Sprechmann, and Ken Perlin. Accelerating Eulerian fluid simulation with convolutional networks. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 3424–3433. PMLR, 06–11 Aug 2017. URL https://proceedings.mlr.press/v70/tompson17a.html.

Kiwon Um, Robert Brand, Yun (Raymond) Fei, Philipp Holl, and Nils Thuerey. Solver-in-the-loop: Learning from differentiable physics to interact with iterative PDE-solvers. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6111–6122, 2020.

Benjamin Ummenhofer, Lukas Prantl, Nils Thuerey, and Vladlen Koltun. Lagrangian fluid simulation with continuous convolutions. In *International Conference on Learning Representations*, 2020.

Aurélien Vadrot, Xiang I. A. Yang, and Mahdi Abkar. Survey of machine learning wall models for large eddy simulation, 2023. URL https://arxiv.org/abs/2211.03614.

Pantelis R. Vlachas, Georgios Arampatzis, Caroline Uhler, and Petros Koumoutsakos. Multiscale simulations of complex systems by learning their effective dynamics. *Nature Machine Intelligence*, 4(4):359–366, April 2022. ISSN 2522-5839. doi: 10.1038/s42256-022-00464-w. URL http://dx.doi.org/10.1038/s42256-022-00464-w.

You Xie, Erik Franz, Mengyu Chu, and Nils Thuerey. tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow. *ACM Transactions on Graphics (TOG)*, 37(4):1–15, 2018.

APPENDIX

A Experimental details **14**
A.1 Sphere suite . . . 14
A.2 Parameter count . . . 15
A.3 2D multi-geometry suite . . . 15
A.4 Spheroid suite . . . 15
A.5 Solver-in-the-loop CNN baseline . . . 15
A.6 Computational cost . . . 15

B Static field: composition and provenance **16**

C POD-restricted versus free band optimum **16**

D Static versus dynamic correction **17**
D.1 Ablations . . . 17
D.2 One-step decomposition . . . 17
D.3 Closed loop, sphere . . . 17
D.4 Closed loop, 2D . . . 18
D.5 Drag under composition . . . 18

E CNN controls that isolate the near-wall prior **18**
E.1 Post-hoc restriction . . . 18
E.2 Retrained controls . . . 18

F Physical drag: coarse-operator consistency versus the fine reference **20**

G Long-horizon rollouts **21**
G.1 Protocol . . . 22
G.2 Results . . . 22

H Additional 3D figures **23**

I Grid convergence of the fine reference **24**

J Full per-trajectory metrics **24**
J.1 Sphere suite . . . 24
J.2 Spheroid suite . . . 25

K 2D multi-geometry study: per-group metrics **25**

L 2D multi-geometry study: full flow-field gallery **27**

M 2D multi-geometry study: solver-in-the-loop CNN flow fields **32**

A EXPERIMENTAL DETAILS

A.1 SPHERE SUITE

The sphere has diameter 1 in a 16^3 domain. Of the sixteen training trajectories (Section 3.1) thirteen run for 150 time units and three for 60 with the sphere offset from the grid, the offset joining $\mathbf{e}_n$ as a fourth conditioning scalar. Every consecutive frame pair over the first 88% of each trajectory ($18{,}584$ frames) is a training sample; the last 12% is a validation tail used only to select architecture

and seed by one-step MSE. The band POD carries 87% of the energy in mode 1, and the band oracle removes 31–42% of the near-field one-step defect on every training trajectory.

A.2 Parameter count

The deployed corrector is a $28 \to 64 \to 64 \to 8$ $\tanh$ network with $6{,}536$ trainable weights. Stored with it are the fitted but non-trainable tables of (2) and (3): PCA whitening, the eight POD modes on the $1{,}432$ band cells and the normalisation constants, $55{,}645$ coefficients in all; the shell-sampling stencils and the harmonic projector are geometric tables, not fitted. The volumetric U-Net has $1{,}159{,}299$ trainable parameters, $177\times$ the corrector's weights and $21\times$ its stored coefficients; `cnnpod` has $639{,}992$ ($98\times$). Ratios in the main text count trainable weights.

A.3 2D multi-geometry suite

Grids: cylinder, square and tandem pairs at $D/\Delta = 8$ on 128^2; NACA0012 at $c/\Delta = 16$ on 256^2; the 3%-thick flat plate at $\alpha = 20^\circ$ on 512^2; fine references at octree levels 10–11, $t \in [0, 150]$. Band, near-field mask and features (loops offset $\{1, 2, 3\}\Delta$ from the surface, Fourier modes $m \le 8$) follow the signed distance; the corrector has two $\tanh$ layers of width 64. $K = 32$ modes carry at least 99.8% of the band-POD energy, and the band oracle removes 32–46% of the near-field one-step defect, except 14% on the inclined plate, whose defect sits in the shed shear layers beyond the band. Evaluation mirrors the sphere protocol ($t_0 = 135$, window $[135, 150]$). One caveat: on the tandem L/D=3 pair at $\mathrm{Re} = 312.5$ the *uncorrected* solver falls into a wrong gap-flow attractor ($2.8\times$ the reference lift fluctuation), and the corrected loop, pulled back to the correct regime, overdamps.

A.4 Spheroid suite

The spheroid (thickness $D = 1$) is pitched nose-up at angle of attack α, with level-9 references filtered to the same 128^3 grid. Replacing the sphere's radial offset $r - R$ by the semi-axis offset $\lambda(x)$ carries band, mask and feature shells over verbatim. POD energy and band-oracle reduction (35–40% steady, 29% decelerated) match the sphere. No static stage is estimated on this suite.

A.5 Solver-in-the-loop CNN baseline

The U-Net of Section 3.4 runs on the full 128^3 grid with seven input channels (u_x, u_y, u_z, p and the broadcast scalars $U_\infty, \dot{U}_\infty, \mathrm{Re}$); its output is normalised by the training-set defect RMS and masked to the fluid. Its 1,280 raw full-field pairs $(\mathbf{x}_n^*, \mathbf{d}_n)$ carry no low-pass: 80 per training trajectory at 0.5 spacing over the 40 time units before the evaluation window, 10% held out for validation, a subset of the corrector's training frames at 0.1 spacing. `sotl1s`: one-step supervision. `sotlnoise`: Gaussian input noise on the velocity channels ($\sigma = 1\times$ the defect RMS, selected on validation from $\{0.5, 1, 2\}$). `sotlur`: the one-step network is rolled out over the training windows, the self-generated states are labelled by pullback against the filtered reference through the slot, and the network is retrained on the union; `sotlur2` repeats this once. `sotlsur`: the same relabelling applied to 3–8-step rollouts from every reference state. All variants use Adam (Kingma & Ba, 2014) at 2×10^{-4} with cosine decay, batch 4, at most 300 epochs and early stopping. At the selected epoch the training loss is below the validation loss, so the network is neither capacity- nor epoch-limited. The 2D port trains one 0.39M-parameter U-Net per geometry.

A.6 Computational cost

Table 2 was measured with exclusive use of the node (AMD EPYC 7R13), on re300 from $t_0 = 135$ with ten frames per time unit and two solver steps per frame; the uncorrected 128^3 loop takes 0.47 s per frame. The corrector's learned stage on CPU splits into 1.7 ms of shell sampling and harmonic projection, $10\,\mu$s of MLP and 4.7 ms of POD spread; adding the static field and writing the band correction cost a further 33 ms, and host–device transfer of the four fields 21 ms. The U-Net forward pass costs 65 ms on the H100 with transfers. The solver-side injection, one face interpolation and pressure projection, costs 0.20 s for the band correction and 0.38 s for the full-field defect; the full-pipeline row of Table 2 sums these pieces for a persistent runtime, whereas the prototype as run launches a fresh process per frame (0.15 s for Python/NumPy, 5.5 s for torch/CUDA) and measures

8.2 and 63 s per time unit. Training the corrector takes about 20 s per model on four CPU threads; the U-Net takes 2.3–3.2 h on the H100 for the one-step variants and 7–10 h for the unrolled ones, excluding rollout generation.

B STATIC FIELD: COMPOSITION AND PROVENANCE

Stacking the two stages estimated independently (`mlpvol`) is worse than the static stage alone and doubles the drag bias (Table 14).

Under the held-out protocol of Section 2.5, re300 ("Re 200/400 field" rows of Table 14) moves from 2.29 to 2.95×10^{-2} **surf**, $93 \to 87\%$ of the oracle's improvement. The loss is in the transverse components, which the planar-symmetric wake orients differently on each trajectory (own-window and interpolated fields correlate at 0.91 streamwise but below 0.15 transversely).

Under deceleration nothing brackets the deployment. The steady $\mathrm{Re} = 300$ static field ("steady Re 300" rows of Table 15; **surf** 20.3×10^{-2}, $\mathrm{d}C_d = -0.142$) fails like the steady band mean, and even as a residual on the frozen corrector the transplanted field is worse than no correction (3.27 vs. 2.35), so the corrector alone is deployed there; a same-trajectory window (†) would improve **surf** $6.5\times$.

C POD-RESTRICTED VERSUS FREE BAND OPTIMUM

The band oracle `opt` of (4) is optimal only within the $K = 32$ POD modes of $R_{\mathcal{B}}\mathcal{P}\mathbf{d}_n$; over the whole band the free minimiser is $\mathbf{s}_n^* = (A^\top W A)^\dagger A^\top W \mathbf{d}_n$ with $A = \mathcal{P} R_{\mathcal{B}}^\top$. To price the restriction we formed A explicitly (one projection per band cell and component, 4,296 columns, once offline; $\mathrm{cond}(A^\top W A) = 2.2 \times 10^3$, no null directions) and solved for 30 evaluation-window frames per sphere trajectory; the free solution is insensitive to regularisation (residuals change by $< 0.01 \times 10^{-3}$ for pseudo-inverse truncation 10^{-3}–10^{-8} or ridge $\leq 10^{-4}$).

Table 5: **One-step near-field residual of band-limited optima** ($r < 1.5$, $\times 10^{-3}$, mean over 30 evaluation-window frames; in parentheses the fraction of the uncorrected defect removed). K rows: optimum in the first K band-POD modes ($K = 8$: the corrector's output space; $K = 32$: `opt`). *Free*: unrestricted least squares on all 4,296 band unknowns. *Captured*: share of the free reduction reached by $K = 32$. *In span* Ψ_{32}: W-energy share of $\mathbf{s}_n^*$ in the 32-mode span. Bottom block: basis from the POD of the *other four* trajectories' free optima.

	re300	re300-decel	re350	re250	re450
`none`	11.09	3.28	11.28	10.89	12.03
$K = 8$	6.78 (38.9%)	3.08 (6.2%)	6.93 (38.5%)	6.57 (39.7%)	7.33 (39.0%)
$K = 32$ (`opt`)	6.53 (41.1%)	3.02 (8.0%)	6.68 (40.8%)	6.40 (41.3%)	7.13 (40.7%)
$K = 128$	6.32 (43.0%)	2.91 (11.5%)	6.44 (42.9%)	6.18 (43.3%)	6.92 (42.5%)
Free (4,296)	5.20 (53.1%)	2.80 (14.7%)	5.28 (53.2%)	5.03 (53.8%)	5.73 (52.4%)
Captured by $K = 32$	77%	54%	77%	77%	78%
$\mathbf{s}^*$ in span Ψ_{32}	29%	29%	28%	26%	29%
Captured, **glob**	84%	56%	83%	83%	84%
Leave-one-out POD of the free optima					
$K = 8$	5.41 (51.2%)	3.05 (7.0%)	5.52 (51.1%)	5.38 (50.6%)	6.20 (48.5%)
$K = 32$	5.32 (52.0%)	3.02 (7.9%)	5.45 (51.7%)	5.28 (51.5%)	6.09 (49.4%)
Captured by $K = 8$	97%	48%	96%	94%	92%

The restriction costs about a quarter (Table 5): the free optimum removes 52–54% of the near-field one-step defect at steady inflow against the 32-mode oracle's 41%, so `opt` reaches 77% of what the band can do (54% under deceleration), and only $\sim$29% of the free optimum's energy lies in $\mathrm{span}\,\Psi_{32}$ (43% with 128 modes). The gap is not rank: the free optima are low-rank (8 modes carry 99.8%) and a basis built from four trajectories' free optima recovers 92–97% of the free reduction on the fifth with $K = 8$ (48% on the regime-dependent deceleration); it is the POD material, $R_{\mathcal{B}}\mathcal{P}\mathbf{d}_n$ rather than $\mathbf{s}_n^*$. Every closed-loop number in the paper uses the original Ψ, and "band oracle" denotes the POD-restricted quantity throughout.

Table 6: One-step near-field residual $\|\mathbf{d}_n - G\mathbf{b}_n\|_W$ ($r < 1.5$, $\times 10^{-2}$, mean over the evaluation window of the reference states) for the coefficient vectors of the ablation ladder; $b*$ is the per-frame optimum in $K_{\text{out}} = 8$ modes. **mean share**: fraction of the near-field G-energy of $\mathbf{b}_n^*$ over the window carried by its mean (2D rows: mean over the deployments).

	none	mean	bandstatic	mlpcond	mlpnocond	mlp	mlpdyn	$b*$	mean share
re300	1.109	0.687	0.679	0.680	0.679	0.679	1.114	*0.678*	0.9994
re300-decel	0.328	0.835	0.308	0.308	0.308	0.308	0.328	*0.308*	0.9999
re350	1.130	0.708	0.698	0.699	0.698	0.698	1.142	*0.696*	0.9985
re250	1.091	0.679	0.663	0.668	0.702	0.664	1.084	*0.661*	0.9982
re450	1.200	0.782	0.737	0.743	0.743	0.737	1.195	*0.734*	0.9952
2D held-out (26 runs)	1.89	1.30	1.24	1.25	1.21	1.21	1.85	*1.21*	0.968
2D in-dist. (13 runs)	1.91	1.28	1.23	1.24	1.20	1.20	1.88	*1.20*	0.969

Table 7: Closed-loop metrics of the static-versus-dynamic ablations on the five sphere trajectories: near-wall band error **surf** ($\times 10^{-2}$), signed consistent drag error and window-mean lift $|C_l|$ (fine: 0.072 on re300, 0.064 on re250). Window means follow Table 1: own window on re300 and re300-decel, bracketing training Re on re350/re250/re450.

	re300			re300-decel			re350			re250			re450		
mode	surf	dCd	$\lvert C_l\rvert$	surf	dCd	$\lvert C_l\rvert$	surf	dCd	$\lvert C_l\rvert$	surf	dCd	$\lvert C_l\rvert$	surf	dCd	$\lvert C_l\rvert$
`bandstatic` (condition-matched band mean, no learning)	4.43	+0.003	0.072	1.33	+0.002	0.005	6.17	+0.015	0.078	3.25	−0.012	0.057	8.38	+0.027	0.074
`mlpcond` (corrector, conditioning scalars only)	5.07	+0.006	0.075	1.32	+0.005	0.004	7.60	+0.027	0.078	2.68	−0.022	0.053	7.59	+0.014	0.076
`mlpnocond` (corrector, flow features only)	5.05	−0.000	0.100	1.97	−0.002	0.004	5.31	+0.011	0.078	3.71	−0.006	0.070	6.58	−0.004	0.071
`mlpdyn` (corrector's state-dependent part alone)	9.77	+0.017	0.061	2.02	+0.019	0.006	11.57	−0.003	0.086	8.97	+0.038	0.037	13.41	−0.036	0.050

D STATIC VERSUS DYNAMIC CORRECTION

D.1 ABLATIONS

All modes share the band basis Ψ_{out} and differ only in the coefficient vector $\mathbf{b}_n$ of (2). $\mathcal{T}$ is the static-field window of Section 2.5:

`bandstatic`:	$\mathbf{b}_n = \langle \mathbf{b}^*_{1:K_{\text{out}}} \rangle_{\mathcal{T}}$	condition-matched window mean;
`mlpcond`:	$\mathbf{b}_n = f_{\theta_e}(\mathbf{e}_n)$	retrained on the conditioning scalars only;
`mlpnocond`:	$\mathbf{b}_n = f_{\theta_\phi}(\boldsymbol{\phi}_n)$	retrained on the flow features only;
`mlpdyn`:	$\mathbf{b}_n = f_\theta(\boldsymbol{\phi}_n, \mathbf{e}_n) - \langle f_\theta(\boldsymbol{\phi}^*_n, \mathbf{e}_n) \rangle_{\mathcal{T}}$	corrector minus its window mean.

The retrained variants use the protocol of (5); `mlpcond` has 4 inputs (5,000 parameters), `mlpnocond` 24 (1,064), the deployed corrector 28 (6,536). At steady inflow the scalars are constant over the window, so `mlpcond` is a *learned* per-condition lookup; `mlpdyn` isolates what the flow-state input adds, and `volstatic` + `mlpdyn` equals the composed configuration.

D.2 ONE-STEP DECOMPOSITION

Table 6 evaluates the ablation ladder on the evaluation-window reference states. Over the 39 2D deployments the mean carries 97% of the optimal coefficients' near-field G-energy on average (median 99.9%; wide tandem pairs lowest at 80–87%). Window mean, deployed corrector and both retrained variants reach the band oracle's reduction to three digits, and the state-dependent part removes nothing: one-step, the band correction is a condition lookup.

D.3 CLOSED LOOP, SPHERE

Table 7 gives the closed-loop metrics of the ablations; the deployed corrector, the composed configuration, the oracles and the remaining modes are in Tables 1 and 14–18. The condition-matched band mean is within 4% of the band oracle on re300, re300-decel and re250, and interpolating it from bracketing Reynolds numbers changes re300 by 2% (4.54 vs. 4.43). `mlpcond` interpolates less evenly than the band mean (re250 2.68, re350 7.60); `mlpnocond` reaches 1.97 on the decelerated case against the corrector's 1.15 and over-excites the shedding ($|C_l| = 0.100$ vs. fine 0.072). The state-dependent part alone removes none of the drag deficit.

D.4 CLOSED LOOP, 2D

Per group (Table 20) the corrector beats the band mean where the band defect fluctuates (square 0.5–$0.6\times$, wide tandem pairs 0.77–$0.95\times$, cylinder 0.90–$0.96\times$) and loses by 20–78% on the NACA airfoils at $\alpha \leq 15^\circ$ and the closest tandem pair, where the dynamic part alone is inert. Aggregated over the suite (held-out / in-distribution), the condition-only lookup reproduces the window mean (5.62 / 5.27 vs. 5.38 / 5.14), the feature-only corrector sits between it and the deployed corrector (5.21 / 4.83 vs. 4.91 / 4.52) and recovers the square's gain (12.0 vs. 19.5), and the state-dependent part alone reaches 13.10 / 13.02 against the uncorrected 13.82 / 13.37.

D.5 DRAG UNDER COMPOSITION

The composed configuration's larger drag error (re300 $+0.004 \rightarrow -0.015$, re250 $-0.012 \rightarrow -0.029$, re350 $+0.007 \rightarrow -0.010$; 2D held-out $\overline{|\mathrm{d}C_d|}$ $0.030 \rightarrow 0.034$) matches the static field alone and the full-field oracle in sign and size (re300 -0.008, -0.010; re250 -0.026, -0.025): the volumetric static stage carries the overshoot of every volumetric correction, which a band impulse cannot cancel.

E CNN CONTROLS THAT ISOLATE THE NEAR-WALL PRIOR

The CNN variants of Section 3.4 differ from the corrector in where the correction is read and written, output rank, low-passed targets and conditioning at once. The controls in Table 8 change one property at a time, either restricting a *trained* volumetric CNN at deployment or retraining the same U-Net with the near-wall structure built in.

E.1 POST-HOC RESTRICTION

`sotlnoise` and `sotlsur` are deployed unchanged but their predicted defect is (i) kept on the band cells only ($R_{\mathcal{B}}^{\top} R_{\mathcal{B}}$) or (ii) additionally projected, W-weighted, on the corrector's eight spread modes $R_{\mathcal{B}}^{\top} \Psi_{\mathrm{out}}$.

E.2 RETRAINED CONTROLS

Four networks are retrained on the full coarse state with the U-Net's optimiser and stopping rule (patience 25 of at most 200 epochs); `cnnband` and `cnnpod` with three seeds, seed 0 deployed (chosen before any closed-loop run) and the others reported beneath it.

- `cnnloc`: the 1.16M-parameter U-Net writing the defect on the whole domain, uniform MSE (no prior).
- `cnnnw`: as `cnnloc`, with the near-field cells $r < 1.5$ weighted to carry half the loss ($301\times$ per cell).
- `cnnband`: output masked to the band, loss on the band cells only (a band-limited CNN).
- `cnnpod`: the U-Net encoder (widths 16/32/64/96, 0.64M parameters), global average pool and a $96 \rightarrow 64 \rightarrow 8$ $\tanh$ head, trained on the corrector's targets and deployed through $R_{\mathcal{B}}^{\top} \Psi_{\mathrm{out}}$ (the corrector's output structure on a convolutional backbone).

`mlp1280` is the corrector retrained on exactly the CNN's 1,280 reference states and its $1{,}152/128$ split, with the deployed corrector's frozen PCA and low-passed targets, so only frame count and spacing change (the deployed corrector sees $18{,}584$ frames at 0.1 spacing, Appendix A); its one-step validation residual equals the deployed corrector's (7.02 vs. 7.02×10^{-3}), as does `cnnpod`'s for every seed (0.700 vs. 0.698×10^{-2}): any reasonable predictor of the eight coefficients reaches the one-step optimum, and only the closed loop separates them.

Reading. *Post-hoc support is not the prior.* Keeping only the band cells of a trained volumetric network removes the far-field damage (global error from 1.4–$3.4\times$ `none` to at or below it) but not the near-wall error (re300 `sotlnoise` $5.19 \rightarrow 6.97$), and under deceleration the restricted networks remain no better than no correction.

Table 8: **Isolating the near-wall prior.** Closed-loop **near** ($r < 1.5$) and **surf** velocity error $\times 10^{-2}$ on the five sphere test trajectories and the global error on re300: post-hoc band and eight-mode restrictions of the two best volumetric networks, the retrained controls (`cnnloc`, `cnnnw`, `cnnband`, `cnnpod`) and `mlp1280`. The unrestricted networks, the deployed corrector and the oracles are in Tables 1 and 14–18. Indented *seed* rows: extra seeds. Bold: lowest deployable error per column, including those reference modes (ties within the 1.3% noise floor share it).

	re300 (in-dist.)		re300-decel		re350 (h.-out)		re250 (h.-out)		re450 (h.-out)		re300
mode	near	surf	near	surf	near	surf	near	surf	near	surf	glob
`sotlnoise`, band cells only	5.58	6.97	1.50	2.64	6.39	7.75	4.49	5.85	7.44	7.87	0.71
`sotlnoise`, band cells, 8 spread modes	2.59	3.56	1.02	1.47	2.67	3.88	1.98	2.90	5.76	5.77	0.52
`sotlsur`, band cells only	4.22	5.48	2.25	3.46	6.15	8.53	4.92	6.89	6.15	8.51	0.64
`sotlsur`, band cells, 8 spread modes	**2.00**	**2.49**	**0.84**	**0.96**	**2.29**	**3.21**	**1.61**	**2.27**	**5.47**	**5.05**	**0.49**
`cnnloc` (same U-Net, whole-field output, retrained)	13.61	14.21	6.01	5.46	8.43	9.25	26.16	28.45	6.71	8.02	11.91
`cnnnw` (same U-Net, near-wall weighted loss)	4.01	4.41	2.49	2.10	4.67	4.91	4.06	4.68	8.11	8.20	2.15
`cnnband` (band-limited U-Net, retrained)	2.67	3.04	1.31	1.80	3.78	4.40	2.63	3.14	5.48	5.33	0.51
seed 1	2.55	2.94	1.33	1.80	3.48	3.98	2.40	2.96	5.28	5.40	0.50
seed 2	2.64	2.84	1.22	1.54	4.25	4.81	2.32	2.75	5.64	5.66	0.53
`cnnpod` (CNN encoder → 8 band coefficients)	3.33	4.32	0.95	1.33	2.92	3.49	2.93	4.88	7.14	7.90	0.56
seed 1	3.42	4.46	0.94	1.30	3.39	4.26	1.69	2.21	8.25	9.85	0.56
seed 2	3.49	4.55	0.95	1.31	6.65	8.48	2.22	2.88	6.72	7.24	0.57
`mlp1280` (corrector on the CNN's 1,280 pairs)	3.30	4.32	1.09	1.82	3.25	3.91	2.85	4.03	7.05	7.15	0.55

Table 9: Force read-outs of the same runs: signed consistent drag error $\mathrm{d}C_d$ (positive: under-correction; negative: overshoot) and the window-mean lift magnitude on re300 (fine $|C_l| = 0.072$; the coarse solver damps it to 0.059). Bold: smallest $|\mathrm{d}C_d|$ among deployed modes per column whose near-wall error is below `none`'s, the reference modes of Table 1 included (values within 0.001 share it).

mode	re300	re300-decel	re350	re250	re450	$\lvert C_l\rvert$ re300
`sotlnoise`, band cells only	$-.025$	$-.\mathbf{000}$	$-.021$	$-.044$	$-.019$	0.075
`sotlnoise`, band cells, 8 spread modes	$-.033$	$+.\mathbf{001}$	$-.033$	$-.040$	$-.031$	0.060
`sotlsur`, band cells only	$-.025$	$+.003$	$-.032$	$-.027$	$-.015$	0.042
`sotlsur`, band cells, 8 spread modes	$-.027$	$+.005$	$-.041$	$-.033$	$-.030$	0.060
`cnnloc` (same U-Net, whole-field output, retrained)	$-.006$	$-.011$	$-.018$	$-.005$	$-.044$	0.089
`cnnnw` (same U-Net, near-wall weighted loss)	$-.019$	$+.003$	$-.014$	$-.018$	$-.031$	0.054
`cnnband` (band-limited U-Net, retrained)	$-.012$	$+.002$	$-.006$	$-.014$	$-.013$	0.063
seed 1	$-.014$	$+.001$	$-.009$	$-.024$	$-.018$	0.063
seed 2	$-.014$	$+.002$	$-.005$	$-.015$	$-.010$	0.062
`cnnpod` (CNN encoder → 8 band coefficients)	$+.\mathbf{003}$	$+.002$	$-.\mathbf{003}$	$-.016$	$+.024$	0.072
seed 1	$+.004$	$+.002$	$-.007$	$-.033$	$+.038$	0.073
seed 2	$+.004$	$+.003$	$+.032$	$-.011$	$+.018$	0.075
`mlp1280` (corrector on the CNN's 1,280 pairs)	$-.006$	$-.004$	$-.\mathbf{003}$	$-.022$	$+.\mathbf{004}$	0.066

Rank is, and so is training with the band in the loss. Projected on the corrector's eight modes, both networks pass every test their volumetric deployment failed; the drag overshoot of the restricted short-unroll network ($\mathrm{d}C_d = -0.027$ to -0.041) comes with damped shedding ($|C_l| = 0.060$), because the band restriction of a predicted defect approximates $R_{\mathcal{B}}\mathbf{d}_n$ rather than the injected-response optimum $\mathbf{b}^*$ of Eq. equation 4. Retraining band-limited (`cnnband`) recovers most of it: every seed beats the corrector's near-wall error on the steady trajectories, passes the deceleration (1.54–1.80 vs. 1.15) and holds $|\mathrm{d}C_d| \leq 0.024$ at a 12% damped shedding.

Given the output prior, the backbone is secondary. `cnnpod` matches the corrector's drag, shedding ($|C_l| = 0.072$–0.075) and field error on re300, re300-decel and re450 within noise, but scatters by seed at the held-out Reynolds numbers (re350 3.49–8.48, re250 2.21–4.88) despite being one-step-identical to the corrector: the convolutional encoder interpolates less reliably in Reynolds number than the shell harmonics. `mlp1280` sits within the corrector's seed spread, so pair count is not what limits the volumetric networks.

Without the output prior the U-Net is the volumetric network again. `cnnloc` is worse than no correction on re300, the deceleration and re250 (**surf** $14.21/5.46/28.45$ vs. $12.41/2.35/10.55$) with 12–$36\times$ `none`'s global error; weighting the loss to the near field (`cnnnw`) recovers the steady trajectories but not the deceleration (2.10) nor the far field ($3\times$). Restricting what the network *writes*, with nothing about what it reads changed, is what turns the same U-Net into a corrector that

Table 10: **Sphere: physical drag error** $\bar{C}_d - \bar{C}_d^{\text{fine}}$ under the three read-outs (fine / +imp / cal, defined above), same runs as Table 1. Fine drag 0.657, 0.617, 0.707, 0.605 and 0.101; operator bias -0.120, -0.115, -0.127, -0.119 and -0.024 (re300, re350, re250, re450, re300-decel). The corrector's impulse is $C_d^{\text{imp}} = 0.13$–0.15 on the steady trajectories and 0.03 under deceleration. Bold: smallest deployable error in the +imp and cal columns; mode names as in Table 1.

	re300 (in-dist.)			re350 (h.-out)			re250 (h.-out)			re450 (h.-out)			re300-decel	
mode	fine	+imp	cal	fine	+imp	cal	fine	+imp	cal	fine	+imp	cal	fine	+imp
`none`	$-.115$	$-.115$	$+.007$	$-.125$	$-.125$	$-.011$	$-.101$	$-.101$	$+.034$	$-.170$	$-.170$	$-.064$	$-.003$	$\mathbf{-.003}$
`mean`	$-.108$	$+.037$	$+.015$	$-.115$	$+.030$	$\mathbf{+.001}$	$-.114$	$+.032$	$+.018$	$-.175$	$-.029$	$-.070$	$-.169$	$-.023$
`sotlsur` (CNN, short unroll)	$-.129$	—	$-.010$	$-.111$	—	$+.007$	$-.116$	—	$+.016$	$-.122$	—	$\mathbf{-.004}$	$-.017$	—
`sotlsur` → 8 band modes	$-.147$	$\mathbf{-.005}$	$-.032$	$-.157$	$\mathbf{-.004}$	$-.050$	$-.161$	$-.010$	$-.039$	$-.149$	$-.008$	$-.037$	$-.019$	$+.005$
`bandstatic` (band mean)	$-.117$	$+.017$	$\mathbf{+.004}$	$-.100$	$+.032$	$+.020$	$-.139$	$+.011$	$-.013$	$-.092$	$+.034$	$+.034$	$-.022$	$+.006$
Ours (corrector)	$-.116$	$+.027$	$+.006$	$-.109$	$+.025$	$+.010$	$-.139$	$+.013$	$\mathbf{-.013}$	$-.101$	$+.027$	$+.022$	$-.022$	$+.005$
Ours + static (bracketing Re)	$-.130$	$+.010$	$-.011$	$-.125$	$+.006$	$-.011$	$-.157$	$\mathbf{-.008}$	$-.034$	$-.125$	$\mathbf{+.003}$	$-.007$	—	—
Ours + static (own)†	$-.135$	$+.001$	$-.017$	—	—	—	—	—	—	—	—	—	$-.019$	$+.009$
opt	$-.118$	$+.015$	$+.003$	$-.103$	$+.027$	$+.017$	$-.140$	$+.009$	$-.014$	$-.089$	$+.037$	$+.037$	$-.022$	$+.006$
vol	$-.131$	—	$-.012$	$-.121$	—	$-.006$	$-.153$	—	$-.030$	$-.103$	—	$+.020$	$-.019$	—

survives the conditioning shift; among carriers of the band prior the memoryless MLP is the best under deceleration and the only one that reproduces the reference shedding amplitude.

F Physical drag: coarse-operator consistency versus the fine reference

The consistent metric of Section 3.1 and the fine run's own drag $\bar{C}_d^{\text{fine}}$ differ by the *operator bias* $\bar{C}_d^{\text{ref}} - \bar{C}_d^{\text{fine}}$, the drag the coarse operator (embedded-boundary stress on the 128^3 grid) fails to read off the correctly filtered fine field: 18–20% of $\bar{C}_d^{\text{fine}}$ on the steady sphere trajectories, 15–18% on the spheroid and -0.11 on average in 2D (from 8% on the cylinder to a factor two on the tandem pairs). Tables 10–12 report three read-outs of the physical error $\bar{C}_d - \bar{C}_d^{\text{fine}}$ for the same runs as the main tables:

- **fine**: $\bar{C}_d^{\text{run}} - \bar{C}_d^{\text{fine}}$, the coarse-operator drag of the run with the correction ignored in the force balance.
- **+imp**: the same with the corrector's impulse booked to the body. The band increment is added once per frame ($h = 0.1$); its pre-projection momentum $I_n = \hat{\mathbf{b}}_n^\top \mathbf{m}$ ($\mathbf{m}_k = \Delta^d \mathbf{1}^\top W \psi_k$, Section 2) is a momentum sink the coarse embedded boundary did not supply, and the projection passes it to the domain boundaries as it does the resolved wall force. Counted as body force, $C_d^{\text{imp}} = -\langle I_{n,x} \rangle / h \,/ (\frac{1}{2} U_0^2 A_{\text{ref}})$, and the error is $\bar{C}_d^{\text{run}} + C_d^{\text{imp}} - \bar{C}_d^{\text{fine}}$. No fitting is involved. The static field carries no net momentum (Π_0); the full-field oracle and the volumetric CNNs do not log their increments (—).
- **cal**: $\rho\, \bar{C}_d^{\text{run}} - \bar{C}_d^{\text{fine}}$ with $\rho = \bar{C}_d^{\text{fine}} / \bar{C}_d^{\text{ref}}$ fitted on training conditions only. The ratio is stationary in time and smooth in Reynolds number (1.195 at $\mathrm{Re} = 100$ to 1.250 at 500 on the sphere). Sphere: the training window $t \in [55, 135)$ of the five steady training Reynolds numbers, interpolated linearly to the held-out Re (measured ratios agree to 0.3%); spheroid, the single training condition $\mathrm{Re} = 300$, $\alpha = 15°$, applied to the held-out $\mathrm{Re} = 350$ (fitted 1.215, measured 1.221; a10 and a20 are training conditions and use their own ratio); 2D, per geometry the six training Reynolds numbers, interpolated to 156.3 and 312.5. Dynamic inflow is not calibrated, its ratio being history-dependent (1.31 on the decelerated sphere trajectory).

Reading. Under the raw *fine* read-out every corrected run, oracles included, is 0.09–0.17 below the fine drag, and on the spheroid (Section 3.6) and in 2D (0.196 vs. 0.136) it is farther from it than the uncorrected solver. This is not a property of the correction: the operator bias is the same for every mode, and the uncorrected solver looks better only because its own error, an over-prediction of the consistent drag by $+0.02$ to $+0.06$, cancels part of it. The corrected field, being closer to the reference, exposes the bias in full.

Booking the corrector's impulse to the body closes most of the gap without fitting. On the sphere the impulse matches the independently measured operator bias, 0.115–0.127, to within 0.01–0.03

Table 11: **Spheroid: physical drag error**, same runs as Table 4. Fine drag 0.606, 0.595, 0.667, 0.567, 0.107; operator bias -0.107, -0.096, -0.097, -0.103, -0.028 (re300, a20, a10, re350, decel). The corrector's impulse is 0.11–0.12 at $\alpha = 15$–20° and 0.16 at 10°. The closed-loop coefficient logs of the spheroid corrector runs were not kept: its +imp column uses the frozen model's open-loop prediction on the reference features of the same window (the corrector sits on its band oracle there; the oracle's impulse, exact, is within 0.002). The calibration for re350 is fitted at $\mathrm{Re} = 300$ alone.

	re300 (in-dist.)			a20 (in-dist.)			a10 (in-dist.)			re350 (h.-out Re)			decel (h.-out inflow)	
mode	fine	+imp	cal	fine	+imp	cal	fine	+imp	cal	fine	+imp	cal	fine	+imp
`none`	$-.043$	$-.043$	$+.078$	$-.033$	$-.033$	$+.075$	$-.046$	$\mathbf{-.046}$	$+.060$	$-.042$	$-.042$	$+.070$	$-.006$	$\mathbf{-.006}$
`mean`	$-.091$	$\mathbf{-.006}$	$+.019$	$-.055$	$\mathbf{-.001}$	$+.048$	$-.026$	$-.090$	$+.083$	$-.087$	$\mathbf{-.001}$	$+.016$	$-.033$	$+.052$
`sotlsur` (CNN unrolled, short)	$-.089$	—	$+.022$	$-.089$	—	$+.008$	$-.079$	—	$+.022$	$-.088$	—	$+.014$	$+.044$	—
Ours (corrector)	$-.103$	$+.015$	$\mathbf{+.005}$	$-.096$	$+.014$	$\mathbf{-.001}$	$-.104$	$+.056$	$\mathbf{-.008}$	$-.101$	$+.014$	$\mathbf{-.001}$	$-.037$	$+.054$
`opt`	$-.105$	$+.012$	$+.003$	$-.098$	$+.012$	$-.003$	$-.114$	$+.044$	$-.020$	$-.102$	$+.011$	$-.002$	$-.015$	$+.003$
`vol`	$-.097$	—	$+.012$	$-.103$	—	$-.009$	$-.090$	—	$+.008$	$-.097$	—	$+.003$	$-.011$	—

Table 12: **2D suite: mean drag-error magnitude** over the 26 held-out-Re and 13 in-distribution deployments of Table 3, under the consistent metric (cons) and the three physical read-outs. Mean operator bias -0.11 (mean magnitude 0.18; ratio ρ from 1.08 on the cylinder to 2.0 on the tandem pairs). Bold: smallest deployable per column; the raw *fine* column is not bolded because every corrected mode is farther from the fine drag than the uncorrected solver there.

	held-out Re (26 runs)				in-dist. Re (13 runs)			
mode	cons	fine	+imp	cal	cons	fine	+imp	cal
`none` (no correction)	0.067	0.136	0.136	0.096	0.053	0.158	0.158	0.086
`mean` (uncond. static)	0.040	0.193	**0.121**	0.064	0.026	0.209	**0.123**	0.047
`sotlnoise` (CNN + noise)	0.052	0.176	—	0.087	0.049	0.231	—	0.083
`bandstatic` (cond.-matched band mean)	**0.024**	0.196	0.161	0.051	0.022	0.212	0.191	**0.045**
Ours (near-wall corrector)	0.030	0.196	0.165	**0.049**	**0.021**	0.208	0.190	0.046
Ours (+ static bias)	0.034	0.194	0.160	0.053	0.036	0.207	0.188	0.064
`opt` (band oracle)	0.030	0.200	0.137	0.046	0.025	0.214	0.146	0.053
`vol` (full-field oracle)	0.036	0.198	—	0.051	0.027	0.211	—	0.056

on all four steady trajectories and to 0.003 under deceleration; the corrector's error becomes $+0.013$ to $+0.027$ (2–4% of $\bar{C}_d^{\mathrm{fine}}$), -0.008 to $+0.010$ with the static field, against -0.10 to -0.17 uncorrected. The corrector is fitted to velocity increments and never sees a force, so this agreement is a physical consistency check: the band impulse is the wall momentum flux the coarse operator misses. The spheroid agrees at $\alpha = 15$–20° and overshoots at 10° ($+0.056$). In 2D the fraction of the bias the impulse explains follows the flow: 65–100% on the cylinder and the aerofoils at $\alpha \le 10^\circ$, 30–50% on the plate and at 15–20°, 10–30% on the tandem pairs, and the wrong sign on the square, whose bias is a base-pressure effect no near-wall impulse represents; the aggregate therefore improves only from 0.196 to 0.165.

The calibrated read-out, with one ρ per condition fitted on training data and shared by all modes, is the fair cross-mode comparison: the corrector is closest to the fine drag on every held-out condition (sphere re450 0.022 vs. 0.064 uncorrected; 2D held-out 0.049 vs. 0.096), and the band-projected CNN's drag overshoot (Appendix E) survives all three read-outs.

Field versus force. The corrector improves the field representation on every suite. It improves the physical force only through a read-out that accounts for the operator bias: by counting its own impulse as body force (sphere and spheroid, within 0.01–0.03) or by a calibration fitted on training conditions (all suites). The raw coarse-operator drag of a corrected run must not be compared with the fine drag, and the main text's drag claims refer to the consistent metric and the impulse-inclusive reading.

G LONG-HORIZON ROLLOUTS

The scored window of the main text, $t \in [135, 150]$, is about two shedding periods at $\mathrm{St} \approx 0.13$. To test whether the closed loop merely stays near the fine state it was started from, we extend it, with the deployed assets of Table 1 unchanged (held-out trajectories use the bracketing-Re static field; re300 uses the Re 200/400 field).

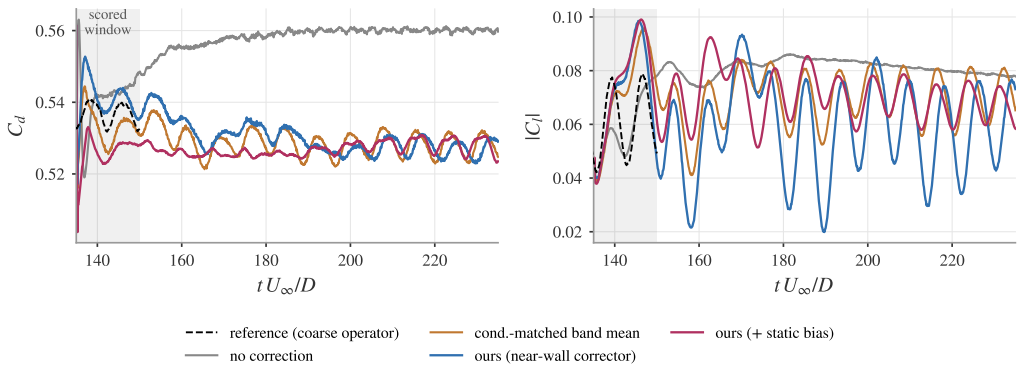


Figure 6: Long-horizon closed loop, sphere at $\mathrm{Re} = 300$, from the filtered reference at $t = 135$ to $t = 235$ (grey: scored window of the main text). Left: drag; right: transverse force magnitude. Dashed: filtered reference under the coarse force operator ($t \le 150$).

Table 13: **Long-horizon rollouts**, $t_0 = 135$, run to $t = 235$. $\mathbf{surf}_{150}$ ($\times 10^{-2}$): band error at $t = 150$ (window means: Table 1); $\bar{C}_d$, $\mathrm{d}C_d$ and St over $t \in [150, 235]$, σ_{C_l} ($\times 10^{-2}$) over $t \in [190, 235]$; $\|\hat{\mathbf{b}}\|$: mean corrector output norm over $[135, 150]$ / $[190, 235]$.

mode	surf_{150}	St	$\bar{C}_d$	$\mathrm{d}C_d$	σ_{C_l}	$\|\hat{\mathbf{b}}\|$
re300 (in-dist.); reference $t \in [75, 150]$: $C_d = 0.538$, $\sigma_{C_l} = 1.30$, $\mathrm{St} = 0.133$						
`none` (no correction)	13.40	0.081	0.558	+0.020	0.21	—
`bandstatic` (cond.-matched band mean)	5.76	0.128	0.529	−0.009	0.81	0.752 / 0.752
Ours (near-wall corrector)	5.31	0.124	0.530	−0.008	1.61	0.772 / 0.773
Ours (+ static bias)	3.36	0.129	0.527	−0.011	1.19	0.772 / 0.777
re350 (held-out); reference $t \in [75, 150]$: $C_d = 0.505$, $\sigma_{C_l} = 2.53$, $\mathrm{St} = 0.128$						
`none` (no correction)	15.46	0.138	0.514	+0.010	0.38	—
Ours (near-wall corrector)	5.02	0.112	0.506	+0.001	3.31	0.806 / 0.796
Ours (+ static bias)	4.17	0.124	0.495	−0.010	3.12	0.791 / 0.793
re250 (held-out); reference $t \in [75, 150]$: $C_d = 0.579$, $\sigma_{C_l} = 0.31$, $\mathrm{St} =$ —						
`none` (no correction)	11.08	0.099	0.618	+0.039	0.11	—
Ours (near-wall corrector)	3.78	0.124	0.559	−0.020	0.16	0.749 / 0.743
Ours (+ static bias)	1.97	0.081	0.546	−0.033	0.43	0.752 / 0.750

G.1 Protocol

The runs of Table 1 are continued past the end of the reference to $t = 235$, 100 time units or $\sim$13 shedding periods in the loop. No reference frames exist after $t = 150$, so beyond the window the runs are judged on force statistics, on the corrector's output and on boundedness rather than on pointwise error. The target statistics are those of the filtered reference under the coarse force operator over $t \in [75, 150]$ (drag $\bar{C}_d$, lift-fluctuation rms σ_{C_l}, Strouhal number of the transverse force); the uncorrected coarse solver is the control.

G.2 Results (Table 13, Figure 6)

Every run is bounded and free of NaNs to $t = 235$. The corrector's output norm is a constant of the loop: its mean over the first 15 units and over the last 45 differ by under 1.5% on all three trajectories, with no drift and no growth of the maximum, so the feedback settles on a fixed operating point rather than ratcheting. The uncorrected solver is the run that leaves the reference's regime: at $\mathrm{Re} = 300$ and 350 its lift fluctuation decays by an order of magnitude over the horizon (σ_{C_l} 0.21 and 0.38×10^{-2} against the reference's 1.30 and 2.53; the shedding dies into a steady, asymmetric wake with a larger steady lift) and its drag drifts $+0.02$ above the reference at $\mathrm{Re} = 300$. With the corrector the shedding persists at the reference's Strouhal number (0.124–0.129 vs. 0.133 at $\mathrm{Re} = 300$) and a fluctuation amplitude within 25% of the reference's when composed with the static field (1.19 vs. 1.30; 3.12 vs. 2.53 at $\mathrm{Re} = 350$); the drag error keeps the sign and size it had in the scored window (-0.008 to -0.011 at $\mathrm{Re} = 300$, $+0.001/-0.010$ at $\mathrm{Re} = 350$) instead of

growing. The corrector alone is over-energetic (1.2–1.3× the reference σ_{C_l}), the unlearned band mean under-energetic (0.81). At $\mathrm{Re} = 250$ the reference wake is steady and planar-asymmetric (no shedding) and every run relaxes to a steady state; the corrected runs sit 0.02–0.03 below the reference drag, the uncorrected one 0.04 above. Inside the scored window the near-wall error is not growing either: at $t = 150$ it is at most 1.3× its window mean (Table 1) on every run.

H ADDITIONAL 3D FIGURES

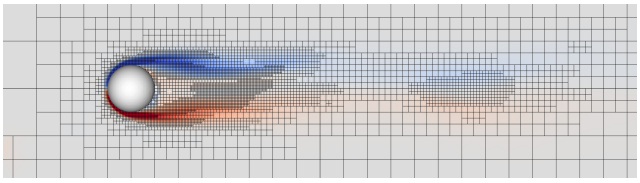


Figure 7: The fine reference: mid-plane vorticity of the level-9 octree simulation at $\mathrm{Re} = 300$ ($t = 150$) with the adaptive mesh overlaid ($D/\Delta = 32$ at the body). Filtered to the uniform 128^3 grid, this field is the target every closed-loop correction is scored against.

Figure 7 shows the fine reference and its adaptive mesh. Figure 8 shows the closed-loop sphere wakes of Section 3.2 as λ_2 skeletons: five to six time units in, the uncorrected solver has lost the wake while the corrected loop is nearly indistinguishable from the filtered reference, including at the held-out Reynolds number. Figure 9 repeats the Figure 3 comparison at the farthest held-out Reynolds number, $\mathrm{Re} = 450$ ($t = 138.5$), with the same ranking: the one-step and noise-trained CNNs grow a spurious core jet, the unrolled variants deform the head and shed debris, and the composed correction tracks the reference structures.

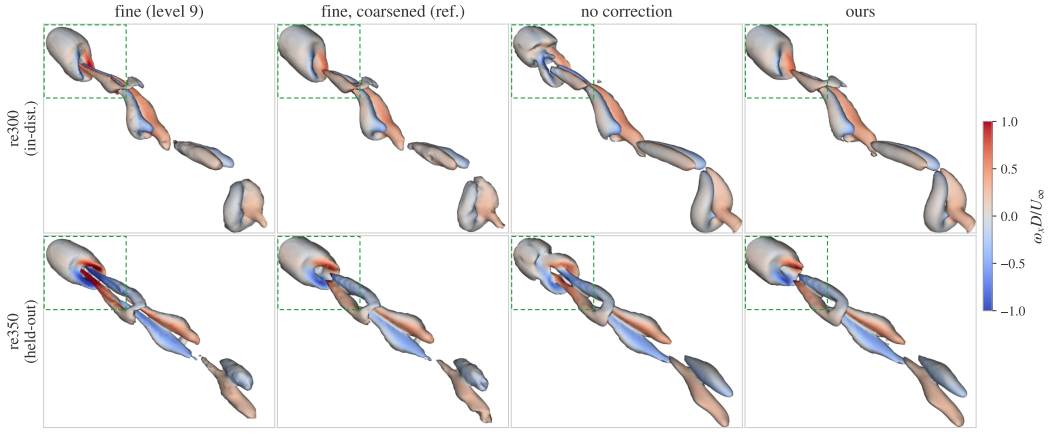


Figure 8: Wake vortical structure: λ_2 isosurfaces ($\lambda_2 = -0.01\,U_\infty^2/D^2$) coloured by streamwise vorticity ω_x, five to six time units into the closed loop, on re300 and the held-out re350 (re300-decel omitted: no coherent vortices remain after the deceleration). Columns: fine simulation; filtered reference; uncorrected coarse solver, which invents a spurious vortex loop around the body (dashed box: the corrector's band) and fuses threads the reference keeps separate; composed correction, which returns the structures in shape, position and phase.

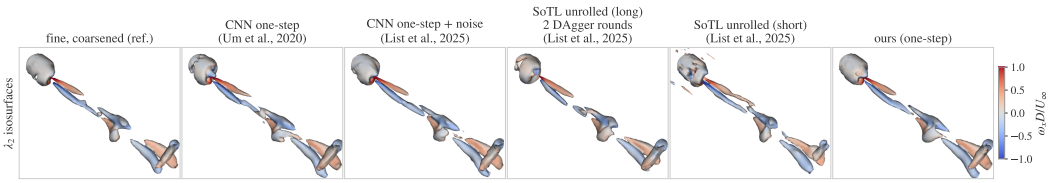
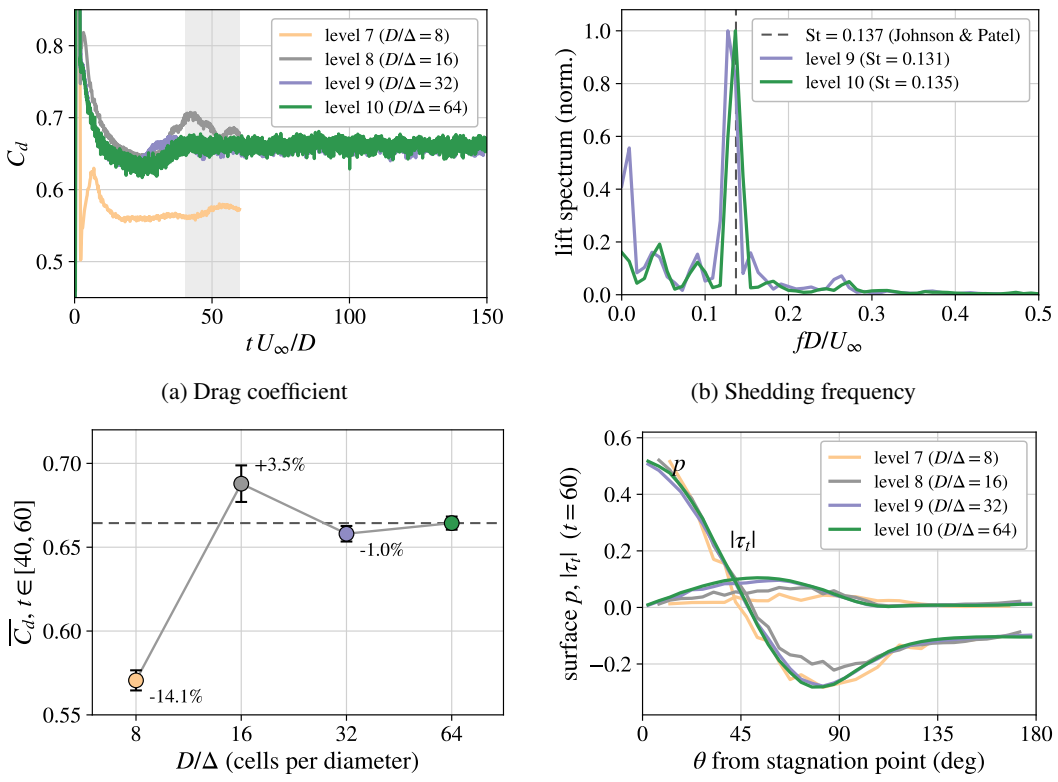


Figure 9: Wake vortical structure in the closed coarse-solver loop at the farthest held-out Reynolds number, $\mathrm{Re} = 450$ ($t = 138.5$): λ_2 isosurfaces ($\lambda_2 = -0.01\,U_\infty^2/D^2$) coloured by streamwise vorticity ω_x. Columns as in Figure 3.



(a) Drag coefficient

(b) Shedding frequency

(c) Mean drag versus resolution

(d) Surface loads

Figure 10: Grid convergence of the fine reference (re300 configuration, levels 7–10, $D/\Delta = 8$–64). (a) Drag coefficient; levels 7–8 run to $t = 60$, shading marks the common statistics window $t \in [40, 60]$. (b) Normalised lift spectrum over $t \in [40, 150]$ for levels 9–10; the dashed line is the reference $\mathrm{St} = 0.137$ of Johnson & Patel (1999). (c) Window-mean drag versus resolution; error bars are the standard deviation over the window, annotations the change relative to level 10, the dashed line the level-10 mean. (d) Azimuthally averaged surface pressure p and tangential traction $|\tau_t|$ versus polar angle from the stagnation point at $t = 60$ (colours as in panel a).

I Grid convergence of the fine reference

The re300 configuration was re-run at octree levels 7–10 ($D/\Delta \in \{8, 16, 32, 64\}$) with everything else fixed; levels 7–8 run to $t = 60$, levels 9–10 to $t = 150$, and cross-level statistics use the common window $t \in [40, 60]$ (Figure 10).

Window-mean drag is $\overline{C_d} = 0.571, 0.688, 0.658$ and 0.664 at levels 7–10 (-14.1%, $+3.5\%$, -1.0% relative to level 10); the level-9-to-10 gap (0.006) matches the shedding-cycle variability within the window (0.004), and both fine levels agree with $C_d = 0.656$ of Johnson & Patel (1999). Lift spectra give $\mathrm{St} = 0.131$ (level 9) and 0.135 (level 10), within one frequency bin of each other and 4% of the reference 0.137; level 7 has no shedding peak. Surface pressure and tangential traction, the quantities the corrector is trained on, collapse between levels 9 and 10 while levels 7–8 misplace separation. Level 10 quadruples the compute for $< 1\%$ change in the scored statistics, so level 9 is used throughout.

J Full per-trajectory metrics

J.1 Sphere suite

Tables 14–18 report all modes and metrics of Section 3.1 on the five test trajectories, including the five CNN variants; Table 1 collects their surf and dCd columns. Oracle rows are italic; bold marks the best deployable mode per column (ties within the 1.3% noise floor); physical drag read-

Table 14: Closed-loop metrics on **re300** (in-distribution, steady inflow). Velocity errors $\times 10^{-2}$ (window means). Consistent reference $C_d^{\text{ref}} = 0.537$; fine $C_d^{\text{fine}} = 0.657$, fine $|C_l| = 0.072$. Italic rows are oracles.

mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l \rvert$
`none`	0.72	7.00	12.41	0.541	+0.005	0.059
`mean`	0.60	4.28	5.66	0.548	+0.012	0.079
`sotl1s`	3.07	5.26	6.89	0.550	+0.014	0.067
`sotlnoise`	1.41	4.65	5.19	0.529	−0.008	0.067
`sotlur`	1.71	8.98	12.15	0.555	+0.018	0.188
`sotlur2`	1.46	9.93	12.18	0.570	+0.033	0.121
`sotlsur`	1.79	8.05	10.68	0.528	−0.009	0.034
`opt`	*0.55*	*3.26*	*4.28*	*0.539*	*+0.002*	*0.071*
`mlp`	0.58	3.80	4.90	0.541	**+0.004**	0.072
`volstatic`	0.43	**1.80**	**2.25**	0.529	−0.008	0.063
`mlpvol`	0.48	2.49	3.31	0.512	−0.025	0.058
`mlp+static` (deployed)	**0.42**	**1.80**	2.29	0.522	−0.015	0.058
`volstatic` (Re 200/400 field)	0.57	2.18	2.63	0.522	−0.015	0.072
`mlp+static` (Re 200/400 field)	0.58	2.42	2.95	0.527	−0.010	0.075
`volmean`	*0.36*	*1.37*	*1.83*	*0.526*	*−0.011*	*0.062*
`vol`	*0.28*	*1.01*	*1.55*	*0.526*	*−0.010*	*0.058*

Table 15: Closed-loop metrics on **re300-decel** (inflow decelerated $U{=}1 \to 0.3$ before the window; $\mathrm{Re}_{\text{eff}} \approx 90$, no shedding). Velocity errors $\times 10^{-2}$. $C_d^{\text{ref}} = 0.077$; $C_d^{\text{fine}} = 0.101$.

mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l \rvert$
`none`	0.25	1.29	2.35	0.098	+0.021	0.006
`mean`	0.71	11.02	19.84	−0.067	−0.144	0.003
`sotl1s`	5.95	5.30	4.13	0.083	+0.006	0.017
`sotlnoise`	1.56	4.58	4.34	0.094	+0.017	0.010
`sotlur`	1.77	3.44	4.33	0.094	+0.017	0.021
`sotlur2`	0.94	4.60	5.49	0.084	+0.007	0.011
`sotlsur`	1.22	2.23	2.47	0.084	+0.007	0.009
`opt`	*0.24*	*0.96*	*1.34*	*0.079*	*+0.002*	*0.005*
`mlp`	0.24	0.91	1.15	0.080	**+0.003**	0.005
`volstatic`	**0.21**	**0.43**	**0.36**	0.082	+0.006	0.004
`mlpvol`	0.22	0.78	1.25	0.079	**+0.003**	0.004
`mlp+static` (deployed)	**0.21**	**0.44**	0.40	0.082	+0.005	0.004
`bandstatic` (steady Re 300 band mean)	0.71	11.04	19.87	−0.063	−0.140	0.005
`volstatic` (steady Re 300 field)	0.89	11.65	20.30	−0.066	−0.142	0.007
`mlp+static` (steady Re 300 field)	0.56	2.75	3.27	0.077	+0.000	0.008
`volmean`	*0.14*	*0.26*	*0.26*	*0.083*	*+0.006*	*0.004*
`vol`	*0.13*	*0.22*	*0.24*	*0.083*	*+0.006*	*0.004*

outs of the same runs are in Appendix F. Rows tagged with a field provenance ("Re 200/400 field", "steady Re 300 field") deploy `mlp+static` with a static field not estimated on the scored trajectory (Appendix B).

J.2 SPHEROID SUITE

Table 19 adds the training incidence re300-a10 and the global/near metrics omitted from Table 4. `sotlsur` edges the corrector on the band metric at $\alpha = 10^\circ$ (2.20 vs. 2.69×10^{-2}) but carries 2–3× its global error on the steady trajectories. $|C_l|$ is dominated by the mean lift at incidence (fine 0.14–0.42).

K 2D MULTI-GEOMETRY STUDY: PER-GROUP METRICS

Table 20 breaks the aggregates of Table 3 down by geometry group and adds the static-versus-dynamic ablations of Appendix D. Every group improves monotonically along `none` → `mean` → corrector → composed correction; the one-step CNN is worse than no correction in 11 of 13 groups

Table 16: Closed-loop metrics on **re350** (held-out Reynolds number; deployed correctors interpolated from $\mathrm{Re} = 300/400$). Velocity errors $\times 10^{-2}$. $C_d^{\mathrm{ref}} = 0.502$; $C_d^{\mathrm{fine}} = 0.617$.

mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l \rvert$
`none`	0.85	8.19	13.83	0.491	−0.010	0.081
`mean`	**0.62**	4.05	5.02	0.502	**+0.000**	0.080
`sotl1s`	4.14	12.39	12.45	0.484	−0.018	0.108
`sotlnoise`	1.80	7.22	7.86	0.489	−0.012	0.112
`sotlur`	1.89	8.55	11.75	0.507	+0.006	0.176
`sotlur2`	1.70	12.63	14.50	0.531	+0.029	0.110
`sotlsur`	3.38	10.31	12.62	0.506	+0.004	0.113
`opt`	*0.62*	*4.25*	*5.60*	*0.514*	*+0.013*	*0.079*
`mlp`	0.65	4.37	5.38	0.508	+0.007	0.069
`volstatic`	0.63	3.52	4.14	0.503	+0.002	0.076
`mlp+static` (deployed)	0.64	**3.23**	**3.74**	0.492	−0.010	0.075
`volmean`	*0.45*	*2.38*	*3.09*	*0.494*	*−0.008*	*0.070*
`vol`	*0.32*	*1.56*	*2.30*	*0.496*	*−0.006*	*0.060*

Table 17: Closed-loop metrics on **re250** (second held-out Reynolds number; correctors interpolated from $\mathrm{Re} = 200/300$). Velocity errors $\times 10^{-2}$. $C_d^{\mathrm{ref}} = 0.580$; $C_d^{\mathrm{fine}} = 0.707$.

mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l \rvert$
`none`	0.57	5.71	10.55	0.606	+0.027	0.030
`mean`	0.50	4.29	6.27	0.593	+0.013	0.074
`sotl1s`	6.44	21.68	23.08	0.610	+0.030	0.177
`sotlnoise`	1.21	4.35	4.54	0.554	−0.026	0.062
`sotlur`	1.91	6.17	8.54	0.609	+0.029	0.164
`sotlur2`	1.64	9.98	13.38	0.604	+0.025	0.117
`sotlsur`	1.53	8.28	10.68	0.591	**+0.012**	0.079
`opt`	*0.41*	*2.24*	*3.15*	*0.567*	*−0.013*	*0.056*
`mlp`	**0.42**	2.50	3.38	0.568	**−0.012**	0.057
`volstatic`	0.44	**1.34**	1.76	0.554	−0.026	0.054
`mlp+static` (deployed)	0.44	**1.34**	**1.72**	0.550	−0.029	0.054
`volmean`	*0.25*	*0.78*	*1.41*	*0.555*	*−0.025*	*0.049*
`vol`	*0.24*	*0.71*	*1.36*	*0.554*	*−0.025*	*0.049*

Table 18: Closed-loop metrics on **re450** (third held-out Reynolds number, the farthest from the training anchors; correctors interpolated from $\mathrm{Re} = 400/500$). Velocity errors $\times 10^{-2}$. $C_d^{\mathrm{ref}} = 0.485$; $C_d^{\mathrm{fine}} = 0.605$.

mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l \rvert$
`none`	1.31	10.46	16.64	0.435	−0.051	0.066
`mean`	**0.92**	**4.08**	**4.85**	0.430	−0.056	0.075
`sotl1s`	4.15	11.12	12.14	0.476	−0.010	0.088
`sotlnoise`	2.12	7.98	8.05	0.479	−0.006	0.084
`sotlur`	2.06	10.55	14.20	0.476	−0.009	0.210
`sotlur2`	2.30	13.21	14.58	0.524	+0.039	0.154
`sotlsur`	2.16	10.48	13.52	0.482	−0.003	0.065
`opt`	*1.41*	*7.87*	*8.98*	*0.516*	*+0.030*	*0.074*
`mlp`	1.42	7.41	7.87	0.503	+0.018	0.077
`volstatic`	1.28	5.63	5.29	0.488	**+0.002**	0.072
`mlp+static` (deployed)	1.32	5.78	5.19	0.480	−0.005	0.076
`volmean`	*1.19*	*5.46*	*5.40*	*0.501*	*+0.016*	*0.073*
`vol`	*0.51*	*2.58*	*3.83*	*0.502*	*+0.016*	*0.065*

(noise variant 9, unrolled rounds 7 and 10), and all four trail the corrector in all 13. Figure 11 shows three representative wakes; Appendix L gives four further frames for six geometry groups.

Table 19: Full closed-loop metrics for the spheroid suite (Section 3.6). Velocity errors $\times 10^{-2}$ (window means); $\mathrm{d}C_d$ against each trajectory's consistent coarse-operator reference. Consistent references C_d^{ref} (fine C_d, fine $|C_l|$): re300 0.499 (0.606, 0.250); re300-a10 0.570 (0.667, 0.143); re300-a20 0.499 (0.595, 0.394); re350 0.464 (0.567, 0.277); re300-decel 0.079 (0.107, 0.034). Italic rows are oracles; bold marks the best deployable mode per column.

	mode	glob↓	near↓	surf↓	C_d	$\mathrm{d}C_d \to 0$	$\lvert C_l\rvert$
re300	none	0.50	3.78	7.16	0.563	+0.064	0.250
	mean	0.39	1.81	3.16	0.515	+0.016	0.195
	mlp	**0.36**	**1.36**	**2.28**	0.503	**+0.004**	0.198
	sotlsur	0.95	2.05	2.44	0.517	+0.018	0.196
	opt	*0.36*	*1.33*	*2.25*	*0.501*	*+0.002*	*0.201*
	vol	*0.25*	*0.54*	*0.79*	*0.509*	*+0.010*	*0.193*
a10	none	0.45	3.35	6.45	0.621	+0.051	0.138
	mean	0.64	5.06	9.41	0.641	+0.071	0.085
	mlp	**0.36**	**1.59**	2.69	0.564	**−0.006**	0.118
	sotlsur	0.76	1.72	**2.20**	0.589	+0.018	0.113
	opt	*0.38*	*1.98*	*3.54*	*0.553*	*−0.017*	*0.119*
	vol	*0.24*	*0.46*	*0.56*	*0.577*	*+0.007*	*0.111*
a20	none	0.56	4.47	8.47	0.562	+0.063	0.416
	mean	0.55	4.36	8.27	0.539	+0.040	0.387
	mlp	**0.38**	**1.56**	**2.62**	0.498	**−0.001**	0.354
	sotlsur	0.92	2.49	3.36	0.505	+0.006	0.344
	opt	*0.38*	*1.60*	*2.67*	*0.497*	*−0.002*	*0.354*
	vol	*0.27*	*0.56*	*0.83*	*0.492*	*−0.007*	*0.343*
re350	none	0.52	4.20	7.91	0.525	+0.060	0.282
	mean	0.41	2.33	4.05	0.480	+0.016	0.220
	mlp	**0.37**	**1.77**	**2.96**	0.466	**+0.001**	0.225
	sotlsur	1.23	2.55	3.27	0.479	+0.014	0.227
	opt	*0.35*	*1.53*	*2.52*	*0.465*	*+0.001*	*0.226*
	vol	*0.23*	*0.60*	*0.92*	*0.470*	*+0.005*	*0.221*
decel	none	**0.24**	**0.92**	**1.29**	0.101	+0.022	0.026
	mean	0.72	7.18	12.02	0.074	**−0.005**	0.021
	mlp	0.53	5.15	9.56	0.070	−0.009	0.016
	sotlsur	10.34	8.22	10.04	0.151	+0.072	0.086
	opt	*0.24*	*0.75*	*0.75*	*0.092*	*+0.013*	*0.026*
	vol	*0.12*	*0.17*	*0.18*	*0.096*	*+0.017*	*0.029*

L 2D MULTI-GEOMETRY STUDY: FULL FLOW-FIELD GALLERY

Figures 12–17 show four closed-loop frames for six of the geometry groups, drawn from the three deployments $\mathrm{Re} \in \{156.3, 234.4, 312.5\}$ (the first and last held out) and spread over the evaluation window; each column is labelled with its Reynolds number and frame time (tandem $L/D{=}3$ at $\mathrm{Re} = 312.5$ is Figure 5). Rows: fine; filtered; uncorrected loop; composed correction. One symmetric colour scale per figure (99.5th percentile of $|\omega_z|$ over the reference row). The released figure scripts render the full four-frame sequence ($t = 138, 142, 146, 150$) of all 39 deployments.

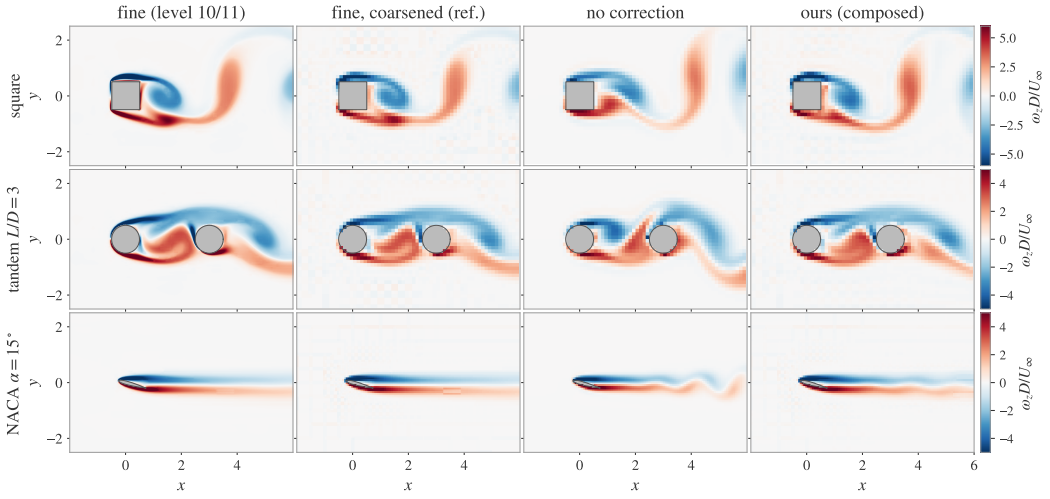


Figure 11: 2D suite: vorticity ω_z on three geometry groups at the held-out $\mathrm{Re} = 312.5$, 7–15 time units into the closed loop (one symmetric colour scale per row). Columns: fine simulation; filtered reference; uncorrected coarse solver; composed correction. The uncorrected failure modes differ per geometry (damped, displaced street on the square; wrong gap-flow attractor of the bistable tandem L/D=3 pair; meandering wake on NACA α=15°); the composed correction returns each wake to the reference structure and phase. Single frames; further frames of every geometry group are in Appendix L.

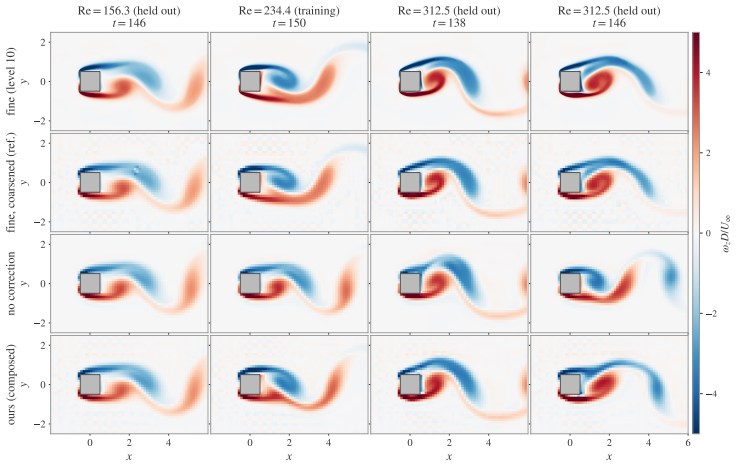


Figure 12: 2D gallery: square, $N = 128$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

Table 20: 2D suite, near-wall band error **surf** $\times 10^{-2}$ per geometry group. Top: held-out Reynolds numbers (mean over $\mathrm{Re} \in \{156.3, 312.5\}$); bottom: the in-distribution $\mathrm{Re} = 234.4$. **Ours** is the deployed near-wall corrector (`mlp`), `bandstatic` the condition-matched band mean, `mlpdyn` the corrector's state-dependent part alone, Ours+static the composition with the static bias field; `sotl1s`/`sotlnoise`/`sotlur`/`sotlur2` are the per-group solver-in-the-loop CNN baselines (Section 3.4); `opt` and `vol` are the band and full-field oracles. $\mathrm{d}C_d$: signed consistent drag error of the corrector.

held-out Re (2 runs per group) geometry	none	mean	bandstatic	Ours	mlpdyn	Ours+static	sotl1s	sotlnoise	sotlur	sotlur2	*opt*	*vol*	$\mathrm{d}C_d \to 0$
cylinder	13.20	6.5	6.38	5.73	12.70	3.51	16.7	11.4	13.8	23.4	*6.38*	*2.72*	−0.076
square	27.01	20.4	19.52	11.61	27.30	10.82	28.7	28.7	16.4	32.2	*7.57*	*2.67*	+0.028
tandem L/D=1.5	9.51	3.9	3.38	4.01	9.68	2.80	17.2	15.9	13.4	12.1	*3.72*	*1.46*	−0.018
tandem L/D=2	9.84	3.7	3.15	3.13	9.31	2.23	22.4	17.9	17.0	33.4	*3.18*	*1.34*	−0.032
tandem L/D=3	19.35	10.8	7.16	7.60	10.93	5.61	22.6	19.3	20.7	27.6	*8.44*	*2.19*	−0.099
tandem L/D=4	13.68	9.2	7.10	6.71	12.68	4.62	24.6	15.8	13.6	23.5	*6.33*	*2.46*	−0.027
tandem L/D=5	15.22	9.7	8.27	6.37	11.78	6.21	21.2	20.1	11.0	24.3	*6.52*	*2.90*	−0.020
NACA α=0°	8.98	2.8	1.87	2.71	9.77	1.65	38.1	19.9	10.0	13.6	*1.79*	*1.02*	−0.009
NACA α=5°	9.07	3.2	2.00	3.23	8.87	1.47	19.3	4.7	8.9	12.6	*1.94*	*1.05*	−0.008
NACA α=10°	12.06	3.5	2.39	3.06	13.78	2.10	18.8	26.7	12.6	12.4	*2.32*	*1.76*	−0.011
NACA α=15°	15.16	4.7	3.53	4.26	14.66	2.32	12.3	25.9	24.3	10.5	*3.37*	*1.61*	+0.005
NACA α=20°	17.31	4.3	3.29	3.54	19.87	1.56	13.6	10.5	12.8	12.8	*3.51*	*1.76*	−0.008
plate α=20°	9.33	2.8	1.95	1.82	8.91	1.30	23.0	16.0	7.6	4.2	*1.81*	*1.02*	−0.015

in-distribution Re (1 run per group) geometry	none	bandstatic	Ours	mlpdyn	*opt*
cylinder	13.70	6.01	5.77	13.30	*6.24*
square	25.88	19.68	10.39	25.52	*7.99*
tandem L/D=1.5	9.71	3.24	3.91	9.80	*3.30*
tandem L/D=2	9.67	4.00	4.00	9.28	*4.13*
tandem L/D=3	11.07	5.15	4.20	9.06	*4.78*
tandem L/D=4	17.75	7.63	6.76	14.75	*7.13*
tandem L/D=5	15.75	6.80	5.89	13.27	*6.23*
NACA α=0°	8.65	1.66	2.25	9.18	*1.65*
NACA α=5°	8.99	1.92	3.43	8.42	*1.93*
NACA α=10°	12.04	2.32	3.00	14.51	*2.32*
NACA α=15°	15.10	3.21	4.14	14.47	*3.21*
NACA α=20°	16.96	3.62	3.33	19.82	*3.46*
plate α=20°	8.58	1.60	1.63	7.87	*1.64*

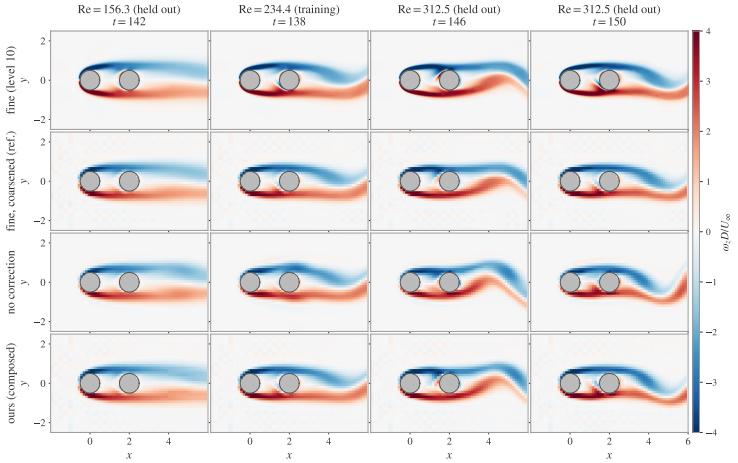


Figure 13: 2D gallery: tandem L/D=2, $N = 128$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

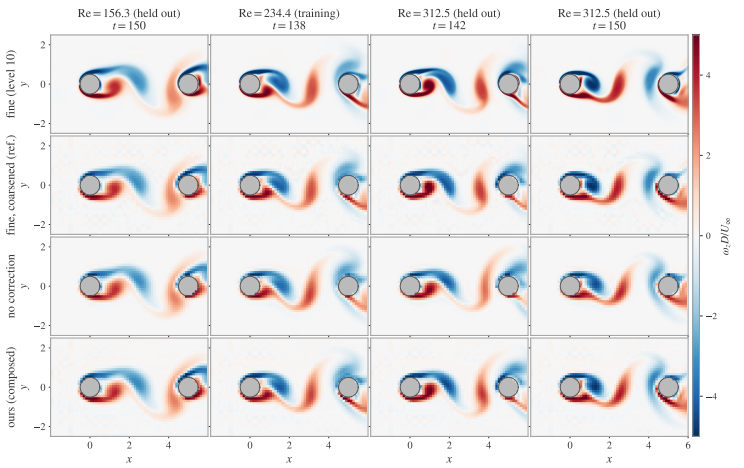


Figure 14: 2D gallery: tandem L/D=5, $N = 128$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

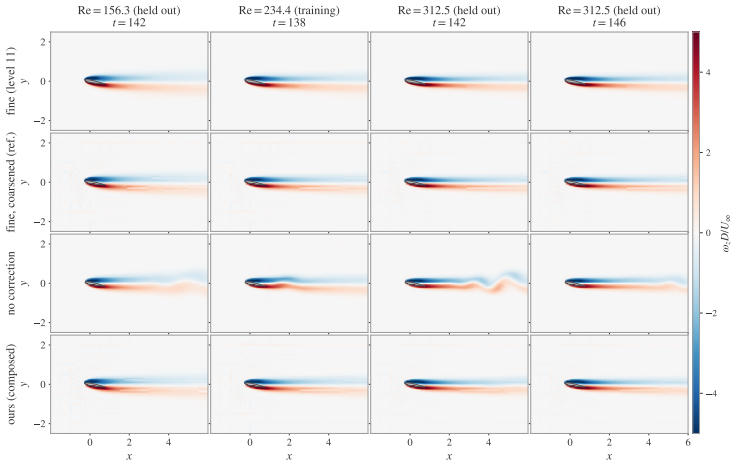


Figure 15: 2D gallery: NACA α=10°, $N = 256$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

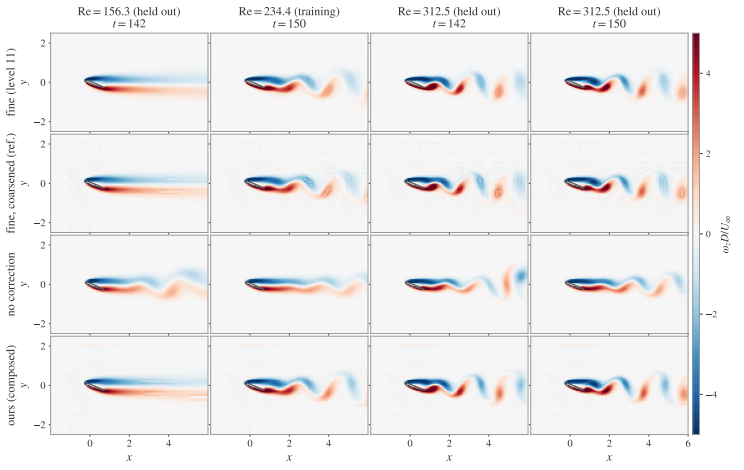


Figure 16: 2D gallery: NACA α=20°, $N = 256$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

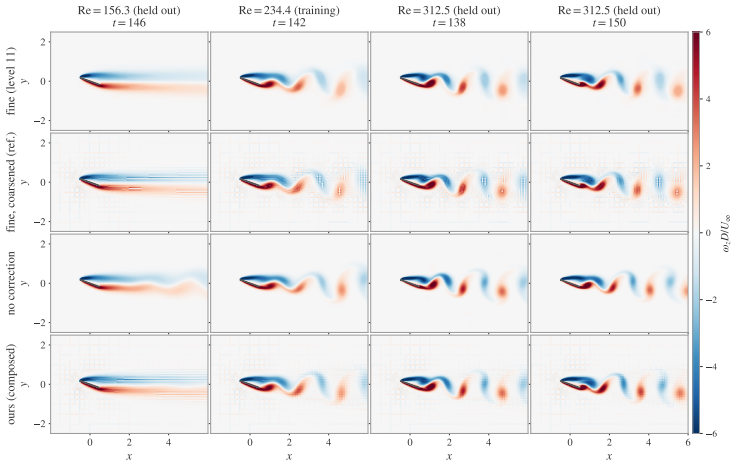


Figure 17: 2D gallery: plate α=20°, $N = 512$. Rows fine / fine coarsened / no correction / composed correction; Reynolds number (held out or training) and frame time per column.

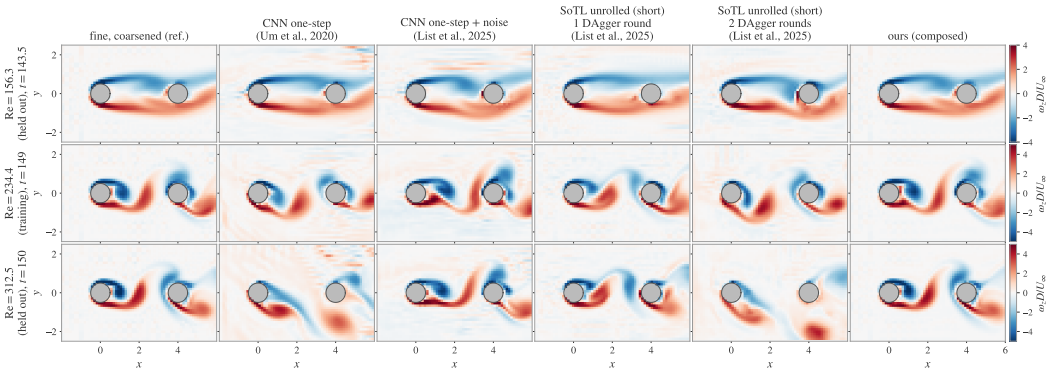
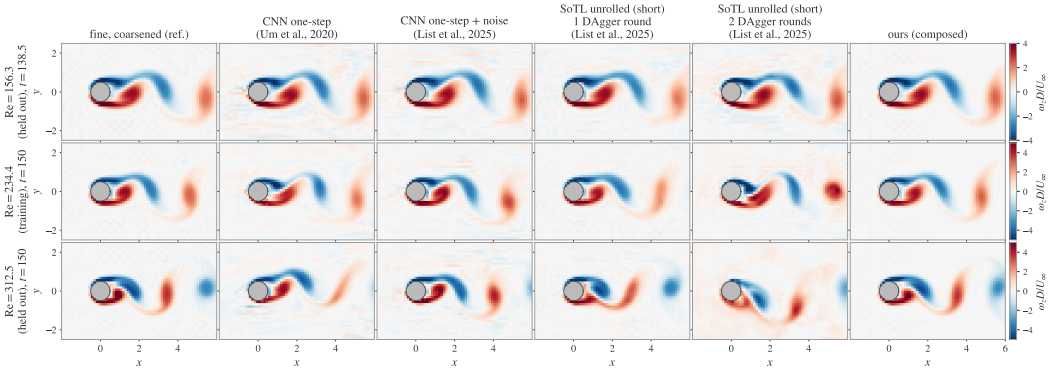


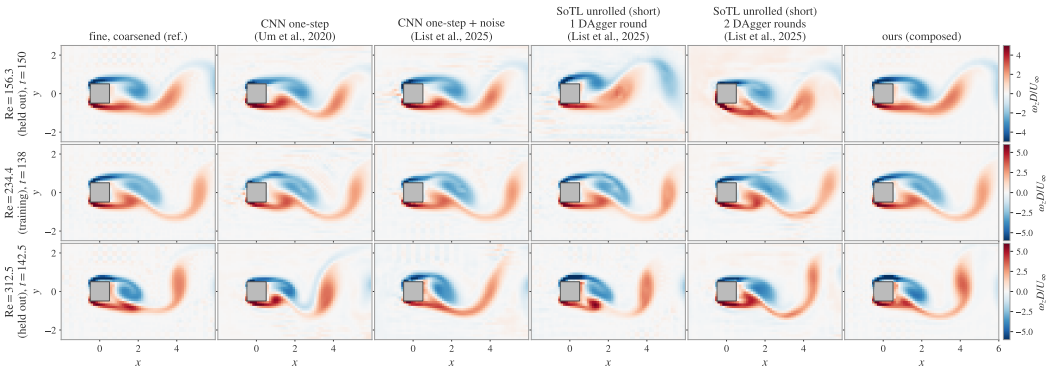
Figure 18: The 2D counterpart of Figure 3: closed-loop vorticity ω_z on the tandem $L/D{=}4$ group. Rows: $\mathrm{Re} = 156.3$ (held out), 234.4 (training), 312.5 (held out). Columns: filtered reference; the four training variants of the per-group solver-in-the-loop CNN; composed near-wall correction. The CNN variants wash out or misplace the gap shear layers and litter the field with invented vorticity at training and held-out Reynolds numbers alike; the composed correction keeps gap flow and shedding locked to the reference.

M 2D MULTI-GEOMETRY STUDY: SOLVER-IN-THE-LOOP CNN FLOW FIELDS

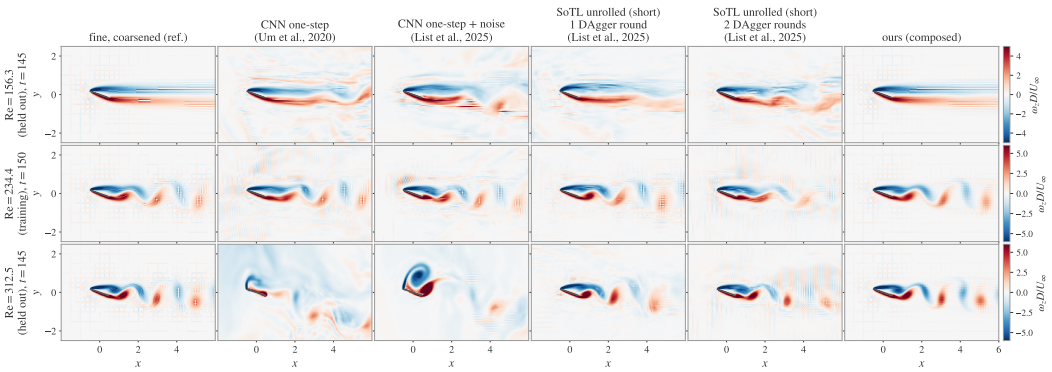
Figures 19–22 show the fields behind the 2D solver-in-the-loop numbers (Tables 3 and 20), one panel per geometry group, rows the three deployments $\mathrm{Re} \in \{156.3, 234.4, 312.5\}$ at the frame of largest CNN-vs-ours contrast. Columns: filtered reference; `sotl1s`; `sotlnoise`; `sotlur`; `sotlur2`; composed correction. All four CNN variants overwrite the wake with fragmented, over-energised vorticity, most starkly on the quasi-steady NACA $\alpha{=}0^\circ$ wing (Figure 21a); round-1 unrolled fields are calmer than one-step but misshapen, and round 2 drifts again.



(a) cylinder, $N = 128$ ($t = 138.5, 150, 150$ by row)

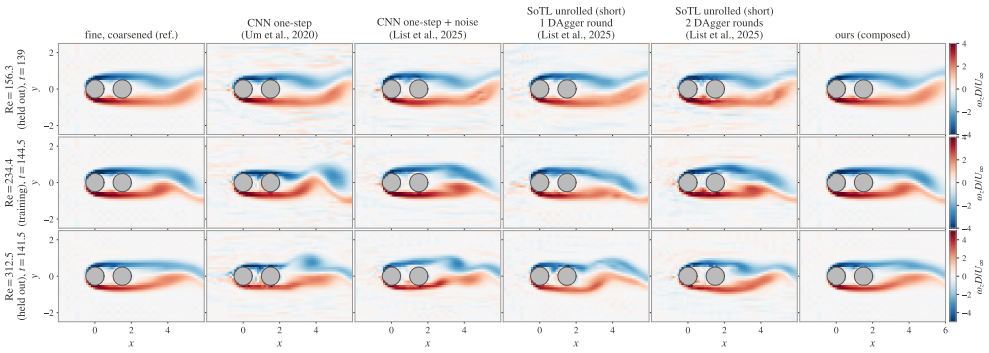


(b) square, $N = 128$ ($t = 150, 138, 142.5$ by row)

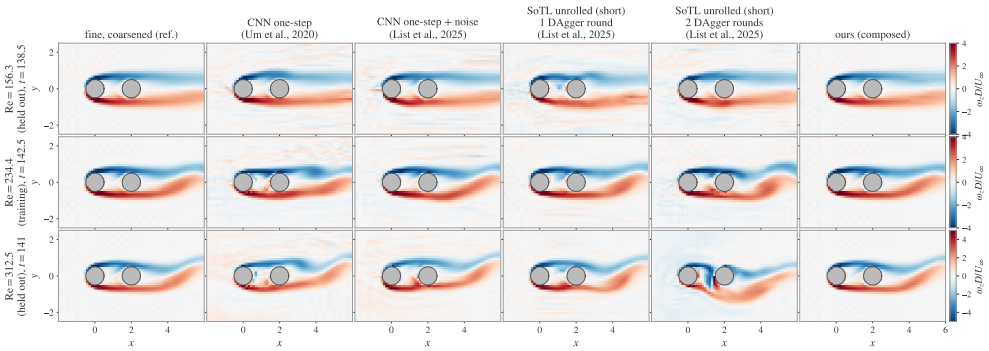


(c) plate α=20°, $N = 512$ ($t = 145, 150, 145$ by row)

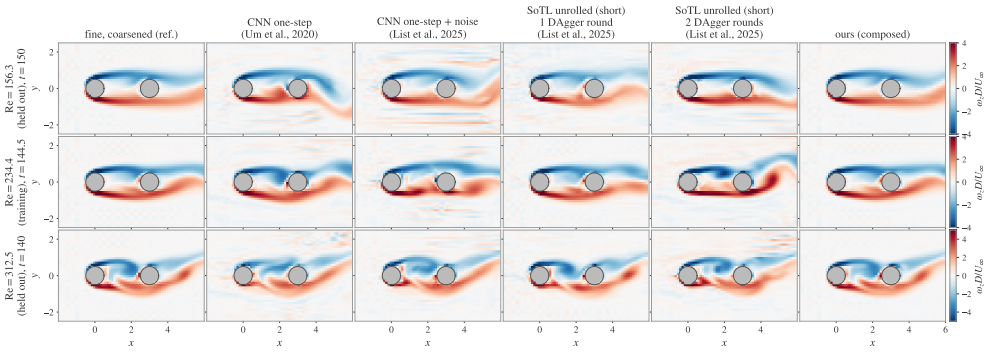
Figure 19: 2D solver-in-the-loop comparison, bluff bodies (cylinder, square, plate). Rows of each panel: Re = 156.3 (held out), 234.4 (training), 312.5 (held out), each at the evaluation-window frame of maximal CNN-vs-ours contrast (frame times in the panel captions). Columns as in Figure 18.



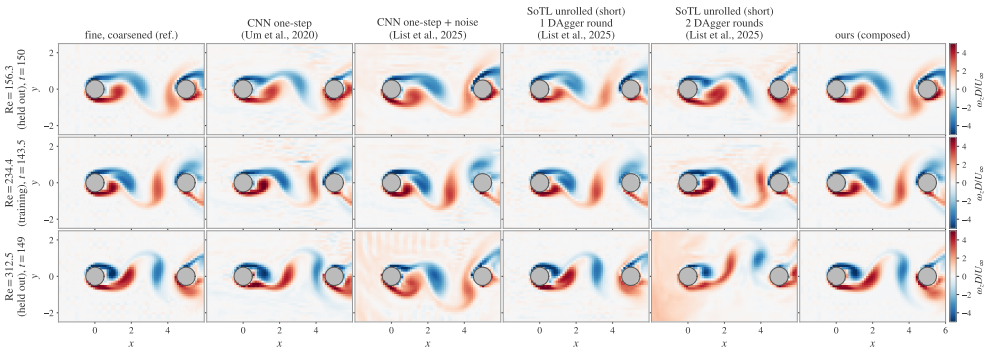
(a) tandem L/D=1.5, $N = 128$ ($t = 139, 144.5, 141.5$ by row)



(b) tandem L/D=2, $N = 128$ ($t = 138.5, 142.5, 141$ by row)

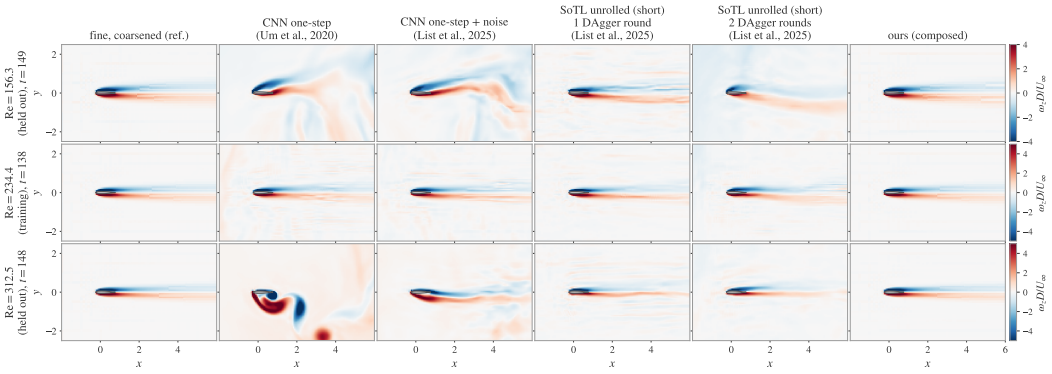


(c) tandem L/D=3, $N = 128$ ($t = 150, 144.5, 140$ by row)

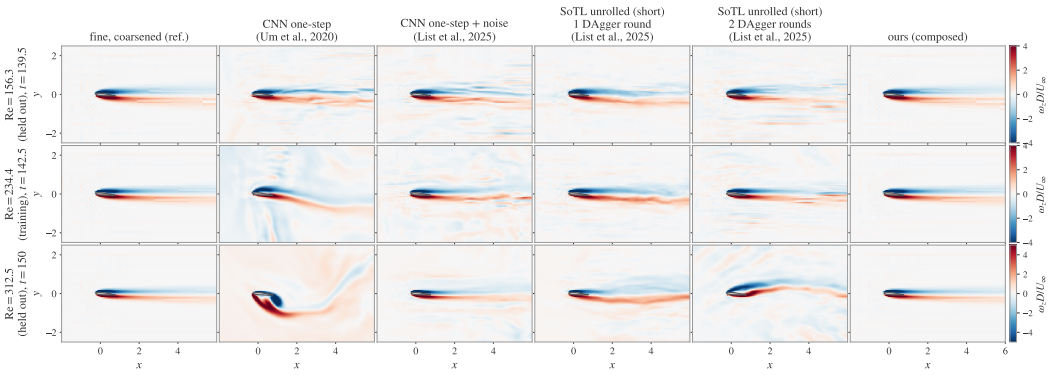


(d) tandem L/D=5, $N = 128$ ($t = 150, 143.5, 149$ by row)

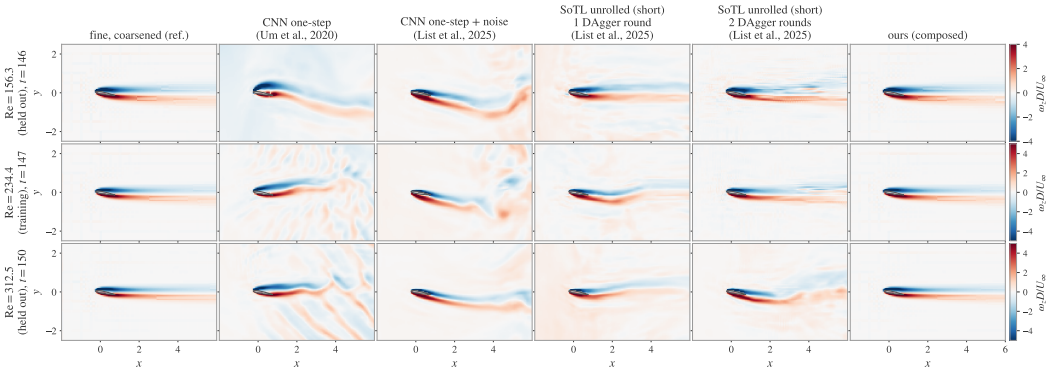
Figure 20: 2D solver-in-the-loop comparison, tandem cylinder pairs. Rows of each panel: $\mathrm{Re} = 156.3$ (held out), 234.4 (training), 312.5 (held out), each at the evaluation-window frame of maximal CNN-vs-ours contrast (frame times in the panel captions). Columns as in Figure 18.



(a) NACA α=0°, $N = 256$ ($t = 149, 138, 148$ by row)

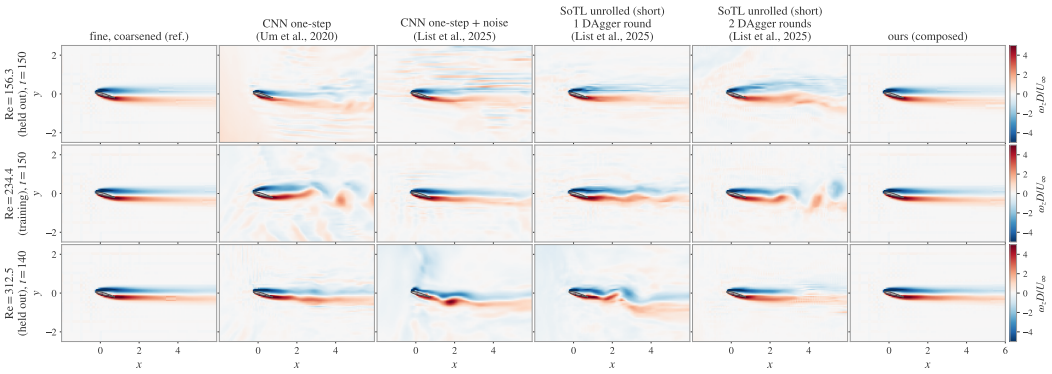


(b) NACA α=5°, $N = 256$ ($t = 139.5, 142.5, 150$ by row)

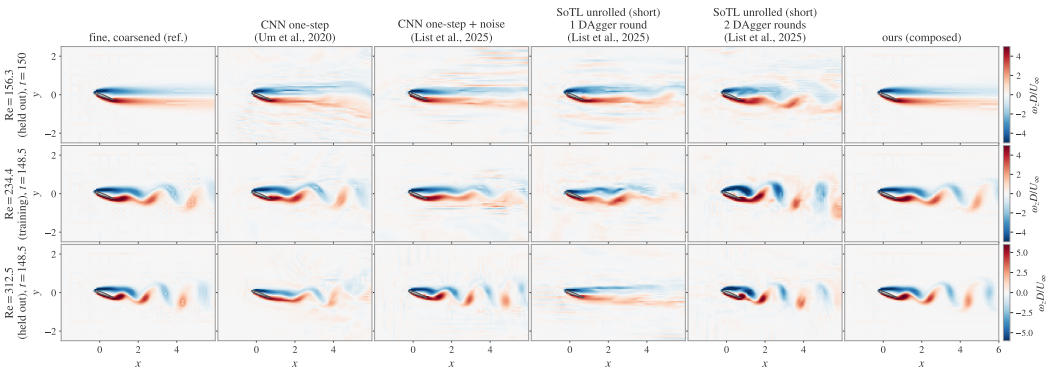


(c) NACA α=10°, $N = 256$ ($t = 146, 147, 150$ by row)

Figure 21: 2D solver-in-the-loop comparison, NACA wing, $\alpha \leq 10°$. Rows of each panel: $\mathrm{Re} = 156.3$ (held out), 234.4 (training), 312.5 (held out), each at the evaluation-window frame of maximal CNN-vs-ours contrast (frame times in the panel captions). Columns as in Figure 18.



(a) NACA α=15°, $N = 256$ ($t = 150, 150, 140$ by row)



(b) NACA α=20°, $N = 256$ ($t = 150, 148.5, 148.5$ by row)

Figure 22: 2D solver-in-the-loop comparison, NACA wing, $\alpha \geq 15°$. Rows of each panel: $\mathrm{Re} = 156.3$ (held out), 234.4 (training), 312.5 (held out), each at the evaluation-window frame of maximal CNN-vs-ours contrast (frame times in the panel captions). Columns as in Figure 18.