EQUIVARIANT DEEPONETS FOR SYMMETRY-AWARE OPERATOR LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural operators constitute a powerful class of deep learning architectures within scientific machine learning which learn mappings between function spaces, yielding natural and effective surrogates for parametric systems of partial differential equations (PDEs). A classical and potent method is the Deep Operator Network (DeepONet), which separates an input-dependent branch and output-coordinate trunk before fusing. Simultaneously, equivariant deep learning has been shown to be a powerful mechanism to imbue networks with knowledge of physical symmetries. In this work, we construct equivariant DeepONets through a representation-theoretic framework in which the observation, branch, trunk, and fusion maps obey compatible group actions. Our construction uses equivariant linear layers and symmetry-compatible branch-trunk fusion. We show that finite observations are compatible with equivariance when they recover a finite-dimensional invariant input representation or when the sensor family is closed under the spatial group and prove a universal approximation theorem for continuous equivariant operators by equivariant DeepONets. We evaluate the framework on numerous PDE systems with different symmetry groups, equivariance actions, and system sizes, including Allen-Cahn dynamics on the sphere, non-rotating spherical shallow water equations, thermoelasticity on an annulus, and a two-component Gross-Pitaevskii system. The equivariant architectures achieve order-of-magnitude reductions in prediction error and equivariance diagnostics over baseline DeepONets while often simultaneously possessing fewer learnable parameters, and the finite-sensor variants achieve comparable results. We also compare favorably to both appropriate FNO architectures and DeepONets with pre-chosen symmetry-aware trunks.

1 INTRODUCTION

Many problems in scientific computing are naturally formulated at the level of an operator. An initial condition, forcing, geometry, or boundary condition is mapped to a solution field, and a conventional numerical method approximates this map by approximately solving a discrete form of the governing equations separately for each input. For challenging systems, these solves may require fine spatial and temporal discretizations, iterative linear or nonlinear solvers, stability-constrained time stepping, or other computationally expensive procedures to resolve multiscale or stiff dynamics.

A neural operator seeks an approximation of the *map* through a finite-dimensional parameterization. In practice, this parameterization is implemented using deep neural networks and evaluated on finite representations of the input and output functions. It is trained on data, such as that generated by high-fidelity simulations, to learn the mapping from input functions to output functions. Once trained, evaluating the learned operator requires only a forward pass rather than a new numerical solution of the governing equations. The operator timestep is also not constrained like traditional numerical methods. Such a surrogate can substantially reduce the cost of repeated solves while retaining useful predictive accuracy, making neural operators particularly attractive in settings that require many evaluations of the same underlying solution operator. See, for example, Kovachki et al. (2023) and the references therein.

There are two seminal families of neural operators, each of which takes a different approach. A Deep Operator Network (DeepONet) (Lu et al., 2021) separates the dependence on the input function from the output coordinate. Its canonical scalar form, motivated by operator approximation theory (Chen

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.

& Chen, 1995), is

$$\widehat{\mathcal{F}}(u)(y) = \sum_{j=1}^{p} b_j(\mathcal{O}u)\tau_j(y), \tag{1.1}$$

where the branch b_j produces coefficients from an observation $\mathcal{O}u$ of the input and the trunk τ_j produces basis values at the query point y. In this way, DeepONets serve as a particularly efficient reduced-order model. The finite-rank reconstruction may become problematic when the output family is dominated by transformations that are expensive to describe in a fixed linear basis (Venturi & Casey, 2023; Lanthaler et al., 2023). DeepONets have been applied to a variety of complex physical systems such as geometry-dependent hypersonic/supersonic flows (Peyvan et al., 2025) and crack nucleation/propagation in brittle materials (Kiyani et al., 2025).

The Fourier Neural Operator (FNO) introduced in Li et al. (2021), in contrast, repeatedly updates a discretized field through learned spectral integral kernels in a manner analogous to a learned pseudospectral method. It provides an efficient route to global spatial interaction on regular grids. Both methods can be formulated in a discretization-convergent manner, allowing a learned operator to be evaluated on discretizations that differ from those used during training; see Kovachki et al. (2023). FNOs are particularly attractive for dense field-to-field dynamics on regular grids, where their spectral convolution structure can efficiently model global interactions. DeepONets, by contrast, are naturally suited to operator learning from finite observations and pointwise evaluation, and can be easier to adapt to irregular sensor locations or complex geometries. In high-dimensional settings, standard FNO architectures may face unfavorable scaling due to their reliance on tensor-product spatial grids and Fourier modes, while DeepONets can be advantageous when the underlying dynamics admit a relatively low-dimensional representation. We refer to Lu et al. (2022) for a thorough, direct comparison of DeepONets and FNOs on a variety of representative benchmark problems.

While neural operators are designed for and motivated by emulating complex physical system evolution, they do not explicitly leverage physical symmetries. Suppose that a compact group G acts on an admissible collection of inputs X and outputs Y, and the physical solution operator $\mathcal{F} : X \to Y$ satisfies

$$\big(\mathcal{F} \circ \rho_X(g)\big)u = \big(\rho_Y(g) \circ \mathcal{F}\big)\, u, \qquad g \in G.$$

This expresses a functional *symmetry* of the operator called *equivariance*. Note that this constraint ties together predictions for all symmetry-related inputs: learning the operator for u simultaneously constrains its prediction for every transformed input $\rho_X(g)u$. Equivariant neural networks provide a finite-dimensional instantiation which leverages this symmetry in the architectural design (Cohen & Welling, 2016; Kondor & Trivedi, 2018; Finzi et al., 2021). Such symmetry has been incorporated into several neural-operator constructions, though primarily for FNOs. Group-equivariant FNOs constrain spectral kernels under rotations, translations, and reflections (Helwig et al., 2023). Spherical FNOs adapt spectral operator learning to the sphere (Bonev et al., 2023). Finally, equivariant operators have been developed for tensorial constitutive maps, graphon convolution, topology optimization, canonicalization, and symmetry-induced shifts (see Han et al. (2023); Cheng & Peng (2023); Erzmann & Dittmer (2024); Shumaylov et al. (2025); Xu et al. (2026)).

Symmetry has also been explored in DeepONet-specific settings, although explicit equivariant architectures remain less developed than in the FNO literature. Venturi & Casey (2023) introduce a moving reference frame for transformation-dominated outputs, while Jin et al. (2025) incorporate an equivariance-preserving unit into the DeepONet branch network to encode problem-specific permutation symmetries. These approaches, however, remain problem-specific, and no common representation-theoretic treatment of equivariant DeepONets has emerged (to the knowledge of the authors). Moreover, techniques developed for equivariant FNOs do not directly carry over to DeepONets. Rather than iteratively applying learned global convolutional integral-kernel approximations in the spectral domain, DeepONets fuse two maps typically parameterized by MLPs. Further, DeepONets may rely on sparse sensor data, which provide an additional obstruction to equivariance.

Contributions. The goal of this paper is to make equivariance compatibility conditions explicit and construct a general framework for equivariant DeepONets, providing physics-based steering without requiring knowledge of exact governing equations. This involves formalizing the theory and performing evaluation across distinct operator-learning settings. Our contributions are the following:

1. We formulate an equivariant DeepONet as a composition of representation-valued observation and branch maps with an equivariant trunk and an intertwining fusion. The ordinary scalar contraction is recovered as a special case.
2. We prove exact equivariance of the factorization and give two concise finite-sensor theorems. The first covers unisolvent reconstruction on finite invariant input spaces, and the complementary theorem considers group-closed finite sensor families.
3. We prove, under an invariant-exhaustion hypothesis on the output representation, that finite-rank equivariant DeepONets uniformly approximate continuous equivariant operators on compact invariant sets.
4. We realize the construction through evaluations on four PDE families. Each benchmark compares a vanilla sparse DeepONet, a symmetry-aware trunk (modulo one experiment, where it is not sensible), a full-observation equivariant DeepONet, a finite-sensor equivariant DeepONet, and an appropriate FNO reference.
5. We demonstrate that equivariant DeepONets enjoy large performance increases over baseline DeepONets and those with symmetry-aware trunks while also frequently reducing the parameter counts. They also obtain competitive or lower reported errors than the selected FNO references under the configurations tested here.

Section 2 introduces the construction and states the main theorem in the form used by the experiments, as well as provides non-rigorous statements of the remaining key theorems. Section 3 records the four symmetry actions and the purpose of each benchmark. Section 4 reports results of the numerical experiments. Section 5 discusses the conclusions and future work. The requisite representation theory, complete theorem statements and proofs, model problems, data generation, model realizations, metric definitions, and training parameterizations are located in the appendices.

2 Equivariant Deep Operator Networks

Here, we introduce the theoretical construction of equivariant DeepONets and rigorous results that follow. We will only introduce the background needed to state the construction and the main result. Appendix A gives more thorough background, and Appendix B contains the complete proofs of the theorems (as well as more precise statements for various results).

2.1 Operator Equivariance and the DeepONet Factorization

Let G be a compact group, X be a Hilbert space carrying a unitary representation $\rho_X : G \to \mathcal{U}(X)$, and $\Omega \subset \mathbb{R}^n$ be the spatial domain of the output field on which G acts via $y \mapsto g \cdot y$. Let W be a finite dimensional vector space in which the output field takes values, and assume that it carries a unitary representation $\rho_W : G \to \mathcal{U}(W)$. Consider the output function space Y, which we take to be a G-invariant space of W-valued functions on Ω carrying the induced action

$$(\rho_Y(g)f)(y) = \rho_W(g)f(g^{-1} \cdot y), \qquad f \in Y, g \in G, y \in \Omega.$$

Vector-valued Sobolev spaces provide a natural example of such a space. An operator $\mathcal{F} : X \to Y$, possibly nonlinear, is said to be *G-equivariant* if

$$\mathcal{F} \circ \rho_X(g) = \rho_Y(g) \circ \mathcal{F}, \qquad g \in G.$$

Specifically, a linear map $A : V \to Z$ between representation spaces V, Z is called an *intertwiner* if

$$A \circ \rho_V(g) = \rho_Z(g) \circ A, \qquad g \in G.$$

A DeepONet requires four core components, namely an observation map, a branch, a trunk, and a fusion rule; none of these are traditionally required to respect a group action. In order to describe the equivariant construction, let V_{obs}, V_B, V_T be finite-dimensional representation spaces of G, carrying representations $\rho_{\text{obs}}, \rho_B, \rho_T$, respectively. We consider an observation map $\mathcal{O} : X \to V_{\text{obs}}$, a branch $B : V_{\text{obs}} \to V_B$, a trunk $\tau : \Omega \to V_T$, as well as a bilinear map $C : V_B \times V_T \to W$. The associated DeepONet is

$$\widehat{\mathcal{F}}(u)(y) = C\left(B(\mathcal{O}(u)), \tau(y)\right). \tag{2.1}$$

Equivalently, the universal property of tensor products provides a unique linear map $\widetilde{C} : V_B \otimes V_T \to W$ such that $\widetilde{C}(b \otimes t) = C(b, t)$, in which case we may express the DeepONet as

$$\widehat{\mathcal{F}}(u)(y) = \widetilde{C}\left(B(\mathcal{O}(u)) \otimes \tau(y)\right).$$

The linearity of $\tilde{C}$ is convenient when expressing the fusion condition in terms of intertwiners.

Thus far, we have merely abstractly expressed a typed DeepONet. In order to obtain an equivariant construction, we will impose hypotheses on each component of the DeepONet. Namely, for $g \in G$, we require that

$$\begin{cases} \mathcal{O} \circ \rho_X(g) = \rho_{\text{obs}}(g) \circ \mathcal{O}, \\ B \circ \rho_{\text{obs}}(g) = \rho_B(g) \circ B, \\ \tau(g \cdot y) = \rho_T(g)\tau(y), \qquad y \in \Omega \\ C\big(\rho_B(g)b, \rho_T(g)t\big) = \rho_W(g)C(b, t), \qquad b \in V_B, t \in V_T. \end{cases} \tag{2.2}$$

The last condition is equivalent to the intertwining condition

$$\widetilde{C} \circ \big(\rho_B(g) \otimes \rho_T(g)\big) = \rho_W(g) \circ \widetilde{C}.$$

Note that, in a chosen basis, the linear maps may be represented by matrices, in which case the compositions are merely matrix multiplication.

Our first theoretical result is the equivariance of the described architecture; the proofs of all theoretical results are contained in Appendix B.

Theorem 2.1 (Equivariant DeepONets)**.** *Let G be compact, and suppose that the collection of conditions (2.2) holds. If (2.1) defines a map $\widehat{F} : X \to Y$, then $\widehat{F}$ is G-equivariant.*

A coordinate expression may also be convenient for interpretability. Assume that $W = \mathbb{R}^q$. Let $b = B(\mathcal{O}(u)) \in \mathbb{R}^{d_B}$ and $t = \tau(y) \in \mathbb{R}^{d_T}$, and (C_{ij}^w) be the coefficients of the fusion map. Then,

$$[\widehat{\mathcal{F}}(u)(y)]_w = \sum_{i=1}^{d_B} \sum_{j=1}^{d_T} C_{ij}^w b_i t_j, \qquad w \in \{1, \cdots, q\}.$$

The standard DeepONet with its trivial output-fiber action is recovered from taking $q = 1, d_B = d_T = p$, and $C_{ij}^1 = \delta_{ij}$, giving that $\widehat{\mathcal{F}}(u)(y) = \langle B(\mathcal{O}(u)), \tau(y)\rangle$, where $\langle \cdot, \cdot \rangle$ is the typical Euclidean dot product. In this case, if G acts non-trivially on X and Ω, equivariance would require that $\widehat{\mathcal{F}}(\rho_X(g)u)(y) = \widehat{\mathcal{F}}(u)(g^{-1} \cdot y)$ for each $g \in G$. A standard DeepONet does not generally satisfy this condition without equivariance constraints imposed on the observation, branch, and trunk maps. The Euclidean dot-product fusion itself, however, is compatible with a non-trivial group action when the branch and trunk features transform under the same orthogonal representation ρ since $\langle \rho(g)b, \rho(g)t\rangle = \langle b, t\rangle$. The non-trivial requirement is not the pairing itself but that the branch and trunk outputs obey these matching transformation laws. A standard DeepONet does not enforce this: its branch and trunk are independent, unconstrained networks, so their latent features need not transform under the same orthogonal representation, or even under a well-defined representation at all. Consequently, equivariance of the full architecture typically fails unless these transformation laws, together with equivariance of the observation map, are imposed explicitly. For $W = \mathbb{R}^q$ with $q > 1$, the more general fusion condition is required since $\rho_W(g)$ may mix the output features.

It can also be convenient to combine the trunk features and the fusion map together into a single operator-valued trunk. For $y \in \Omega$, define

$$T(y) : V_B \to W, \qquad T(y)b = C(b, \tau(y)).$$

Note that $T(y)$ is linear in b by bilinearity of C. The prior equivariance conditions imply that

$$T(g \cdot y) \circ \rho_B(g) = \rho_W(g) \circ T(y),$$

or, equivalently, the covariant statement

$$T(g \cdot y) = \rho_W(g) \circ T(y) \circ \rho_B(g)^{-1}.$$

The same relation may be viewed as ordinary equivariance into the representation space $\mathrm{Hom}(V_B, W)$, equipped with

$$\rho_{\mathrm{Hom}}(g)A = \rho_W(g) \circ A \circ \rho_B(g)^{-1}. \tag{2.3}$$

Conversely, an operator-valued trunk satisfying the covariance condition is a special case of the bilinear formulation where $V_T = \mathrm{Hom}(V_B, W), \tau = T$, and fusion is the evaluation map $C_{\mathrm{ev}}(b, A) = A(b)$. See Appendix B for details. The distinction between the two frameworks is simply if one wishes to separate out the trunk and fusion maps or package them together.

In practice, B and T are parameterized by finite-dimensional neural networks. The representation spaces specify how their features transform under G, while the network layers are constructed to respect these actions. Equivariant linear layers are intertwiners between the corresponding representation spaces, and equivariant nonlinearities are built from invariant gating or symmetry-compatible tensor products. Thus, the equivariance conditions are enforced by construction.

2.2 Finite Sensing, Equivariance Preservation, and Approximation Theory

The construction above assumes an equivariant finite-dimensional observation map $\mathcal{O} : X \to V_{\mathrm{obs}}$. In practice, a DeepONet branch receives finitely many measurements of the input field, which may be sparse. A generic collection need not automatically have the given equivariance property. Here, we consider the case of such a finite set realizing the full observation map and consider when these measurements generate an exactly equivariant observation.

For fields with well-defined point values, a fixed finite family of sensors $\Xi = \{x_i\}_{i=1}^m$ defines a sampling map $\mathcal{S}_\Xi u = (u(x_1), \ldots, u(x_m))$. For a continuous spatial group, this map is generally not equivariant. The transformed samples $g^{-1}x_i$ need not belong to the original family. However, there are two exact constructions relevant here.

The first case occurs when the admissible inputs lie in a finite-dimensional G-invariant subspace $E \subset X$ and the samples uniquely determine every element of E. The sampling map restricted to E is then injective, so the coefficient vector of $u \in E$ in a chosen basis can be reconstructed exactly from the samples. Since E is G-invariant, these coefficients carry the corresponding finite-dimensional representation of G. Thus, although the raw sensor values need not themselves carry a closed group action, exact reconstruction yields an equivariant representation-valued observation.

For the second case, suppose G preserves the sensor set, in the sense that transforming the field only permutes the locations at which its samples are evaluated (and, for vector-valued fields, also transforms the sampled values). Consequently, $\mathcal{S}_\Xi$ is already an equivariant finite-dimensional observation map in this case, and no reconstruction of the complete field is required. In other words, if the group only moves sensor locations to other sensor locations, then transforming a field and sampling it is equivalent to sampling it first and then permuting/transforming the samples.

We summarize these notions in an informal theorem containing these observations, with the precise technical statements and proofs being located in Appendix B.

Theorem 2.2 ((Informal) Equivariance from finite observations)**.** *Finite sensor measurements define an exactly equivariant observation whenever either*

1. *The admissible inputs belong to a finite-dimensional invariant subspace and are uniquely determined by the measurements, so that an equivariant finite-dimensional representation can be reconstructed exactly, or*
2. *The symmetry group preserves the finite sensor set, so that its action induces a representation directly on the sensor measurements.*

The precise statements and proofs of both parts of this theorem are located in Appendix B, as well as a stability estimate in terms of the smallest singular value of A from Theorem B.1. It also contains a statement and proof of a separate operator approximation theorem. Once again, we provide an informal statement, with the precise statement and proof being given in Appendix B.

Theorem 2.3 ((Informal) Finite-rank equivariant approximation)**.** *A continuous equivariant operator on a compact, symmetry-invariant input set can be approximated arbitrarily well by equivariant DeepONets with finite-dimensional observations and finite-rank outputs, provided the output space admits a dense family of finite-dimensional invariant subspaces.*

Table 3.1: The benchmark groups and their actions

Problem	Group	Action
Allen–Cahn	$\mathrm{SO}(3) \times \mathbb{Z}_2$	$((R, s) \cdot u)(x) = s\, u(R^{-1}x)$
Shallow water	$\mathrm{SO}(3)$	$R \cdot (\eta, v)(x) = \big(\eta(R^{-1}x), Rv(R^{-1}x)\big)$
Thermoelasticity	$\mathrm{O}(2)$	$Q \cdot (\theta, \mathbf{u}, \mathbf{v})(x) = \big(\theta(Q^{-1}x), Q\mathbf{u}(Q^{-1}x), Q\mathbf{v}(Q^{-1}x)\big)$
Gross–Pitaevskii	$\mathrm{U}(2) \times D_2$	$((U, g) \cdot \Psi)(x) = U\Psi(g^{-1}x)$

2.3 Equivariant Linear and Nonlinear Blocks

The implementations of the branch and trunk networks use equivariant linear maps and nonlinearities. These linear layers admit two equivalent descriptions. In an analytic Schur construction, features are decomposed into irreducible representations and trainable maps act on multiplicity spaces. In a generator-nullspace construction, an admissible weight tensor W is parameterized by the solutions of

$$d\rho_{\text{out}}(A_k)W - W d\rho_{\text{in}}(A_k) = 0$$

for the Lie-algebra generators A_k, together with finite constraints for disconnected components (see Finzi et al. (2021)). These are two parameterizations of the same intertwining requirement. The former is easier if a known Schur decomposition exists and is exact to numerical precision, though the latter produces highly similar results with a one-time upfront cost of solving a system of equations. In our experiments, the results were equivalent up to precision of the results displayed here.

Arbitrary coordinate-wise activations do not generally commute with a non-trivial irreducible representation, complicating the ability to equivariantly perform nonlinear mixing. In this work, we perform mixing via tensor products. Here, products of representation-valued features are projected via fixed intertwiners. Depending on the representation, these are implemented either by Clebsch–Gordan contractions or an efficient Gaunt-product construction. This permits direct nonlinear coupling between representation types while preserving equivariance. For $\mathrm{U}(2)$, invariant Hermitian Gram products are instead used to modulate a third doublet feature, yielding an equivariant cubic coupling.

One may also consider invariant gating, where invariant quantities such as norms are passed through an MLP to produce scalar gates. Empirically, we observed better performance by the tensor product framework for strongly-coupled nonlinear features, which may reflect its ability to directly couple features transforming under different representations. For this reason, we will focus specifically on the former method or combinations thereof, though all are described in Appendix A.3.

3 Benchmark Problems

In order to both provide numerical evidence for the theory and demonstrate the impact of the architectural modifications, we evaluate the efficacy of the equivariant DeepONet in emulating numerical solvers for four distinct families of PDEs. Each problem satisfies different equivariance properties that generate different challenges. Table 3.1 records the groups and their actions on the physical fields. The full equations, symmetry arguments, and data generation procedures for each benchmark are located in Appendix C.

Allen–Cahn on $\mathbb{S}^2$: Allen–Cahn provides the simplest setting amongst the listed experiments, namely scalar phase-field evolution via a reaction-diffusion equation. Rotations of the sphere give an $\mathrm{SO}(3)$ action, while the oddness of the reaction term also preserves the sign symmetry $u \mapsto -u$, yielding $\mathrm{SO}(3) \times \mathbb{Z}_2$ equivariance. Although the state itself is scalar, an equivariant operator must simultaneously respect a continuous geometric symmetry and an independent internal parity. This makes Allen–Cahn a clean setting in which to separate exact symmetry enforcement from the use of spherical geometry alone. Furthermore, the equation is non-linear, meaning that equivariant non-linearity must still be accounted for.

Non-rotating shallow water on $\mathbb{S}^2$: The shallow-water equations model the coupled evolution of fluid height and horizontal velocity on a spherical surface under a hydrostatic approximation. In the absence of planetary rotation, the equations are equivariant under the full $\mathrm{SO}(3)$ action. Unlike

Allen–Cahn, however, the states carry different representations. Namely, the height perturbation is a scalar field, whereas the velocity is a tangent vector field, which satisfy different transformation laws with respect to rotations.

Thermoelasticity on an annulus: Thermoelasticity models the coupled evolution of temperature, displacement, and velocity in an elastic medium. The coupling requires traditional numerical models to resolve physics that occur on drastically different spatiotemporal scales. On the annular domain considered here, rotational and reflection symmetry give an $\mathrm{O}(2)$ action. Once again, the state contains different representation types: temperature is a scalar field, whereas displacement and velocity are vector fields. Hence, they carry different transformation laws. Along with reflection symmetry being necessary to consider here, unlike in the prior problems, another interesting feature of this benchmark is that the underlying dynamics are linear but strongly coupled across thermal and mechanical variables.

Two-component Gross–Pitaevskii on a square: The Gross–Pitaevskii system models the evolution of a two-component complex wavefunction with both internal and spatial symmetries. Component symmetry gives a global $\mathrm{U}(2)$ action on the complex doublet, while the square domain, Dirichlet boundary conditions, and potential term preserve the two coordinate reflections, giving $D_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. Hence, the full symmetry group is $\mathrm{U}(2) \times D_2$. This case is qualitatively different from the preceding geometric examples, since the dominant non-Abelian symmetry acts on the *field values* rather than on spatial coordinates. It does not give a useful harmonic spatial decomposition nor substantial spatial compression, and the finite-observation construction is correspondingly information-limited. Further, the spatial group is *finite*. This problem tests whether exact equivariance remains useful when the principal symmetry organizes internal channels rather than the geometry of the domain.

The datasets consist of distinct trajectories with Gaussian random field initial data. Every trajectory contains an initial condition and 5 subsequent operator-level timesteps (with numerous solver timesteps omitted from the trajectories). We train on all five adjacent one-step transitions and evaluate both one-step prediction and a five-step autoregressive rollout. Throughout the rollout, equivariance metrics are computed to determine how well the models adhere to the physical symmetries. The training/validation/testing splits are by trajectory, with 4096 training, 256 validation, and 512 held-out test trajectories for each experiment. For more on the training set-up and evaluation metrics, see Appendix D.

4 Results

The comparison contains up to five models for each benchmark: a conventional sparse DeepONet, a DeepONet with a symmetry-aware trunk but an unconstrained branch, a full-observation equivariant model, its finite-sensor equivariant counterpart, and a geometry-appropriate FNO reference. The full and finite-sensor equivariant rows have identical trainable parameter counts; the finite observation is implemented by fixed sampling and reconstruction maps. The reported uncertainty is the standard error across held-out examples. We remark that the Gross–Pitaevskii problem does not have a symmetry-aware trunk model, as it does not possess a clean harmonic representation; the group action from $U(2)$ is internal. While the purpose of this work is to, in part, highlight the performance gains from introducing equivariance to DeepONets, we still provide comparisons to symmetry-relevant FNOs due to their prevalence as a high-performing benchmark model within the neural operator literature (and the primary model in the class where equivariance has been studied, albeit non-exhaustively). We underscore that the goal of these experiments is not to provide a comprehensive FNO comparison.

At a high level, full-observation equivariant models have the lowest one-step and five-step error in all four benchmarks. The same is true of their equivariant defects, which remain near machine-ε precision (in FP32), supporting the equivariance of our construction. These models also remain stable across the rollout, retaining a similar order of accuracy throughout. The finite-sensor results support the conclusion of Theorem 2.2: The exactly reconstructive variants recover the same representation-valued input, while the Gross–Pitaevskii sensors preserve equivariance despite information loss.

The sparse DeepONets and symmetry-aware trunks are typically orders of magnitude worse than the equivariant versions across all performance metrics, though they have more parameters in all but the Gross–Pitaevskii case (more on this shortly). The symmetry-aware trunk performs better than

Table 4.1: Held-out predictions, symmetry metrics, and model sizes across each experiment; E_h indicates the relative L^2 error at timestep h; errors are reported as the mean with the SEM shown parenthetically over held-out examples; Eq_h is the symmetry-factor-maximum equivariance defect at step h; lower is better throughout; bold denotes the best value within an experiment.

Model	E_1 rel. L^2	E_5 rel. L^2	Eq_1	Eq_5	Params. (M)
Allen–Cahn					
Vanilla DeepONet	3.4×10^{-1} (2.5×10^{-3})	5.8×10^{-1} (2.1×10^{-3})	4.0×10^{-1}	7.9×10^{-1}	2.628
Symmetry-aware trunk DeepONet	1.7×10^{-1} (1.4×10^{-3})	4.7×10^{-1} (2.5×10^{-3})	3.0×10^{-1}	7.3×10^{-1}	1.331
Equivariant DeepONet (ours)	$\mathbf{2.9\times10^{-3}}$ (4.0×10^{-5})	$\mathbf{9.6\times10^{-3}}$ (4.0×10^{-5})	$\mathbf{8.9\times10^{-7}}$	$\mathbf{1.3\times10^{-6}}$	**0.381**
Finite-sensor equivariant DeepONet (ours)	$\mathbf{2.9\times10^{-3}}$ (4.0×10^{-5})	9.8×10^{-3} (5.0×10^{-5})	2.8×10^{-6}	9.8×10^{-6}	**0.381**
SFNO	5.1×10^{-3} (7.0×10^{-5})	1.2×10^{-2} (3.3×10^{-4})	1.0×10^{-2}	1.6×10^{-2}	0.804
Shallow water					
Vanilla DeepONet	8.9×10^{-1} (6.0×10^{-4})	1.0×10^{0} (1.1×10^{-3})	6.9×10^{-1}	1.1×10^{0}	4.727
Symmetry-aware trunk DeepONet	4.7×10^{-1} (7.0×10^{-4})	1.0×10^{0} (1.0×10^{-3})	1.6×10^{-1}	4.5×10^{-1}	12.788
Equivariant DeepONet (ours)	$\mathbf{6.5\times10^{-3}}$ (8.0×10^{-5})	$\mathbf{2.2\times10^{-2}}$ (1.4×10^{-4})	$\mathbf{1.8\times10^{-6}}$	$\mathbf{2.8\times10^{-6}}$	**0.560**
Finite-sensor equivariant DeepONet (ours)	$\mathbf{6.5\times10^{-3}}$ (8.0×10^{-5})	$\mathbf{2.2\times10^{-2}}$ (1.4×10^{-4})	$\mathbf{1.8\times10^{-6}}$	2.9×10^{-6}	**0.560**
SFNO	6.1×10^{-2} (1.6×10^{-4})	1.6×10^{-1} (7.0×10^{-4})	4.1×10^{-2}	1.3×10^{-1}	1.100
Thermoelasticity					
Vanilla DeepONet	1.3×10^{-1} (7.0×10^{-4})	1.7×10^{-1} (2.2×10^{-3})	8.5×10^{-2}	1.2×10^{-1}	5.775
Symmetry-aware trunk DeepONet	1.3×10^{-1} (8.0×10^{-4})	1.6×10^{-1} (2.7×10^{-3})	7.0×10^{-2}	9.7×10^{-2}	24.126
Equivariant DeepONet (ours)	$\mathbf{2.4\times10^{-3}}$ (1.0×10^{-5})	$\mathbf{5.3\times10^{-3}}$ (5.0×10^{-5})	$\mathbf{3.7\times10^{-7}}$	$\mathbf{1.4\times10^{-6}}$	**1.029**
Finite-sensor equivariant DeepONet (ours)	$\mathbf{2.4\times10^{-3}}$ (1.0×10^{-5})	$\mathbf{5.3\times10^{-3}}$ (5.0×10^{-5})	4.5×10^{-7}	1.9×10^{-6}	**1.029**
Polar FNO	1.7×10^{-2} (4.0×10^{-5})	4.1×10^{-2} (2.2×10^{-4})	1.0×10^{-2}	2.9×10^{-2}	9.747
Gross–Pitaevskii					
Vanilla DeepONet	3.2×10^{-1} (4.0×10^{-4})	3.6×10^{-1} (1.0×10^{-3})	2.0×10^{-1}	2.8×10^{-1}	5.776
Equivariant DeepONet (ours)	$\mathbf{2.1\times10^{-2}}$ (3.0×10^{-5})	$\mathbf{2.3\times10^{-2}}$ (2.4×10^{-4})	$\mathbf{2.3\times10^{-6}}$	$\mathbf{5.0\times10^{-6}}$	13.948
Finite-sensor equivariant DeepONet (ours)	$\mathbf{2.1\times10^{-2}}$ (3.0×10^{-5})	2.4×10^{-2} (9.7×10^{-4})	2.4×10^{-6}	5.1×10^{-6}	13.948
FNO	2.2×10^{-2} (3.0×10^{-5})	2.8×10^{-2} (7.0×10^{-5})	1.1×10^{-2}	1.8×10^{-2}	**5.163**

the vanilla DeepONet, though it does not provide as substantial a jump as the transition from the symmetry-aware trunk to the equivariant model. This provides evidence that a symmetry-aware trunk is not a sufficient substitute for an equivariant observation and branch.

The FNO/SFNO references are substantially more accurate than the non-equivariant DeepONets, though the equivariant DeepONet outperforms despite having smaller parameter counts in all but the Gross–Pitaevskii case. We should note that strong performance of FNOs is expected in these cases, as they are akin to learned pseudospectral methods; vanilla DeepONets would require large latent ranks and/or long training times to capture the spatially-complex, multiscale solution features (especially for the nonlinear benchmarks). Nonetheless, the equivariant DeepONets are able to outperform with (typically) fewer parameters.

The selected full and finite-sensor equivariant models for Allen–Cahn, shallow water, and thermoelasticity are smaller than both the sparse DeepONet and the corresponding spectral reference while obtaining the lowest reported prediction errors. Gross–Pitaevskii is the important exception: the selected equivariant model attains the lowest prediction error while retaining near-machine-precision equivariance, but it does feature more parameters. The internal $\mathrm{U}(2)$ action fails to provide compression of the solution's spatial dependence; the D_2 factor contributes mild compression. Consequently, the DeepONet must compress a spatially-rich complex input through a finite equivariant branch representation while simultaneously representing the output through its finite branch–trunk factorization. In practice, competitive accuracy required increasing the hidden multiplicity of the equivariant representation channels, exposing a capacity challenge for equivariant DeepONets when the available symmetry provides little spatial compression. We must underscore that increasing the capacity of vanilla DeepONets did not yield errors comparable to those of the equivariant models, so this is not merely an artifact of scaling model size.

Figure 4.1 illustrates a qualitative comparison for each model type, using the spherical FNO (SFNO) of Bonev et al. (2023) as the reference FNO model. Similar comparisons are contained in Appendix E for each experiment, though omitted here due to space. Appendix E also contains plots of the mean relative L^2 error and equivariance defects over the autoregressive window for each experiment.

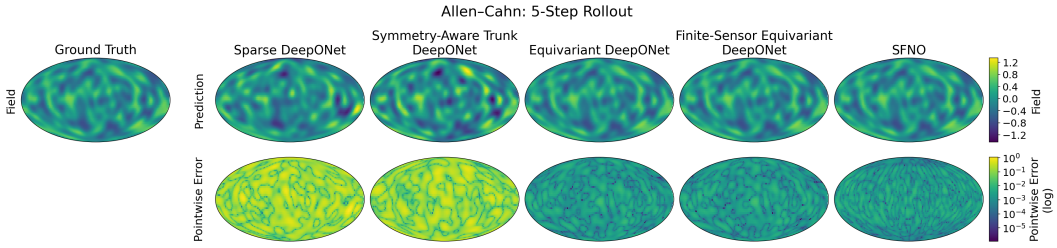


Figure 4.1: Five-step autoregressive rollout of a single spherical Allen-Cahn initial condition for the baseline DeepONet, DeepONet with a symmetry-aware trunk, equivariant DeepONet, finite-sensor equivariant DeepONet, and spherical FNO

It is important to note that the SFNO is designed to respect spherical rotations through its spherical-harmonic parameterization, and its *continuous* scalar-field construction is $\mathrm{SO}(3)$-equivariant. However, this does not imply exact equivariance to the full action after discretization or for heterogeneous field types. Finite spherical transforms, spectral truncation, and pointwise nonlinearities introduce non-zero discrete equivariance defects (this is described explicitly in Bonev et al. (2023)).

5 Discussion and Conclusions

A core conclusion is that equivariance in a DeepONet is a property of the full factorization. The observation, branch, trunk, and fusion maps must satisfy compatible transformation laws, as formalized in Theorem 2.1. In particular, a symmetry-aware trunk does not by itself imply theoretical equivariance of the resulting operator nor learned equivariance. The branch–trunk factorization is compatible with exact equivariance once the appropriate conditions are imposed. Numerical experiments support the equivariance of the construction, which remained small across the symmetry factors and rollout. Such networks were also shown to uniformly approximate continuous equivariant operators on compact G-invariant sets under suitable conditions.

Finite observations require an additional compatibility condition. For continuous spatial groups, a generic fixed sensor set does not necessarily carry the group action. However, exact equivariance can be obtained by reconstructing a finite-dimensional invariant representation from the samples. Alternatively, whenever the spatial group preserves a finite sensor family, the sampled values carry the induced finite-dimensional representation directly. This does not require the acting group itself to be finite, although finite spatial groups provide the most direct examples. The close agreement between the full- and finite-observation models is consistent with these results.

The four benchmarks instantiate different representation structures. Allen–Cahn combines spherical $\mathrm{SO}(3)$ representations with a fiberwise sign action $\mathbb{Z}_2$. Shallow water couples scalar and tangent-vector types under $\mathrm{SO}(3)$. Thermoelasticity requires both rotational and reflection structure under $\mathrm{O}(2)$ on scalar and vector components. Gross–Pitaevskii contains an internal $\mathrm{U}(2)$ action together with a finite spatial reflection group D_2. The same abstract construction applies in each case, although the corresponding finite-dimensional representations differ.

In summary, we have constructed an equivariant DeepONet by imposing compatible group actions throughout the observation, branch, trunk, and fusion maps. We have also identified two exact mechanisms for finite equivariant observations and shown that finite-rank equivariant approximation is compatible with the DeepONet factorization. The numerical examples demonstrate the construction for continuous rotations, mixed scalar and vector fields, disconnected reflection groups, and internal unitary symmetries with favorable performance over baseline DeepONets and appropriately-chosen FNOs. This highlights the power of injecting physical symmetries into DeepONet architectures for improved performance with (typically) reduced model complexity.

For future work, natural extensions include quantifying when physical symmetries reduce finite-rank representation complexity, extending the finite-observation theory to noisy and approximately band-limited inputs, and developing equivariant DeepONets for settings in which the governing dynamics possess only partial or approximate symmetry.

Reproducibility Statement

The appendices state the representation conventions and assumptions, record the explicit PDEs formulations, describe the dataset generation, outline the particulars of the representation-theoretic instantiations, and specify the training protocol, model parameterization, and metric formulas.

AI Use Statement

Generative AI assistants were used to support with optimization of code for the equivariant models (e.g., speed-ups in the Clebsch–Gordan/Gaunt implementation); reviewing code implementations (e.g., the solvers for the thermoelastic and Gross–Pitaevskii systems); and initial drafting, revision recommendations, and proofreading of the manuscript. The latter includes assistance with mathematical formalization, particularly in the technical framing of specific group representation components of the architecture instantiations in Appendix C and the framing of Theorem B.5. It also provided an initial proof attempt of the latter result, which was then heavily edited by the authors. The authors have checked all details of the mathematical statements, generated code, and writing recommendations. The authors take responsibility for all the content in the paper.

References

Boris Bonev, Thorsten Kurth, Christian Hundt, Jaideep Pathak, Maximilian Baust, Karthik Kashinath, and Anima Anandkumar. Spherical fourier neural operators: Learning stable dynamics on the sphere. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 2806–2823. PMLR, 2023.

Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks*, 6(4):911–917, 1995. doi: 10.1109/72.392253.

Chaoran Cheng and Jian Peng. Equivariant neural operator learning with graphon convolution. In *Advances in Neural Information Processing Systems*, volume 36, pp. 61960–61984, 2023.

Taco S. Cohen and Max Welling. Group equivariant convolutional networks. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 2990–2999. PMLR, 2016.

David Erzmann and Sören Dittmer. Equivariant neural operators for gradient-consistent topology optimization. *Journal of Computational Design and Engineering*, 11(3):91–100, 2024. doi: 10.1093/jcde/qwae039.

Marc Finzi, Max Welling, and Andrew Gordon Wilson. A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3318–3328. PMLR, 2021.

Gerald B. Folland. *A Course in Abstract Harmonic Analysis*. CRC Press, Boca Raton, Florida, 1995.

William Fulton and Joe Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics*. Springer, New York, 1991.

E. P. Gross. Structure of a quantized vortex in boson systems. *Il Nuovo Cimento*, 20(3):454–477, 1961. doi: 10.1007/BF02731494.

Jiequn Han, Xu-Hui Zhou, and Heng Xiao. An equivariant neural operator for developing nonlocal tensorial constitutive models. *Journal of Computational Physics*, 488:112243, 2023. doi: 10.1016/j.jcp.2023.112243.

Jacob Helwig, Xuan Zhang, Cong Fu, Jerry Kurtin, Stephan Wojtowytsch, and Shuiwang Ji. Group equivariant fourier neural operators for partial differential equations. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 12907–12930. PMLR, 2023.

Hanxun Jin, Boyu Zhang, Qianying Cao, Enrui Zhang, Aniruddha Bora, Sridhar Krishnaswamy, George E. Karniadakis, and Horacio D. Espinosa. Characterization and inverse design of stochastic mechanical metamaterials using neural operators. *Advanced Materials*, 37(29):e2420063, 2025. doi: 10.1002/adma.202420063.

Elham Kiyani, Manav Manav, Nikhil Kadivar, Laura De Lorenzis, and George Em Karniadakis. Predicting crack nucleation and propagation in brittle materials using deep operator networks with diverse trunk architectures. *Computer Methods in Applied Mechanics and Engineering*, 441: 117984, 2025.

Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2747–2755. PMLR, 2018.

Jean Kossaifi, Nikola Kovachki, Zongyi Li, David Pitt, Miguel Liu-Schiaffini, Robert J. George, Boris Bonev, Kamyar Azizzadenesheli, Julius Berner, Valentin Duruisseaux, and Anima Anandkumar. A library for learning neural operators. *Journal of Machine Learning Research*, 27(205):1–6, 2026.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.

Samuel Lanthaler, Siddhartha Mishra, and George E. Karniadakis. Error estimates for deeponets: A deep learning framework in infinite dimensions. *Transactions of Mathematics and Its Applications*, 6(1):tnac001, 2022. doi: 10.1093/imatrm/tnac001.

Samuel Lanthaler, Roberto Molinaro, Patrik Hadorn, and Siddhartha Mishra. Nonlinear reconstruction for operator learning of pdes with discontinuities. In *International Conference on Learning Representations*, 2023.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3:218–229, 2021. doi: 10.1038/s42256-021-00302-5.

Lu Lu, Xuhui Meng, Shengze Cai, Zhiping Mao, Somdatta Goswami, Zhongqiang Zhang, and George Em Karniadakis. A comprehensive and fair comparison of two neural operators (with practical extensions) based on fair data. *Computer Methods in Applied Mechanics and Engineering*, 393:114778, 2022.

Shengjie Luo, Tianlang Chen, and Aditi Krishnapriyan. Enabling efficient equivariant operations in the fourier basis via gaunt tensor products. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=mhyQXJ6JsK.

Witold Nowacki. *Thermoelasticity*. Pergamon Press, 2 edition, 1986. ISBN 978-0-08-024767-0. doi: 10.1016/C2013-0-03247-1.

Ahmad Peyvan, Varun Kumar, and George Em Karniadakis. Fusion-deeponet: A data-efficient neural operator for geometry-dependent hypersonic and supersonic flows. *Journal of Computational Physics*, pp. 114432, 2025.

L. P. Pitaevskii. Vortex lines in an imperfect bose gas. *Soviet Physics JETP*, 13(2):451–454, 1961.

Zakhar Shumaylov, Peter Zaika, James Rowbottom, Ferdia Sherry, Melanie Weber, and Carola-Bibiane Schönlieb. Lie algebra canonicalization: Equivariant neural operators under arbitrary lie groups. In *International Conference on Learning Representations*, 2025.

Gilbert Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968. doi: 10.1137/0705041.

Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum Theory of Angular Momentum*. World Scientific, Singapore, 1988.

Simone Venturi and Tiernan Casey. Svd perspectives for augmenting deeponet flexibility and interpretability. *Computer Methods in Applied Mechanics and Engineering*, 403:115718, 2023. doi: 10.1016/j.cma.2022.115718.

Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco S. Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. In *Advances in Neural Information Processing Systems*, volume 31, pp. 10402–10413, 2018.

Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, and Gabriel J. Brostow. Harmonic networks: Deep translation and rotation equivariance. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7168–7177, 2017. doi: 10.1109/CVPR.2017.758.

Jiaxiao Xu, Changhong Mou, Yeyu Zhang, and Fengxiang He. Physics-aligned canonical equivariant fourier neural operator under symmetry-induced shifts. *arXiv preprint arXiv:2605.18606*, 2026. doi: 10.48550/arXiv.2605.18606.

Dmitry Yarotsky. Universal approximations of invariant maps by neural networks. *Constructive Approximation*, 55(1):407–474, 2022. doi: 10.1007/s00365-021-09546-1.

CONTENTS

1 Introduction **1**

2 Equivariant Deep Operator Networks **3**

2.1 Operator Equivariance and the DeepONet Factorization 3

2.2 Finite Sensing, Equivariance Preservation, and Approximation Theory 5

2.3 Equivariant Linear and Nonlinear Blocks 6

3 Benchmark Problems **6**

4 Results **7**

5 Discussion and Conclusions **9**

A Requisite Theory **15**

A.1 Groups, Representation Spaces, and Intertwiners 15

A.2 Field Representations and Typed States 15

A.3 Equivariant Nonlinearities and Tensor-Product Couplings 16

A.4 Key Groups and Their Representations 18

A.5 Neural Operators, DeepONets, and FNOs 20

A.6 Equivariant Multilayer Perceptrons 20

B Abstract Construction and Proofs **22**

B.1 Equivariance of the Branch–Trunk Factorization 22

B.2 Unisolvent Finite Sensors 23

B.3 Group-Closed Sensor Families 25

B.4 Finite-Rank Equivariant Approximation 25

C Benchmarks: PDE Formulations, Model Realizations, and Dataset Construction **30**

C.1 Allen–Cahn Equation on $\mathbb{S}^2$ 30

C.2 Nonrotating Shallow-Water Equations on $\mathbb{S}^2$ 31

C.3 Linear Thermoelasticity on an Annulus 33

C.4 Two-Component Gross–Pitaevskii System 35

C.5 Dataset and Observation Summaries 38

C.6 Model Parameterization Summary 38

D Training and Evaluation **38**

D.1 Training Setup 38

D.2 Prediction Metrics 39

D.3 Equivariance Metrics 41

E Additional Figures **43**

A Requisite Theory

This appendix records the group- and representation-theoretic conventions used throughout the paper. Standard references include Fulton & Harris (1991), Folland (1995), and Varshalovich et al. (1988).

A.1 Groups, Representation Spaces, and Intertwiners

Definition A.1 (Groups and group actions)**.** A *group* is a set G equipped with an associative product, an identity element e, and inverses. A *left action* of G on a set M is a map $(g, x) \mapsto g \cdot x$ satisfying

$$e \cdot x = x, \qquad (g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x).$$

Definition A.2 (Representations)**.** A *representation* of G on a real or complex vector space V is a homomorphism

$$\rho_V : G \to \mathrm{GL}(V).$$

We refer to the pair (V, ρ_V) as a *representation space*. A subspace $U \subset V$ is invariant when $\rho_V(g)U \subset U$ for every $g \in G$. The representation is *irreducible* when its only invariant subspaces are $\{0\}$ and V.

Definition A.3 (Equivariant maps and intertwiners)**.** Let (V, ρ_V) and (Z, ρ_Z) be representation spaces. A map $f : V \to Z$, not necessarily linear, is *equivariant* if

$$f \circ \rho_V(g) = \rho_Z(g) \circ f, \qquad g \in G.$$

If f is linear, it is called an *intertwiner*. The space of linear intertwiners is denoted $\mathrm{Hom}_G(V, Z)$.

If G is compact, any finite-dimensional representation may be equipped with a G-invariant inner product. Hence, we will take the finite-dimensional representations below to be orthogonal or unitary. They decompose into isotypic components,

$$V \cong \bigoplus_{\lambda \in \widehat{G}} M_\lambda \otimes V_\lambda, \tag{A.1}$$

where $\widehat{G}$ is the unitary dual of G (the set of equivalence classes of irreducible representations of G), V_λ is an irreducible representation space, and M_λ is its multiplicity space. The group acts trivially on M_λ and through ρ_λ on V_λ.

Schur's lemma implies that a linear intertwiner cannot mix inequivalent irreducible types. In the real-type case, its restriction to an isotypic component has the form

$$L\big|_{M_\lambda \otimes V_\lambda} = L_\lambda \otimes I_{V_\lambda}. \tag{A.2}$$

Thus, the trainable part acts on multiplicity coordinates, while the irrep coordinate carries the prescribed group action. For real representations of complex or quaternionic type, a small fixed commutant accompanies the multiplicity map. The $\mathrm{O}(2)$ and $\mathrm{U}(2)$ constructions below make the relevant complex structures explicit.

If (V, ρ_V) and (Z, ρ_Z) are representation spaces, then $V \otimes Z$ carries the tensor-product representation

$$(\rho_V \otimes \rho_Z)(g)(v \otimes z) = \rho_V(g)v \otimes \rho_Z(g)z.$$

For a product group $G_1 \times G_2$, representations V_1 of G_1 and V_2 of G_2 define the external tensor-product representation $V_1 \boxtimes V_2$ by

$$(\rho_1 \boxtimes \rho_2)(g_1, g_2) = \rho_1(g_1) \otimes \rho_2(g_2).$$

A.2 Field Representations and Typed States

Definition A.4 (Induced field actions)**.** Let G act on a domain M, and let (W, ρ_W) be a representation space. The induced action on W-valued fields is

$$[\rho_Y(g)f](x) = \rho_W(g) f(g^{-1} \cdot x).$$

The inverse in the argument follows from the left-action convention. A scalar field has a trivial action on its values, so only the argument is pulled back. A vector field has both a pullback in its argument and a nontrivial action on its value. For example, a Euclidean vector field transforms as

$$[\rho(Q)v](x) = Qv(Q^{-1}x), \qquad Q \in \mathrm{O}(d).$$

Definition A.5 (Typed states)**.** A field is *typed* when its physical components occupy different representation spaces. If the state has components in $W_1, \dots, W_r$, then the output fiber is the direct sum

$$W = W_1 \oplus \cdots \oplus W_r,$$

with the block action $\rho_W = \rho_{W_1} \oplus \cdots \oplus \rho_{W_r}$.

The shallow-water state is the clearest example: height is scalar, whereas velocity is tangent vector valued. Thermoelasticity combines a scalar temperature with two vector fields. In Gross–Pitaevskii, the spatial action is accompanied by an internal action on the complex component doublet. These actions must be represented separately; treating stored coordinate channels as unrelated scalars gives the wrong transformation law.

For a product of an internal group G_{int} and a spatial group G_{sp}, a common field action is

$$[\rho(g_{\mathrm{int}}, g_{\mathrm{sp}})f](x) = \rho_{\mathrm{int}}(g_{\mathrm{int}})\rho_{\mathrm{fib}}(g_{\mathrm{sp}})f(g_{\mathrm{sp}}^{-1}x).$$

The two factors commute in the Gross–Pitaevskii model considered here.

A.3 Equivariant Nonlinearities and Tensor-Product Couplings

A coordinate-wise scalar activation generally fails to commute with a nontrivial irreducible representation. Implementations may instead use invariant gating and tensor-product projection (or a bilinear layer, as in Finzi et al. (2021), which we do not consider here).

Invariant Gating. Suppose $z_1, \dots, z_K$ are copies of a unitary representation space V which transform under an action ρ_V. The pairwise inner products

$$\mathcal{G}_{ab}(z) = \langle z_a, z_b \rangle_V$$

are invariant under the action $z_a \mapsto \rho_V(g)z_a$. If scalar functions γ_a depend only on such invariants, then

$$z_a \mapsto \gamma_a(\mathcal{G}(z))z_a$$

is equivariant. In practice, the functions γ_a may be parameterized by an MLP acting only on the invariant quantities. A special case is norm gating, in which the gate depends only on invariant feature norms $\|z_a\|_V$ (i.e., we take the square roots of the diagonal of the Gram matrix $\mathcal{G}$). See Worrall et al. (2017); Weiler et al. (2018) for related constructions.

In the spherical models, the implemented gate computes the root-mean-square coefficient magnitude separately for each spherical degree and multiplicity channel, passes the collection of invariant scalars through an MLP, and applies one scalar gate to all $2\ell + 1$ basis components in the corresponding copy of V_ℓ. In the $\mathrm{O}(2)$ model, the magnitude of each complex Fourier coefficient is invariant under rotations and reflections and is gated channelwise. In the $\mathrm{U}(2) \times D_2$ model, norms are taken over the real and imaginary coordinates of each complex doublet.

Clebsch–Gordan Products. To introduce more complex nonlinear interactions while preserving equivariance, we use tensor products. If $z_\alpha \in V_\alpha$ and $z_\beta \in V_\beta$, then $z_\alpha \otimes z_\beta$ contains their pairwise bilinear interactions and transforms under the tensor-product representation $\rho_\alpha \otimes \rho_\beta$. Decomposing this tensor product into irreducible components identifies the representation types that can be produced equivariantly.

To that end, let V_α and V_β be irreducible representation spaces. Their tensor product decomposes as

$$V_\alpha \otimes V_\beta \cong \bigoplus_\gamma M_{\alpha\beta}^\gamma \otimes V_\gamma.$$

That is, after a change of basis, the tensor product splits into invariant subspaces that transform as irreducible representation spaces V_γ. For each output type V_γ, a Clebsch–Gordan map $C_{\alpha\beta}^{\gamma,r}$:

$V_\alpha \otimes V_\beta \to V_\gamma$ is a projection map which selects one copy, indexed by r. Since the subspace is invariant, this map is an intertwiner, i.e.

$$C_{\alpha\beta}^{\gamma,r}(\rho_\alpha(g) \otimes \rho_\beta(g)) = \rho_\gamma(g) C_{\alpha\beta}^{\gamma,r}, \qquad g \in G.$$

The index r distinguishes the different copies of V_γ. Therefore, $(z_\alpha, z_\beta) \mapsto C_{\alpha\beta}^{\gamma,r}(z_\alpha \otimes z_\beta)$ is an equivariant bilinear map whose output lives in V_γ. This defines the equivariant rule with which we may mix representations to obtain new representations.

For $\mathrm{SO}(3)$ specifically, the degree-ℓ irreducible representation space V_ℓ satisfies

$$V_{\ell_1} \otimes V_{\ell_2} \cong \bigoplus_{\ell=|\ell_1-\ell_2|}^{\ell_1+\ell_2} V_\ell.$$

Each admissible V_ℓ appears with multiplicity one, so the index r is unnecessary in this case. In a spherical-harmonic basis, the corresponding Clebsch–Gordan contraction is

$$c_{\ell m} = \sum_{m_1,m_2} C_{\ell_1 m_1, \ell_2 m_2}^{\ell m} a_{\ell_1 m_1} b_{\ell_2 m_2}.$$

If a_{ℓ_1} and b_{ℓ_2} transform under Wigner matrices, $D^{\ell_1}(R)$ and $D^{\ell_2}(R)$, then c_ℓ transforms under $D^\ell(R)$. This is the "angular-momentum" addition rule used by the spherical models.

The tensor products introduce nonlinear interactions between representation-valued features. The Clebsch–Gordan contractions are fixed projections onto symmetry-compatible outputs, and a neural network may be used to learn how these contractions are combined. In particular, each irreducible representation may occur in multiple copies. If $z_{\alpha,a} \in V_\alpha$ denotes the a'th copy, a learned tensor-product coupling may take the form

$$[\mathcal{Q}(z,z)]_{\gamma,c} = \sum_{\alpha,\beta} \sum_{a,b} \sum_{r} W_{(\alpha,a),(\beta,b),r}^{\gamma,c} \, C_{\alpha\beta}^{\gamma,r} \left(z_{\alpha,a} \otimes z_{\beta,b}\right),$$

where a, b, c index feature copies, r indexes the multiplicity of V_γ in $V_\alpha \otimes V_\beta$, the Clebsch–Gordan maps $C_{\alpha\beta}^{\gamma,r}$ are fixed, and the coefficients $W_{(\alpha,a),(\beta,b),r}^{\gamma,c}$ are learned. Thus, the representation theory determines which bilinear couplings are allowed, while the network learns how to combine them. For more on Clebsch–Gordan products in equivariant neural networks, see (Thomas et al., 2018).

Gaunt Projection. For scalar spherical harmonics, the coefficient of a pointwise product fg is

$$\widehat{(fg)}_{\ell m} = \sum_{\ell_1,m_1} \sum_{\ell_2,m_2} \mathcal{G}_{\ell_1 m_1, \ell_2 m_2}^{\ell m} \widehat{f}_{\ell_1 m_1} \widehat{g}_{\ell_2 m_2}, \tag{A.3}$$

where

$$\mathcal{G}_{\ell_1 m_1, \ell_2 m_2}^{\ell m} = \int_{\mathbb{S}^2} Y_{\ell_1 m_1}(x) Y_{\ell_2 m_2}(x) \overline{Y_{\ell m}(x)} \, \mathrm{d}x$$

are the *Gaunt* coefficients. Up to normalization and basis conventions, these coefficients are proportional to a product of Clebsch–Gordan coefficients. Thus, projection of a scalar pointwise product onto retained spherical modes is a particular $\mathrm{SO}(3)$-equivariant bilinear intertwiner. For more on the efficient construction of the Gaunt coefficients, see Luo et al. (2024).

The spherical implementation used in this work does not store the dense Gaunt tensor in (A.3). Instead, spherical harmonic coefficients are transformed to an oversampled spatial grid, pointwise products are formed there, and the result is transformed back to spectral coefficients. This provides an efficient, dealiased evaluation of the Gaunt contraction which does not explicitly form the full tensor. In the shallow-water model, the height coefficients and two tangent vector coefficients each transform as a copy of V_ℓ at a fixed degree ℓ, and the scalar Gaunt product gives an equivariant manner to couple these features, though only those arising from pointwise multiplication rather than every possible bilinear equivariant map.

The $\mathrm{O}(2)$ Gaunt block is the one-dimensional Fourier analogue: retained angular modes are transformed to an oversampled angular grid, multiplied pointwise, and transformed back to Fourier coefficients. Under a reflection, each feature has either even or odd parity, and the parity product is defined by sign multiplication.

Cubic Gram Coupling for $\mathrm{U}(2)$. The methodology applied to Gross–Pitaevskii does not utilize a Clebsch–Gordan/Gaunt product. Under the standard $\mathrm{U}(2)$ internal action on $\mathbb{C}^2$, the tensor product $\mathbb{C}^2 \otimes \mathbb{C}^2$ crucially does not contain another copy of $\mathbb{C}^2$. Equivalently, there is no non-zero $\mathrm{U}(2)$-equivariant bilinear map from $\mathbb{C}^2 \times \mathbb{C}^2$ to $\mathbb{C}^2$. Instead, we use the following cubic construction: If $z_a, z_b, z_c \in \mathbb{C}^2$ transform by the standard $\mathrm{U}(2)$ action $z \mapsto Uz$, then $z_a^\dagger z_b$ is invariant under $U \in \mathrm{U}(2)$, and so $(z_a^\dagger z_b) z_c$ transforms in the same manner as z_c. The accompanying D_2 reflection parities in the Gross-Pitaevskii symmetry group are combined by multiplying their signs.

A.4 Key Groups and Their Representations

The sign group $\mathbb{Z}_2$: We identify $\mathbb{Z}_2$ with $\{+1, -1\}$ under multiplication. It has two real irreducible representation spaces: the trivial character $\chi_+(s) = 1$ and the sign character $\chi_-(s) = s$. Their tensor products multiply parities:

$$\chi_- \otimes \chi_- = \chi_+, \qquad \chi_+ \otimes \chi_- = \chi_-.$$

If the input and output representations contain $k_\pm^{\mathrm{in}}$ and $k_\pm^{\mathrm{out}}$ copies of $\chi_\pm$, respectively, then every $\mathbb{Z}_2$-equivariant linear map is block diagonal with respect to this decomposition, i.e. it is of the form $\mathrm{diag}(A_+, A_-)$, where $A_\pm \in \mathbb{R}^{k_\pm^{\mathrm{out}} \times k_\pm^{\mathrm{in}}}$.

The rotation group $\mathrm{SO}(3)$: The group

$$\mathrm{SO}(3) = \{R \in \mathbb{R}^{3\times 3} : R^\top R = I,\ \det R = 1\}$$

acts on $\mathbb{S}^2$ by rotations. Its finite-dimensional irreducible representation spaces are indexed by $\ell \in \mathbb{N}_0$ and have dimension $2\ell + 1$. A standard realization is given by the degree-ℓ spherical harmonics $(Y_{\ell m})_{m=-\ell}^{\ell}$, defined as eigenfunctions of the Laplace-Beltrami operator on $\mathbb{S}^2$, i.e.

$$-\Delta_{\mathbb{S}^2} Y_{\ell m} = \ell(\ell+1) Y_{\ell m}.$$

For a fixed ℓ, these span an $\mathrm{SO}(3)$-invariant irreducible space V_ℓ. In a complex basis, these transform as

$$Y_{\ell m}(R^{-1}x) = \sum_{m'=-\ell}^{\ell} D^\ell_{mm'}(R) Y_{\ell m'}(x),$$

where $D^\ell(R)$ is the Wigner D-matrix. In other words, $\rho_\ell(R) = D^\ell(R)$. Our implementations use an equivalent real orthonormal basis. We choose the convention $Y_\ell(Rx) = D^\ell(R) Y_\ell(x)$, so the coefficient column transforms by $D^\ell(R)$ and the real coefficient–trunk pairing is invariant under the simultaneous action. A scalar field bandlimited through degree L has coefficient space $\bigoplus_{\ell=0}^{L} V_\ell$.

Tangent vector fields on $\mathbb{S}^2$ decompose into spheroidal (gradient-type) and toroidal (curl-type) vector spherical harmonics. For every $\ell \geq 1$, the spheroidal and toroidal coefficient families each transform as a copy of the same irreducible representation V_ℓ. There is no tangent vector $\ell = 0$ mode.

More generally, suppose an input contains k_ℓ^{in} copies of V_ℓ and an output contains k_ℓ^{out} copies. An $\mathrm{SO}(3)$-equivariant linear map may mix these copies arbitrarily, but it must apply the same scalar mixing to every coordinate within V_ℓ. Its degree-ℓ block has the form $A_\ell \otimes I_{V_\ell}$ where $A_\ell \in \mathbb{R}^{k_\ell^{\mathrm{out}} \times k_\ell^{\mathrm{in}}}$ mixes the copies. Equivalently, if $z_{\ell,a} \in V_\ell$ denotes the a'th input copy, then

$$z'_{\ell,b} = \sum_{a=1}^{k_\ell^{\mathrm{in}}} (A_\ell)_{ba} z_{\ell,a}.$$

The orthogonal group $\mathrm{O}(2)$: The group

$$\mathrm{O}(2) = \{Q \in \mathbb{R}^{2\times 2} : Q^\top Q = I\}$$

contains rotations and reflections and may be written as $\mathrm{SO}(2) \rtimes \mathbb{Z}_2$. The real irreducible representations may be organized by angular frequency m. At $m = 0$, the two one-dimensional irreducible representations are the trivial and determinant characters (i.e. $\chi_e(Q) = 1$ and $\chi_{\det}(Q) = \det(Q)$). For $m \geq 1$, the Fourier pair $(\cos(m\theta), \sin(m\theta))$ spans a two-dimensional irreducible representation.

Let R_α denote rotation by an angle α and κ denote the reflection map $\theta \mapsto -\theta$. Under the induced action on functions, R_α maps $f(\theta)$ to $f(\theta - \alpha)$. In the Fourier basis, one may write

$$\rho_{m,p}(R_\alpha) = \begin{pmatrix} \cos(m\alpha) & -\sin(m\alpha) \\ \sin(m\alpha) & \cos(m\alpha) \end{pmatrix}, \qquad \rho_{m,p}(\kappa) = p \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $p \in \{-1, 1\}$ specifies the reflection choice. For $m \geq 1$, the two reflection conventions are equivalent. In the chosen bases, an intertwiner between equal conventions is a scalar multiple of I_2, while an intertwiner between opposite conventions is a scalar multiple of

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Suppose there are k_{in} input copies and k_{out} output copies at a fixed frequency m. Between copies with the same reflection convention, an equivariant linear map has the form $A_{\text{same}} \otimes I_2$, where $A_{\text{same}} \in \mathbb{R}^{k_{\text{out}} \times k_{\text{in}}}$. Between copies with opposite reflection conventions, it has the form $A_{\text{opp}} \otimes J$, where $A_{\text{opp}} \in \mathbb{R}^{k_{\text{out}} \times k_{\text{in}}}$. The matrices A_{same} and A_{opp} mix the copy indices, while I_2 and J give the corresponding intertwiner on the two Fourier coordinates. At $m = 0$, the trivial and determinant characters are inequivalent, so no nonzero intertwiner connects them.

The unitary group $\mathrm{U}(2)$: The group

$$\mathrm{U}(2) = \{U \in \mathbb{C}^{2\times 2} : U^\dagger U = I_2\}$$

acts on $\mathbb{C}^2$ by its standard representation $z \mapsto Uz$. This representation is irreducible. Hence, by Schur's lemma, a complex-linear map $L : \mathbb{C}^2 \to \mathbb{C}^2$ commutes with every $U \in \mathrm{U}(2)$ if and only if

$$L = \lambda I_2, \qquad \lambda \in \mathbb{C}.$$

More generally, for several copies of $\mathbb{C}^2$, equivariant linear maps may mix copies with arbitrary complex coefficients while acting as the identity on the $\mathbb{C}^2$ representation coordinate. Thus, between K_{in} input copies and K_{out} output copies, they have the form

$$L = A \otimes I_2, \qquad A \in \mathbb{C}^{K_{\text{out}} \times K_{\text{in}}}.$$

Any real-linear map commuting with the full $U(2)$ action also commutes with the central element iI_2, making it complex-linear. Hence, the complex Schur statement also characterizes the real commutant after realification.

The implementation uses the realification of this same representation. For $z = (z_1, z_2) \in \mathbb{C}^2$, we use the coordinate ordering

$$q(z) = (\Re z_1, \Im z_1, \Re z_2, \Im z_2)^\top \in \mathbb{R}^4.$$

For a complex number $u = \alpha + i\beta$, let

$$R(u) = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

denote the real matrix representing multiplication by u. Then, the corresponding real representation is given by $\rho_{\mathbb{R}}$ defined by

$$q(Uz) = \rho_{\mathbb{R}}(U) q(z), \qquad \rho_{\mathbb{R}}(U) = \begin{pmatrix} R(U_{11}) & R(U_{12}) \\ R(U_{21}) & R(U_{22}) \end{pmatrix}.$$

To characterize the equivariant linear maps in these coordinates, first note that on a single copy of $\mathbb{C}^2$, any complex linear map commuting with the standard $\mathrm{U}(2)$ action is a complex scalar multiple of I_2. Under our realification procedure, multiplication by i is represented by

$$J_{\mathbb{C}} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

Multiplication by a complex scalar $\lambda = c + id$ is represented in these coordinates by $cI_4 + dJ_{\mathbb{C}}$. Thus, the complex intertwiner λI_2 corresponds exactly to the real intertwiner $cI_4 + dJ_{\mathbb{C}}$. For several

copies of $\mathbb{C}^2$, an equivariant linear map combines the copies using complex scalar coefficients. Each such coefficient acts on one copy as λI_2, and under realification $\lambda = c + id$ becomes the 4×4 block $cI_4 + dJ_{\mathbb{C}}$. Thus, the real-valued representation of $A \otimes I_2$ is a block matrix obtained by replacing each entry of A with its corresponding block $cI_4 + dJ_{\mathbb{C}}$.

The reflection group D_2: In this paper,

$$D_2 = \langle r_x, r_y : r_x^2 = r_y^2 = e,\ r_x r_y = r_y r_x \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$$

is the four-element group generated by the two coordinate reflections. It has four real one-dimensional irreducible representation spaces,

$$\chi_{p_x,p_y}, \qquad (p_x, p_y) \in \{\pm 1\}^2,$$

where r_x acts by p_x and r_y acts by p_y. Products of parity types are component-wise products of the signs. Since $D_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, an equivariant linear map is similar to the $\mathbb{Z}_2$ case. In particular, if the input and output representations contain $k^{\text{in}}_{p_x,p_y}$ and $k^{\text{out}}_{p_x,p_y}$ copies of χ_{p_x,p_y}, respectively, then each D_2-equivariant linear map may be written in the form $\operatorname{diag}(A_{++}, A_{+-}, A_{-+}, A_{--})$, where $A_{p_x,p_y} \in \mathbb{R}^{k^{\text{out}}_{p_x,p_y} \times k^{\text{in}}_{p_x,p_y}}$.

Hence, copies with the same reflection parities may mix arbitrarily, though different parity types may not mix linearly.

A.5 Neural Operators, DeepONets, and FNOs

A neural operator is a parametric map between spaces of functions, intended to approximate an operator rather than a fixed finite-dimensional map (Kovachki et al., 2023). The DeepONet form in (1.1) is motivated by operator universal approximation (Chen & Chen, 1995) and developed in its modern branch–trunk form by Lu et al. (2021). Approximation and generalization in infinite-dimensional settings are studied by Lanthaler et al. (2022).

The reconstruction by an ordinary DeepONet may require high rank on transformation-dominated solution families. Venturi & Casey (2023) relate this phenomenon to singular-value structure and introduce a learned moving reference frame, while Lanthaler et al. (2023) give lower bounds for linear reconstruction on certain discontinuous families and develop nonlinear alternatives.

An FNO involves a channel lifting, a sequence of applications of integrals with learned convolution kernels in spectral space, and a channel projection. The Fourier layer has the form

$$v_{k+1}(x) = \sigma\big(W_k v_k(x) + \mathcal{F}^{-1}\big(R_k(\xi)\mathcal{F}v_k(\xi)\big)(x)\big)\,,$$

where R_k is a learned spectral multiplier (Li et al., 2021). Such networks are akin to a learned pseudospectral method and do not suffer from the explicit low-rank bottleneck. They are also efficient on regular grids, though they do require more complex machinery otherwise. Spherical FNOs replace the Euclidean transform by a spherical one (Bonev et al., 2023); group-equivariant FNOs additionally tie the multipliers under other prescribed transformations (Helwig et al., 2023).

A.6 Equivariant Multilayer Perceptrons

An equivariant MLP is a finite-dimensional network whose hidden features live in representation spaces and whose layers commute with the prescribed group actions. If a hidden feature space has an isotypic decomposition as in (A.1), an analytic Schur layer applies trainable maps to multiplicity coordinates and leaves each irreducible representation coordinate unchanged, as in (A.2). Biases may occur only in copies of the trivial representation.

A second construction solves the intertwining equations directly. For a connected matrix Lie group G with Lie algebra $\mathfrak{g}$ and basis $A_1, \dots, A_s$, a linear map $W : V_{\text{in}} \to V_{\text{out}}$ is an intertwiner when

$$d\rho_{\text{out}}(A_k)W - W d\rho_{\text{in}}(A_k) = 0, \qquad k = 1, \dots, s.$$

After choosing bases and vectorizing W, these are homogeneous linear equations whose nullspace parameterizes the equivariant linear maps. If the group has disconnected or finite components, one additionally imposes $\rho_{\text{out}}(h)W - W\rho_{\text{in}}(h) = 0$ for generators h of those components. This is the generator-nullspace construction of Finzi et al. (2021); we refer to this work for further details.

For completeness, we record the generator equations for the continuous groups that arise in the subsequent experiments. For $\mathrm{SO}(3)$, take the standard basis of $\mathfrak{so}(3)$,

$$A_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad A_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad A_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On the degree-ℓ irreducible representation V_ℓ, write

$$J_i^{(\ell)} = d\rho_\ell(A_i), \qquad i \in \{x, y, z\}.$$

If the input and output contain k_ℓ^{in} and k_ℓ^{out} copies of V_ℓ, respectively, then the degree-ℓ block W_ℓ must satisfy

$$\left(I_{k_\ell^{\mathrm{out}}} \otimes J_i^{(\ell)}\right) W_\ell - W_\ell \left(I_{k_\ell^{\mathrm{in}}} \otimes J_i^{(\ell)}\right) = 0, \qquad i \in \{x, y, z\}.$$

These equations recover the Schur form $W_\ell = A_\ell \otimes I_{V_\ell}$ described above.

For $\mathrm{O}(2)$, the connected component is $\mathrm{SO}(2)$, whose Lie algebra has the single generator

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

For the frequency-m representation in (A.4),

$$d\rho_{m,p}(A) = m \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Hence, a linear map W_m between frequency-m blocks must satisfy the constraint

$$\left(I_{k_{\mathrm{out}}} \otimes d\rho_{m,p_{\mathrm{out}}}(A)\right) W_m - W_m \left(I_{k_{\mathrm{in}}} \otimes d\rho_{m,p_{\mathrm{in}}}(A)\right) = 0.$$

Since $\mathrm{O}(2)$ is disconnected, this infinitesimal condition must be supplemented by the reflection constraint

$$\left(I_{k_{\mathrm{out}}} \otimes \rho_{m,p_{\mathrm{out}}}(\kappa)\right) W_m - W_m \left(I_{k_{\mathrm{in}}} \otimes \rho_{m,p_{\mathrm{in}}}(\kappa)\right) = 0.$$

Together, the infinitesimal and reflection constraints give the $\mathrm{O}(2)$ intertwiners described in Appendix A.4.

For $\mathrm{U}(2)$, a convenient real basis of the Lie algebra $\mathfrak{u}(2)$ is

$$A_0 = iI_2, \qquad A_1 = i\sigma_1, \qquad A_2 = i\sigma_2, \qquad A_3 = i\sigma_3,$$

where the matrices σ_j are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

For the standard representation on $\mathbb{C}^2$,

$$d\rho(A_j) = A_j, \qquad j = 0, 1, 2, 3.$$

Thus, if the input and output contain K_{in} and K_{out} copies of the standard doublet, an equivariant linear map W satisfies

$$\left(I_{K_{\mathrm{out}}} \otimes A_j\right) W - W \left(I_{K_{\mathrm{in}}} \otimes A_j\right) = 0, \qquad j = 0, 1, 2, 3.$$

In the realified implementation, the same equations hold after replacing each A_j by its 4×4 realification under the coordinate convention described in Appendix A.4.

The groups $\mathbb{Z}_2$ and D_2 have no nontrivial Lie algebra generators, as they are finite. Instead, their equivariance constraints are imposed directly through the finite-group intertwining equations. For $\mathbb{Z}_2$, it is sufficient to impose the constraint for the non-identity element. For $D_2 = \langle r_x, r_y \rangle$, it is enough to impose that

$$\rho_{\mathrm{out}}(r_x) W - W \rho_{\mathrm{in}}(r_x) = 0, \qquad \rho_{\mathrm{out}}(r_y) W - W \rho_{\mathrm{in}}(r_y) = 0.$$

For product groups containing both continuous and finite factors, the corresponding infinitesimal and finite constraints are imposed simultaneously.

Linear intertwiners produce a *linear* network. Equivariant nonlinear blocks combine them with the mechanisms in Appendix A.3. A gated block has the schematic form

$$z^{(k+1)} = \Gamma_k\big(L_k z^{(k)}\big),$$

where L_k is an intertwiner and Γ_k is an invariant gating map previously described. A tensor-product block has the form

$$z^{(k+1)} = L_k z^{(k)} + \mathcal{Q}_k\big(L_k z^{(k)}, L_k z^{(k)}\big),$$

where $\mathcal{Q}_k$ is assembled from fixed Clebsch–Gordan or Gaunt intertwiners and learned multiplicity mixing. The Gross–Pitaevskii block replaces the quadratic term by the cubic Gram construction. The two equivariant mechanisms may also be computed via, e.g.,

$$z^{(k+1)} = \Gamma_k\left(L_k z^{(k)} + \mathcal{Q}_k\big(L_k z^{(k)}, L_k z^{(k)}\big)\right).$$

Since each constituent map is equivariant, any finite composition of these blocks remains equivariant.

B Abstract Construction and Proofs

We now prove the assertions stated in Section 2. Throughout, the notation is the same as in the main text: $\mathcal{O}$ is the observation map, B is the branch map, τ is the representation-valued trunk feature map, C is the bilinear fusion, and $\widetilde{C}$ is the associated linear map on the tensor product.

B.1 Equivariance of the Branch–Trunk Factorization

Let X carry the representation ρ_X, let Ω be a G-space, and let W carry the output-fiber representation ρ_W. Let V_{obs}, V_B, and V_T be finite-dimensional representation spaces with actions ρ_{obs}, ρ_B, and ρ_T. Consider

$$\mathcal{O} : X \to V_{\mathrm{obs}}, \qquad B : V_{\mathrm{obs}} \to V_B, \qquad \tau : \Omega \to V_T,$$

and a bilinear map $C : V_B \times V_T \to W$. Define

$$\widehat{\mathcal{F}}(u)(y) = C\left(B(\mathcal{O}(u)), \tau(y)\right). \tag{B.1}$$

Let $\widetilde{C} : V_B \otimes V_T \to W$ be the unique linear map satisfying $\widetilde{C}(b \otimes t) = C(b, t)$. We recall the equivariance conditions placed in the assumptions of Theorem 2.1, namely that for every $g \in G$,

$$\mathcal{O} \circ \rho_X(g) = \rho_{\mathrm{obs}}(g) \circ \mathcal{O}, \tag{B.2}$$
$$B \circ \rho_{\mathrm{obs}}(g) = \rho_B(g) \circ B, \tag{B.3}$$
$$\tau(g \cdot y) = \rho_T(g)\tau(y), \tag{B.4}$$
$$C\left(\rho_B(g)b, \rho_T(g)t\right) = \rho_W(g)C(b, t). \tag{B.5}$$

(Proof of Theorem 2.1). Fix $g \in G$, $u \in X$, and $y \in \Omega$. Applying the four hypotheses (B.2)-(B.5) in order gives

$$\begin{aligned}
\widehat{\mathcal{F}}(\rho_X(g)u)(g \cdot y) &= C\left(B(\mathcal{O}(\rho_X(g)u)), \tau(g \cdot y)\right) \\
&= C\left(B(\rho_{\mathrm{obs}}(g)\mathcal{O}(u)), \tau(g \cdot y)\right) \\
&= C\left(\rho_B(g)B(\mathcal{O}(u)), \rho_T(g)\tau(y)\right) \\
&= \rho_W(g)C\left(B(\mathcal{O}(u)), \tau(y)\right) \\
&= \rho_W(g)\widehat{\mathcal{F}}(u)(y).
\end{aligned}$$

Replacing y by $g^{-1} \cdot y$ yields

$$\widehat{\mathcal{F}}(\rho_X(g)u)(y) = \rho_W(g)\widehat{\mathcal{F}}(u)(g^{-1} \cdot y) = [\rho_Y(g)\widehat{\mathcal{F}}(u)](y).$$

Thus, $\widehat{\mathcal{F}} \circ \rho_X(g) = \rho_Y(g) \circ \widehat{\mathcal{F}}$. □

Corollary B.1 (Autoregressive equivariance)**.** *If $\widehat{\mathcal{F}} : X \to X$ is equivariant, then every finite iterate $\widehat{\mathcal{F}}^k$ is equivariant.*

Since the composition of equivariant maps is equivariant, the proof follows trivially by induction.

Proposition B.1 (Equivalence with an operator-valued trunk)**.** *Under* (B.4)–(B.5)*, define* $T : \Omega \to \mathrm{Hom}(V_B, W)$ *by*

$$T(y)b = C(b, \tau(y)).$$

Then,

$$T(g \cdot y) \circ \rho_B(g) = \rho_W(g) \circ T(y). \tag{B.6}$$

Conversely, if T *satisfies* (B.6)*, then the bilinear formulation is recovered by taking* $V_T = \mathrm{Hom}(V_B, W)$*,* $\tau = T$*, and* $C(b, A) = A(b)$*.*

Proof. For $b \in V_B$,

$$\begin{aligned}[T(g \cdot y) \circ \rho_B(g)]b &= C\left(\rho_B(g)b, \tau(g \cdot y)\right) \\ &= C\left(\rho_B(g)b, \rho_T(g)\tau(y)\right) \\ &= \rho_W(g)C(b, \tau(y)) \\ &= [\rho_W(g) \circ T(y)]b.\end{aligned}$$

For the converse, equip $\mathrm{Hom}(V_B, W)$ with the action in (2.3). Rearranging (B.6) gives

$$T(g \cdot y) = \rho_W(g) \circ T(y) \circ \rho_B(g)^{-1} = \rho_{\mathrm{Hom}}(g)T(y).$$

Thus, taking $V_T = \mathrm{Hom}(V_B, W)$ and $\tau = T$, the trunk satisfies

$$\tau(g \cdot y) = \rho_{\mathrm{Hom}}(g)\tau(y).$$

Now, define the bilinear evaluation map

$$C(b, A) = A(b), \qquad A \in \mathrm{Hom}(V_B, W).$$

Then,

$$C\left(\rho_B(g)b, \rho_{\mathrm{Hom}}(g)A\right) = [\rho_{\mathrm{Hom}}(g)A](\rho_B(g)b) = \rho_W(g)A(b) = \rho_W(g)C(b, A),$$

so the fusion condition is also satisfied. □

B.2 Unisolvent Finite Sensors

Definition B.1 (Finite sensor family)**.** Let E be a finite-dimensional space of functions with well-defined pointwise evaluations on a domain M. A finite sensor family $\Xi = (x_1, \ldots, x_m)$ defines the sampling map

$$\mathcal{S}_\Xi : E \to Z, \qquad \mathcal{S}_\Xi u = (u(x_1), \ldots, u(x_m)),$$

where Z is the finite-dimensional space of sampled values. The family is *unisolvent* for E if $\mathcal{S}_\Xi$ is injective.

The relevant theorem is Theorem B.1, which is the formal analogue of Theorem 2.2 (1).

Theorem B.1 (Equivariant DeepONets with Finite Observations)**.** *Let* G *be compact and suppose that the equivariance properties in Theorem 2.1 hold. Further, let* $E \subset X$ *be a finite-dimensional* G*-invariant subspace,* $Q : E \to V_{\mathrm{obs}}$ *be an invertible intertwiner,* Z *be a finite-dimensional inner-product space, and* $\mathcal{S}_\Xi : E \to Z$ *be a sampling map with sensor locations* Ξ *in the input domain and point evaluation well-defined on* E*. If*

$$A := \mathcal{S}_\Xi Q^{-1} : V_{\mathrm{obs}} \to Z$$

is injective, then $\mathcal{O}_{\mathrm{sens}} := A^+ \mathcal{S}_\Xi$ *satisfies* $\mathcal{O}_{\mathrm{sens}} = Q$ *on* E*, where* A^+ *is the Moore–Penrose pseudoinverse of* A*. Consequently, the finite-sensor and coefficient-observation DeepONets coincide on* E *and are exactly* G*-equivariant there.*

Before proving the theorem, let us provide a computational demonstration for intuition. Let E be spanned by basis vectors $\{\phi_j\}_{j=1}^n$. In this case, we may write any $u \in E$ as $u = \sum_{j=1}^{n} u_j \phi_j$. A natural choice of Q is given by

$$Q(u) = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Due to the invariance,

$$\rho_X(g)\phi_k = \sum_{j=1}^{n} [\rho_E(g)]_{jk}\phi_j,$$

and so $Q(\rho_X(g)u) = \rho_E(g)Q(u)$, showing that Q is an intertwiner. Given sensors $\Xi = \{x_1, \cdots, x_m\}$, we may write

$$\mathcal{S}_\Xi u = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_m) \end{pmatrix} = \underbrace{\begin{pmatrix} \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & & \vdots \\ \phi_1(x_m) & \cdots & \phi_n(x_m) \end{pmatrix}}_{A} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

That is, $\mathcal{S}_\Xi = AQ$. If A has full column rank, then we may write $Q = A^+\mathcal{S}_\Xi$. If we are considering $SO(3)$ functions with bandlimit L, then

$$E = \bigoplus_{\ell=0}^{L} V_\ell$$

is a valid G-invariant subspace. $Q(u)$ is the vector of the retained spherical harmonic coefficients, and ρ_E is block-diagonal with Wigner D-matrices along the diagonal.

Now, we formally prove Theorem B.1.

(Proof of Theorem B.1). Since A is injective, its Moore–Penrose pseudoinverse $A^+ := (A^*A)^{-1}A^*$ is a left inverse of A on V_{obs}. For $u \in E$,

$$\begin{aligned} \mathcal{O}_{\text{sens}}(u) &= A^+(\mathcal{S}_\Xi(u)) \\ &= A^+\left(\mathcal{S}_\Xi \circ Q^{-1})(Q(u))\right) \\ &= (A^+ \circ A)(Q(u)) \\ &= Q(u). \end{aligned}$$

Since E is G-invariant, $Q \circ \rho_X(g) = \rho_{\text{obs}}(g) \circ Q$. Based on the above, the same identity holds for $\mathcal{O}_{\text{sens}}$. The final assertion follows by substituting $\mathcal{O}_{\text{sens}} = A^+\mathcal{S}_\Xi$ into (B.1). □

Corollary B.2 (Sensor stability)**.** *Under the assumptions of Theorem B.1, suppose the measured sensor vector is $\mathcal{S}_\Xi(u) + \eta$. Then*

$$\left\|A^+(\mathcal{S}_\Xi(u) + \eta) - Q(u)\right\| \le \frac{\|\eta\|}{\sigma_{\min}(A)}.$$

If $u = v + r$ with $v \in E$ and $\mathcal{S}_\Xi(r)$ well-defined, then

$$\left\|A^+\mathcal{S}_\Xi(u) - Q(v)\right\| \le \frac{\|\mathcal{S}_\Xi(r)\|}{\sigma_{\min}(A)}.$$

Proof. Using the expression for $\mathcal{O}_{\text{sens}}$,

$$A^+(\mathcal{S}_\Xi(u) + \eta) - Q(u) = A^+\eta.$$

The first estimate follows from the standard identity $\|A^+\|_2 = \sigma_{\min}(A)^{-1}$ and the operator norm inequality. For the second, we note that

$$A^+\mathcal{S}_\Xi(u) - Q(v) = A^+\mathcal{S}_\Xi(r),$$

then proceed as in the prior estimate. □

B.3 Group-Closed Sensor Families

Exact reconstruction is not the only way to obtain an equivariant finite observation. We now formalize Theorem 2.2 (2) precisely. Below, let $W^m = \prod_{i=1}^{m} W$.

Theorem B.2 (Group-closed sensors)**.** *Let a group H act on Ω, W carry a representation ρ_W, and $\Xi = \{x_1, \dots, x_m\} \subset \Omega$ satisfy $h\Xi = \Xi$ for every $h \in H$. For each h, let σ_h be the permutation determined by $hx_i = x_{\sigma_h(i)}$. Define $\pi_\Xi(h) : W^m \to W^m$ by*

$$[\pi_\Xi(h)z]_i = \rho_W(h) z_{\sigma_{h^{-1}}(i)}.$$

Then, the sampling map $\mathcal{S}_\Xi u = (u(x_i))_{i=1}^m$ satisfies

$$\mathcal{S}_\Xi \circ \rho(h) = \pi_\Xi(h) \circ \mathcal{S}_\Xi.$$

The same conclusion holds for a commuting internal group acting fiberwise on W.

(Proof of Theorem B.2). For every i,

$$\begin{aligned}
[\mathcal{S}_\Xi(\rho(h)f)]_i &= [\rho(h)f](x_i) \\
&= \rho_W(h) f(h^{-1}x_i) \\
&= \rho_W(h) f(x_{\sigma_{h^{-1}}(i)}) \\
&= [\pi_\Xi(h)\mathcal{S}_\Xi(f)]_i.
\end{aligned}$$

Thus, the following diagram commutes:

$$\begin{array}{ccc}
X & \xrightarrow{\rho(h)} & X \\
{\scriptstyle \mathcal{S}_\Xi}\downarrow & & \downarrow{\scriptstyle \mathcal{S}_\Xi} \\
W^m & \xrightarrow{\pi_\Xi(h)} & W^m.
\end{array}$$

Equivalently,

$$\mathcal{S}_\Xi \circ \rho(h) = \pi_\Xi(h) \circ \mathcal{S}_\Xi,$$

so $\mathcal{S}_\Xi$ is an intertwiner between the field and sensor representations.

A commuting internal action is applied identically at every sensor and therefore commutes with the spatial permutation.

□

This does not require the samples to determine the complete field. It is the relevant observation principle for the Gross–Pitaevskii sensor family. By contrast, a generic finite subset cannot be closed under a continuous transitive group, so Theorem B.1 supplies the useful finite-dimensional alternative in the spherical and annular examples.

B.4 Finite-Rank Equivariant Approximation

We now move on from input observation and turn to output reconstruction. We first prove the finite-rank approximation result used in the main text. We then progressively strengthen the theorem at the cost of additional abstraction. The content is summarized in Table B.1.

Theorem B.3 (Finite-rank equivariant approximation)**.** *Let G be a compact Hausdorff group, let X and Y be Hilbert spaces carrying continuous unitary representations, and let $K \subset X$ be compact and G-invariant. Suppose $\mathcal{F} : K \to Y$ is continuous and equivariant. Assume that Y has a dense, increasing family of finite-dimensional G-invariant subspaces $\{Y_n \subset Y\}$. Then, for every $\varepsilon > 0$, there exist a finite-dimensional representation space V, an isometric intertwiner $\iota : V \to Y$, and a continuous equivariant map $b : K \to V$ such that*

$$\sup_{u \in K} \|\mathcal{F}(u) - \iota(b(u))\|_Y < \varepsilon. \tag{B.7}$$

If the range of ι admits a continuous basis of functions on Ω, then $\iota \circ b$ has an equivariant finite branch–trunk expansion.

Result	Finite output space	Finite input observation	Branch guarantee
Theorem B.3	Prescribed through $\{Y_n\}$	Not constructed	Continuous equivariant finite-rank map
Finite-observation remark	As above	Assumed injective on K	Continuous equivariant branch on $\mathcal{O}(K)$
Theorem B.4	Constructed from compactness and Peter–Weyl	Not constructed	Continuous equivariant finite-rank map
Theorem B.5	Chosen from $\{Y_n\}$	Constructed as P_E	Equivariant polynomial; finite tensor-product realization

Table B.1: Summary of the approximation results. The first two results isolate the output and finite-observation requirements, while the final result also constructs a branch in an explicit equivariant function class.

The idea is to project the equivariant operator output onto a sufficiently large G-invariant subspace.

(Proof of Theorem B.3). Let $P_n : Y \to Y_n$ be the orthogonal projection. Since Y_n is invariant and $\rho_Y(g)$ is unitary, $Y_n^\perp$ is invariant, as well. Hence,

$$P_n \circ \rho_Y(g) = \rho_Y(g) \circ P_n.$$

The density assumption implies that $P_n y \to y$ for every $y \in Y$. By continuity and compactness of K, $\mathcal{F}(K)$ is compact. Taking this with $\|I - P_n\| \le 1$, the convergence is uniform on $\mathcal{F}(K)$. In particular,

$$\sup_{u \in K} \|(I - P_n)\mathcal{F}(u)\|_Y \longrightarrow 0.$$

Choose n large enough so that the above supremum is less than ε. Set $V = Y_n$ with the restricted representation, let $\iota : V \hookrightarrow Y$ be the inclusion map, and define

$$b = P_n \circ \mathcal{F}.$$

Then, b is continuous and equivariant, ι is an isometric intertwiner, and (B.7) follows.

If $\{\phi_1, \dots, \phi_p\}$ is a continuous basis of Y_n, write

$$\iota(b(u))(y) = \sum_{j=1}^{p} b_j(u)\phi_j(y).$$

With

$$\beta(u) = (b_1(u), \dots, b_p(u)), \qquad \tau(y) = (\phi_1(y), \dots, \phi_p(y)),$$

and

$$C(c, t) = \sum_{j=1}^{p} c_j t_j,$$

we obtain the finite branch–trunk expansion

$$\iota(b(u))(y) = C(\beta(u), \tau(y)).$$

Since Y_n is G-invariant and b is equivariant, the induced actions on the coefficient and trunk spaces make β, τ, and C satisfy the corresponding equivariance conditions. □

Remark B.1. In Theorem B.3, the map $b = P_n\mathcal{F}$ is finite-dimensional and equivariant, but a DeepONet also requires the branch to depend on a *finite-dimensional* observation of the input. This follows whenever the observation is sufficient to distinguish the inputs in K. In particular, injectivity of an equivariant observation map restricted to K is sufficient.

More precisely, suppose that there is a finite-dimensional representation space V_{obs} and a continuous equivariant observation $\mathcal{O} : X \to V_{\text{obs}}$ such that $\mathcal{O}|_K$ is injective. Since K is compact and V_{obs} is Hausdorff, the map $\mathcal{O}|_K : K \to \mathcal{O}(K)$ is a homeomorphism. Therefore, for the n chosen as in the prior theorem, we may define a branch on the set of observations that actually occur on K by

$$B_K(z) = P_n\mathcal{F}\left((\mathcal{O}|_K)^{-1}(z)\right), \qquad z \in \mathcal{O}(K).$$

This map is continuous and equivariant. Indeed, if $z = \mathcal{O}(u)$, then equivariance of $\mathcal{O}$, $\mathcal{F}$, and P_n gives

$$\begin{aligned} B_K(\rho_{\text{obs}}(g)z) &= B_K(\mathcal{O}(\rho_X(g)u)) \\ &= P_n\mathcal{F}(\rho_X(g)u) \\ &= \rho_{Y_n}(g)P_n\mathcal{F}(u) \\ &= \rho_{Y_n}(g)B_K(z). \end{aligned}$$

Consequently, $B_K(\mathcal{O}(u)) = P_n\mathcal{F}(u), \qquad u \in K$, so the approximation in Theorem B.3 factors through the finite observation $\mathcal{O}$ as needed for a fully-valid DeepONet construction.

For approximation on K, this is all that is needed. If one additionally requires the branch to be defined on all of V_{obs}, then B_K can be extended equivariantly. Since $\mathcal{O}(K)$ is compact in V_{obs}, there is a continuous extension $\overline{B} : V_{\text{obs}} \to Y_n$ coordinate-wise. Let μ denote the normalized Haar measure on G, and define

$$B(z) = \int_G \rho_{Y_n}(g)^{-1}\overline{B}(\rho_{\text{obs}}(g)z)\, d\mu(g).$$

A change of variables shows that B is equivariant. Since $\mathcal{O}(K)$ is G-invariant and B_K is already equivariant there, the averaging leaves B_K unchanged on $\mathcal{O}(K)$. Thus, $B(\mathcal{O}(u)) = P_n\mathcal{F}(u)$ for every $u \in K$.

In particular, if $K \subset E$ under the assumptions of Theorem B.1, then Q is injective on K and $A^+\mathcal{S}_\Xi = Q$ on E. Hence, the same finite-observation factorization applies to the unisolvent sensor construction.

The assumption of an increasing family $\{Y_n\}$ is natural in the problems considered here since it is provided by harmonic truncation. It is not, however, strictly *required* for finite-rank approximation itself.

Theorem B.4 (Finite-rank approximation without a prescribed exhaustion)**.** *Let G be a compact Hausdorff group, let X and Y be Hilbert spaces carrying continuous unitary representations, and let $K \subset X$ be compact and G-invariant. Suppose $\mathcal{F} : K \to Y$ is continuous and equivariant. Then, for every $\varepsilon > 0$, there exists a finite-dimensional G-invariant subspace $V \subset Y$ such that, if $P_V : Y \to V$ denotes the orthogonal projection,*

$$\sup_{u \in K} \|\mathcal{F}(u) - P_V\mathcal{F}(u)\|_Y < \varepsilon.$$

In particular, $P_V\mathcal{F} : K \to V$ is continuous and equivariant.

The proof is quite similar to Theorem B.3, but we now must construct the space Y_n rather than being provided it by assumption.

(Proof of Theorem B.4). A vector $v \in Y$ is called G-finite if its orbit spans a finite-dimensional subspace. For a continuous unitary representation of a compact group, the G-finite vectors are dense in Y by Peter–Weyl theory; see, for example, (Folland, 1995, Chapter 5). Since $\mathcal{F}(K)$ is compact, choose a finite $\varepsilon/3$-net $\{f_1, \dots, f_N\} \subset \mathcal{F}(K)$. For each i, choose a G-finite vector $v_i \in Y$ satisfying

$$\|f_i - v_i\|_Y < \varepsilon/3.$$

Let

$$V = \sum_{i=1}^{N} \text{span}\{\rho_Y(g)v_i : g \in G\}.$$

Each orbit span is finite-dimensional, so V is finite-dimensional and G-invariant. Let $P_V : Y \to V$ be the orthogonal projection. Since the representation is unitary and V is invariant, P_V is an intertwiner. For $f \in \mathcal{F}(K)$, choose f_i with $\|f - f_i\|_Y < \varepsilon/3$. Since $v_i \in V$ and the orthogonal projection gives the best approximation in V,

$$\|f - P_V f\|_Y \le \|f - v_i\|_Y \le \|f - f_i\|_Y + \|f_i - v_i\|_Y < \frac{2\varepsilon}{3}.$$

Taking $f = \mathcal{F}(u)$ and taking the supremum proves the claim. □

If the subspace V in Theorem B.4 admits a continuous basis of functions on Ω, then the same basis construction used in Theorem B.3 gives an equivariant finite branch–trunk expansion. Likewise, if a finite observation $\mathcal{O}$ is injective on K, then $P_V\mathcal{F}$ factors through $\mathcal{O}(K)$ by the same argument as above. The abstract theorem does not guarantee that V consists of continuous functions. In the spherical and annular settings considered here, the explicit harmonic and Fourier subspaces already provide such spaces, so this distinction does not arise.

Finally, we consider the stronger statement in which both the finite observation and the finite-dimensional branch are constructed as part of the approximation. The output reduction is already provided by Theorem B.3. On the input side, the same compactness and Peter–Weyl argument used in Theorem B.4 produces a finite-dimensional G-invariant subspace that uniformly approximates the compact input set. The additional step here is to show that the resulting finite-dimensional branch may be chosen to be an equivariant polynomial. This, in turn, places the branch in a function class generated by finite tensor products and equivariant linear maps. The polynomial symmetrization argument is standard in compact-group equivariant approximation; see, for example, Yarotsky (2022).

Theorem B.5 (Finite-observation approximation with tensor-product branches)**.** *Let G be a compact Hausdorff group, let X and Y be real Hilbert spaces carrying continuous orthogonal representations, and let $K \subset X$ be compact and G-invariant. Suppose $\mathcal{F} : K \to Y$ is continuous and equivariant. Assume that Y has a dense, increasing family of finite-dimensional G-invariant subspaces $\{Y_n \subset Y\}$ and that each Y_n admits a continuous basis of functions on Ω.*

Then, for every $\varepsilon > 0$, there exist an index n, a finite-dimensional G-invariant subspace $E \subset X$, and an equivariant polynomial map

$$B : E \to Y_n$$

such that

$$\sup_{u \in K} \|\mathcal{F}(u) - B(P_E u)\|_Y < \varepsilon,$$

where $P_E : X \to E$ is the orthogonal projection. Thus, $\mathcal{O} = P_E$ gives a finite-dimensional equivariant observation, and a continuous basis of Y_n gives the corresponding equivariant branch–trunk representation.

Moreover, if

$$B = \sum_{r=0}^{R} B_r$$

is the decomposition of B into homogeneous polynomial components, then, for each $r \geq 1$, there is a G-intertwiner

$$L_r : \mathrm{Sym}^r(E) \to Y_n$$

such that

$$B_r(z) = L_r(z^{\otimes r}).$$

The last part of the theorem states that a branch function class which allows for arbitrary finite tensor powers and equivariant linear maps contains the polynomial branch B exactly.

(Proof of Theorem B.5). We separate the construction into the output reduction, input reduction, and finite-dimensional branch approximation.

First, by Theorem B.3, choose n sufficiently large that

$$\sup_{u \in K} \|\mathcal{F}(u) - P_n\mathcal{F}(u)\|_Y < \frac{\varepsilon}{3}.$$

Define

$$b = P_n\mathcal{F} : K \to Y_n.$$

Then, b is continuous and equivariant.

To compare $b(u)$ with a function of a finite-dimensional projection of u, we first extend b off of K. In a manner nearly identical to the remark following Theorem B.3, one can continuously extend to $\widetilde{b} : X \to Y_n$. Transforming by $\rho_{Y_n}^{-1}$ and averaging over G via the Haar measure on G converts this to

an equivariant map $\bar{b}$ which matches $\widetilde{b}$ on K. By continuity of $\bar{b}$ and compactness of K, there exists $\delta > 0$ such that

$$\|\bar{b}(u) - \bar{b}(z)\|_Y < \frac{\varepsilon}{3}$$

whenever $u \in K$ and $\|u - z\|_X < \delta$.

We now apply the same compactness and Peter–Weyl construction used in the proof of Theorem B.4, this time to the compact set $K \subset X$. Since the G-finite vectors are dense in X, there is a finite-dimensional G-invariant subspace $E \subset X$ such that

$$\sup_{u \in K} \|u - P_E u\|_X < \delta.$$

Consequently,

$$\sup_{u \in K} \|\bar{b}(u) - \bar{b}(P_E u)\|_Y < \frac{\varepsilon}{3}.$$

Since E is G-invariant and the representation is orthogonal, P_E is an intertwiner. Hence, $P_E K$ is compact and G-invariant.

It remains only to approximate the finite-dimensional equivariant map $\bar{b}|_E : E \to Y_n$ on the compact set $P_E K$. Since both E and Y_n are finite-dimensional, polynomial maps are dense on $P_E K$ by Stone–Weierstrass applied coordinate-wise. Thus, there is a polynomial map $p : E \to Y_n$ such that

$$\sup_{z \in P_E K} \|\bar{b}(z) - p(z)\|_Y < \frac{\varepsilon}{3}.$$

The polynomial p need not be equivariant. We enforce equivariance by averaging over G:

$$B(z) = \int_G \rho_{Y_n}(g)^{-1} p(\rho_E(g) z)\, d\mu(g).$$

Since the action on E is linear, B is again a polynomial. A change of variables in the Haar integral shows that B is equivariant. Moreover, since $P_E K$ is G-invariant and $\bar{b}$ is equivariant once again:

$$\begin{aligned} \|\bar{b}(z) - B(z)\|_Y &\le \int_G \|\bar{b}(z) - \rho_{Y_n}(g)^{-1} p(\rho_E(g) z)\|_Y \, d\mu(g) \\ &= \int_G \|\bar{b}(\rho_E(g) z) - p(\rho_E(g) z)\|_Y \, d\mu(g) \\ &< \frac{\varepsilon}{3} \end{aligned}$$

for every $z \in P_E K$.

Now, we may combine our approximations: For every $u \in K$,

$$\begin{aligned} \|\mathcal{F}(u) - B(P_E u)\|_Y \le \|\mathcal{F}(u) - b(u)\|_Y \\ + \|\bar{b}(u) - \bar{b}(P_E u)\|_Y \\ + \|\bar{b}(P_E u) - B(P_E u)\|_Y \\ < \varepsilon. \end{aligned}$$

Finally, write

$$B = \sum_{r=0}^{R} B_r$$

as a sum of homogeneous polynomials. Since the group actions are linear and B is equivariant, each B_r is equivariant. For $r \ge 1$, polarization gives a symmetric r-linear map, equivalently a linear map

$$L_r : \mathrm{Sym}^r(E) \to Y_n,$$

such that

$$B_r(z) = L_r(z^{\otimes r}).$$

Equivariance of B_r implies that L_r is an intertwiner. The degree-zero term lies in the invariant subspace Y_n^G. Thus, the polynomial branch is realized by finite tensor powers, equivariant linear maps, and invariant constants. □

C BENCHMARKS: PDE FORMULATIONS, MODEL REALIZATIONS, AND DATASET CONSTRUCTION

All four datasets are generated from numerical solvers for specific PDE systems. Each trajectory contains six stored states, including the initial condition (so, five adjacent one-step pairs). The train/validation/test split is by trajectory, with 4096/256/512 trajectories in every problem. Normalization is fit on the training split and respects the physical representation types: components mixed by a symmetry share scales, and symmetry-related vector or doublet components are not assigned incompatible shifts.

For all proofs involving equivariance, we will assume that the problem is sufficiently regular to generate a well-defined flow map (as is needed for the operator learning problem to be well-posed).

C.1 ALLEN–CAHN EQUATION ON $\mathbb{S}^2$

Equations and Symmetry Groups. Let $\mathbb{S}^2 \subset \mathbb{R}^3$ be the unit sphere. We consider the *Allen–Cahn* equation

$$\partial_t u = \varepsilon^2 \Delta_{\mathbb{S}^2} u + u - u^3, \qquad \varepsilon = 0.08. \tag{C.1}$$

The symmetry group is

$$G_{\mathrm{AC}} = \mathrm{SO}(3) \times \mathbb{Z}_2.$$

For $(R, s) \in G_{\mathrm{AC}}$, with $s \in \{\pm 1\}$, define

$$[\rho_{\mathrm{AC}}(R, s)u](x) = s\, u(R^{-1}x). \tag{C.2}$$

Proposition C.1 (Allen–Cahn equivariance)**.** *Let Φ_t^{AC} denote the solution operator for (C.1). Then,*

$$\Phi_t^{\mathrm{AC}} \circ \rho_{\mathrm{AC}}(R, s) = \rho_{\mathrm{AC}}(R, s) \circ \Phi_t^{\mathrm{AC}}$$

for every $(R, s) \in \mathrm{SO}(3) \times \mathbb{Z}_2$ for which the solution is defined.

Proof. We only focus on the notable terms, as all others are clear. Rotations of $\mathbb{S}^2$ are isometries, so $\Delta_{\mathbb{S}^2}(u \circ R^{-1}) = (\Delta_{\mathbb{S}^2} u) \circ R^{-1}$. The reaction term $r(u) = u - u^3$ is odd, so $r(su) = sr(u)$ for $s \in \{\pm 1\}$. Therefore, if u solves (C.1), then $\rho_{\mathrm{AC}}(R, s)u$ solves the same equation with transformed initial data. Uniqueness of the initial-value problem gives the stated identity for the flow map. □

Instantiated Equivariant DeepONet. Let V_ℓ denote the degree-ℓ irreducible representation space of $\mathrm{SO}(3)$, and let χ_+ and χ_- denote the trivial and sign representation spaces of $\mathbb{Z}_2$. Define

$$E_{\mathrm{AC},L}^{\pm} = \bigoplus_{\ell=0}^{L} V_\ell \boxtimes \chi_\pm.$$

Given a function $u : \mathbb{S}^2 \to \mathbb{R}$ with bandlimit L, the external $\mathbb{Z}_2$ action in (C.2) induces the sign representation on its spherical-harmonic coefficients, so the retained coefficients transform in $E_{\mathrm{AC},L}^-$. Let

$$\mathcal{H}_{\mathrm{AC},L} : L^2(\mathbb{S}^2) \to E_{\mathrm{AC},L}^-$$

denote the real spherical harmonic transform with its coefficient space tensored with the sign representation, and let

$$\mathcal{S}_{\mathrm{AC},L} : E_{\mathrm{AC},L}^- \to L^2(\mathbb{S}^2)$$

denote the inverse spherical harmonic transform with the sign-representation factor contracted out. A full-observation coefficient-space model may then be written as

$$\widehat{\Phi}_{\mathrm{AC}} = \mathcal{S}_{\mathrm{AC},L} \circ B_\theta \circ \mathcal{H}_{\mathrm{AC},L}, \tag{C.3}$$

where $B_\theta : E_{\mathrm{AC},L}^- \to E_{\mathrm{AC},L}^-$ is equivariant. In this framework, the trunk is given by the spherical harmonics up to degree L, and the intertwining map is the typical $\mathrm{SO}(3)$-invariant contraction between matching irreducible representations.

More generally, and in the form used by the DeepONet implementation, the branch and trunk take values in K copies of the retained representation. In particular, let

$$b = B_\theta(a) \in \mathbb{R}^K \otimes E_{\mathrm{AC},L}^-, \qquad \tau(x) \in \mathbb{R}^K \otimes E_{\mathrm{AC},L}^+,$$

where K is the latent multiplicity. The trunk is constructed equivariantly from the spherical-harmonic features and is even under the external $\mathbb{Z}_2$ action. The DeepONet output is then

$$\widehat{\Phi}_{\mathrm{AC}}(u)(x) = \sum_{k=1}^{K} C\big(b_k(u), \tau_k(x)\big) = \sum_{k=1}^{K} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} b_{k,\ell m}(u)\, \tau_{k,\ell m}(x), \tag{C.4}$$

where

$$C : E^{-}_{\mathrm{AC},L} \times E^{+}_{\mathrm{AC},L} \to \chi_{-}$$

is the bilinear intertwiner obtained by contracting matching $\mathrm{SO}(3)$ types. Degree-wise, this contraction is

$$C_\ell : (V_\ell \boxtimes \chi_{-}) \times (V_\ell \boxtimes \chi_{+}) \to \chi_{-}, \qquad C_\ell(b_{k,\ell}, \tau_{k,\ell}) = \sum_{m=-\ell}^{\ell} b_{k,\ell m} \tau_{k,\ell m}.$$

The contraction is invariant under the simultaneous $\mathrm{SO}(3)$ action, while $\chi_{-} \otimes \chi_{+} \cong \chi_{-}$ leaves the sign carried by the branch unchanged. Consequently, the output transforms in the same χ_{-} type as the physical field.

This is the one-dimensional, nontrivial-fiber specialization of Section 2. In the selected configuration, the trunk is a learned linear equivariant map of the spherical-harmonic features into K latent copies of $E^{+}_{\mathrm{AC},L}$, while the branch maps the spherical-harmonic coefficients into K latent copies of $E^{-}_{\mathrm{AC},L}$. The branch uses degree-wise Schur intertwiners, invariant norm gates, and $\mathbb{Z}_2$-graded Gaunt blocks. Internally, odd–odd products create even features and even–odd products return odd features; the final branch output has the odd parity required by (C.2). The Gaunt contractions are evaluated pseudospectrally using spherical-harmonic transforms from `torch_harmonics` (Bonev et al., 2023), as described in Appendix A.3.

We remark that, while the above construction is technically dense, it may be described quite simply: the branch is an equivariant network acting on spherical-harmonic coefficients, the trunk is an equivariant linear map of the spherical-harmonic features into the latent representation, and the pairing contracts matching spherical-harmonic types and latent multiplicities to produce the scalar output.

The finite-sensor model observes 640 point values. Sensor locations are selected by a QR decomposition on the real degree-22 spherical-harmonic evaluation matrix. The resulting 640×529 matrix has rank 529=$(22+1)^2$. Its pseudoinverse reconstructs $\mathcal{H}_{22}u$ before the unchanged branch, trunk, and fusion in (C.3); hence the full- and finite-observation models have identical trainable parameter counts.

Comparison Models. The sparse DeepONet uses 1024 fixed sensors, four branch and trunk hidden layers of width 512, and latent dimension 512. The symmetry-aware-trunk model uses spherical-harmonic output features but does not impose equivariance on the observation nor the branch. The SFNO reference has width 64, four spectral layers, and a 23×23 spherical spectral truncation threshold.

Data Generation. The data is generated on a 64×128 equiangular grid using the spherical Allen–Cahn solver in `torch_harmonics` (Bonev et al., 2023) with an imposed maximum degree of $L = 22$. The internal integrator is RK4 with $\Delta t_{\mathrm{solver}} = 0.01$; 20 solver steps form one learned operator time step, so $\Delta t_{\mathrm{operator}} = 0.2$ and the stored times are $0, 0.2, \ldots, 1.0$. Initial conditions are drawn by the spherical solver's random spectral routine over the given spectral range.

C.2 Nonrotating Shallow-Water Equations on $\mathbb{S}^2$

Equations and Symmetry Groups. Let h denote total layer height and v a tangent vector field on a sphere of radius a. In the nonrotating inviscid case,

$$\begin{cases} \partial_t h + \mathrm{div}_{\mathbb{S}^2}(hv) = 0, \\ \partial_t v + \nabla_v v + g\, \mathrm{grad}_{\mathbb{S}^2}\, h = 0, \end{cases} \tag{C.5}$$

where g denotes the gravitational acceleration.

The stored scalar field is $\eta = h - H$, where H is the mean height. The symmetry group is $G_{\mathrm{SWE}} = \mathrm{SO}(3)$. For $R \in \mathrm{SO}(3)$, define

$$[\rho_0(R)h](x) = h(R^{-1}x), \qquad [\rho_1(R)v](x) = Rv(R^{-1}x).$$

Proposition C.2 (Shallow-water equivariance)**.** *Let Φ_t^{SWE} denote the solution operator for (C.5). Then,*

$$\Phi_t^{\mathrm{SWE}} \circ (\rho_0(R) \oplus \rho_1(R)) = (\rho_0(R) \oplus \rho_1(R)) \circ \Phi_t^{\mathrm{SWE}}$$

for every $R \in \mathrm{SO}(3)$ for which the solution is defined.

Proof. Rotation by R is an orientation-preserving isometry of $\mathbb{S}^2$. The Levi–Civita connection is preserved by isometries, and so

$$\begin{aligned}
\nabla_{\rho_1(R)v}(\rho_1(R)v) &= \rho_1(R)(\nabla_v v), \\
\mathrm{grad}_{\mathbb{S}^2}(\rho_0(R)h) &= \rho_1(R)(\mathrm{grad}_{\mathbb{S}^2} h), \\
\mathrm{div}_{\mathbb{S}^2}\big((\rho_0(R)h)(\rho_1(R)v)\big) &= \rho_0(R)\,\mathrm{div}_{\mathbb{S}^2}(hv).
\end{aligned}$$

Applying these identities to (C.5) shows that the transformed state solves the same system with transformed initial data; the result follows by uniqueness theory. □

As a note, (C.5) neglects the Coriolis term to preserve $\mathrm{SO}(3)$ symmetry.

Instantiated Equivariant DeepONet. Let

$$X = L^2(\mathbb{S}^2) \oplus L^2(T\mathbb{S}^2)$$

denote the state space for (η, v). Scalar spherical harmonics represent η, while the tangent velocity is represented using the spheroidal (gradient-type) modes $\nabla_{\mathbb{S}^2} Y_{\ell m}$ and the toroidal (curl-type) modes $\hat{n} \times \nabla_{\mathbb{S}^2} Y_{\ell m}$. For every $\ell \geq 1$, the scalar, spheroidal, and toroidal coefficient families are three copies of the same degree-ℓ irreducible representation space V_ℓ of $\mathrm{SO}(3)$, while only the scalar copy is present at $\ell = 0$. Thus, at bandlimit L,

$$E_{\mathrm{SWE},L} = V_0^{(\eta)} \oplus \bigoplus_{\ell=1}^{L} \left(V_\ell^{(\eta)} \oplus V_\ell^{(S)} \oplus V_\ell^{(T)} \right).$$

Let $\mathcal{H}_{\mathrm{SWE},L} : X \to E_{\mathrm{SWE},L}$ denote the typed scalar/vector spherical harmonic transform and $\mathcal{H}_{\mathrm{SWE},L}^{-1} : E_{\mathrm{SWE},L} \to X$ the corresponding inverse transform, implemented using the scalar and vector spherical harmonic transforms provided by `torch_harmonics` (Bonev et al., 2023). The full-observation DeepONet is given by the composition

$$\widehat{\Phi}_{\mathrm{SWE}} = \mathcal{H}_{\mathrm{SWE},L}^{-1} \circ B_\theta \circ \mathcal{H}_{\mathrm{SWE},L}. \tag{C.6}$$

Here, $B_\theta : E_{\mathrm{SWE},L} \to E_{\mathrm{SWE},L}$ is an equivariant branch acting directly on the typed harmonic coefficients.

Equivalently, this may be viewed as a fixed-basis branch–trunk DeepONet. If

$$b = B_\theta(a) \in E_{\mathrm{SWE},L},$$

with coefficient blocks $b_{\ell m}^{(\eta)}$, $b_{\ell m}^{(S)}$, and $b_{\ell m}^{(T)}$, then

$$\begin{aligned}
\widehat{\Phi}_{\mathrm{SWE}}(\eta, v)(x) = \Bigg(& \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} b_{\ell m}^{(\eta)}(\eta, v) Y_{\ell m}(x), \\
& \sum_{\ell=1}^{L} \sum_{m=-\ell}^{\ell} \Big[b_{\ell m}^{(S)}(\eta, v) \nabla_{\mathbb{S}^2} Y_{\ell m}(x) \\
& + b_{\ell m}^{(T)}(\eta, v) \hat{n} \times \nabla_{\mathbb{S}^2} Y_{\ell m}(x) \Big] \Bigg).
\end{aligned}$$

Thus, the branch predicts typed scalar and vector spherical-harmonic coefficients, while the scalar, spheroidal, and toroidal spherical harmonics provide the fixed trunk functions.

The selected branch acts directly on $E_{\mathrm{SWE},L}$ using Schur intertwiners, invariant norm gates, and, for the tensor-product variant, Gaunt interactions. At each $\ell \geq 1$, the Schur intertwiners mix the three multiplicity copies $V_\ell^{(\eta)}$, $V_\ell^{(S)}$, and $V_\ell^{(T)}$ while preserving the common V_ℓ representation factor. The Gaunt interactions provide nonlinear equivariant coupling among retained harmonic degrees and among the three representation copies. Since the scalar, spheroidal, and toroidal families at fixed ℓ are copies of the same V_ℓ under proper rotations, these interactions can be constructed as $\mathrm{SO}(3)$-equivariant tensor-product couplings.

The finite-sensor observation keeps all 64 latitudes and 64 uniformly spaced longitudes. A real field with longitudinal content $|m| \leq 31$ is determined by these 64 samples on every latitude. FFT interpolation reconstructs the complete retained grid field before applying $\mathcal{H}_{\mathrm{SWE}}$, so the remainder of the model in (C.6) is unchanged.

Comparison Models. The sparse DeepONet uses 1024 fixed spatial sensors. The symmetry-aware-trunk model uses harmonic output features without an equivariant branch. The SFNO reference is the spherical spectral operator reported in the main comparison.

Data Generation. Data generation uses the shallow-water solver in `torch_harmonics` (Bonev et al., 2023) on a 64×128 Legendre–Gauss grid with orthonormal spherical harmonics through degree $L = 31$, radius $a = 6{,}371{,}220$ m, gravitational acceleration $g = 9.80616$ m/s^2, mean height $H = 10{,}000$ m, and $\Omega = 0$. Initial conditions are drawn by the solver's random initial-condition law at Mach number 0.2. Five operator steps cover 9000 s, so $\Delta t_{\mathrm{operator}} = 1800$ s; each operator step contains 12 solver steps, giving an internal step of 150 s.

C.3 Linear Thermoelasticity on an Annulus

Equations and Symmetry Groups. Consider the annulus defined by

$$\Omega = \{x \in \mathbb{R}^2 : r_i < |x| < r_o\}, \qquad r_i = 0.018 \text{ m}, \quad r_o = 0.05 \text{ m}.$$

The states of interest are the temperature perturbation θ, displacement $\mathbf{u}$, and velocity $v = \partial_t u$. We will consider the coupling of a thermal equation and linear elasticity equation under plane-strain and isotropic linear elasticity assumptions. In brief, first define Young's modulus E, Poisson ratio ν, thermoelastic stress-temperature modulus β, and linear thermal expansion coefficient α. Then, the Lamé parameters λ and μ may be defined as

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}, \qquad \beta = (3\lambda + 2\mu)\alpha.$$

The 2D plane-strain system with reversible thermo-mechanical coupling is given by

$$\begin{cases} \rho c\, \partial_t \theta - \kappa \Delta \theta + \beta T_0 \operatorname{Tr}(\partial_t \boldsymbol{\varepsilon}) = 0, \\ \rho\, \partial_t^2 \mathbf{u} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{0}, \\ \boldsymbol{\sigma} := 2\mu \boldsymbol{\varepsilon}(\mathbf{u}) + \lambda (\nabla \cdot \mathbf{u}) I - \beta \theta I, \\ \boldsymbol{\varepsilon}(\mathbf{u}) := \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^\top), \end{cases} \tag{C.7}$$

where c is the specific heat capacity, ρ is the density, κ is the thermal conductivity (held constant here), T_0 is the reference temperature, $\boldsymbol{\varepsilon}$ is the strain tensor and $\boldsymbol{\sigma}$ is the stress tensor.

This is classical linear isotropic thermoelastic coupling (Nowacki, 1986). The inner mechanical boundary is clamped, the outer mechanical boundary is traction-free, and both circular boundaries use uniform Robin thermal data with coefficient 50 and zero ambient perturbation. Let Γ_{in} and Γ_{out} denote the inner and outer circular boundaries, respectively. The mechanical boundary conditions are

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_{\mathrm{in}}, \qquad \boldsymbol{\sigma}\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_{\mathrm{out}},$$

corresponding to a clamped inner boundary and a traction-free outer boundary. On both boundaries, the temperature satisfies the Robin condition

$$-\kappa \nabla \theta \cdot \mathbf{n} = 50\, \theta.$$

The symmetry group is $G_{\mathrm{th}} = \mathrm{O}(2)$. For $Q \in \mathrm{O}(2)$, define

$$[\rho_{\mathrm{th}}(Q)(\theta, \mathbf{u}, \mathbf{v})](x) = \big(\theta(Q^{-1}x), Q\mathbf{u}(Q^{-1}x), Q\mathbf{v}(Q^{-1}x)\big). \tag{C.8}$$

Proposition C.3 (Thermoelastic equivariance)**.** *Let Φ_t^{th} denote the solution operator for (C.7) with the stated boundary conditions. Then,*

$$\Phi_t^{\mathrm{th}} \circ \rho_{\mathrm{th}}(Q) = \rho_{\mathrm{th}}(Q) \circ \Phi_t^{\mathrm{th}}$$

for every $Q \in \mathrm{O}(2)$ for which the solution is defined.

Proof. Let $\mathbf{u}_Q(x) = Q\mathbf{u}(Q^{-1}x)$. Orthogonality of Q gives the identities

$$\begin{aligned}
\nabla \mathbf{u}_Q(x) &= Q(\nabla \mathbf{u})(Q^{-1}x)Q^\top, \\
\boldsymbol{\varepsilon}(\mathbf{u}_Q)(x) &= Q\boldsymbol{\varepsilon}(\mathbf{u})(Q^{-1}x)Q^\top, \\
\nabla \cdot \mathbf{u}_Q(x) &= (\nabla \cdot \mathbf{u})(Q^{-1}x).
\end{aligned}$$

It follows from the isotropy equation that

$$\boldsymbol{\sigma}(\theta_Q, \mathbf{u}_Q)(x) = Q\boldsymbol{\sigma}(\theta, \mathbf{u})(Q^{-1}x)Q^\top,$$

and, hence,

$$\nabla \cdot \boldsymbol{\sigma}(\theta_Q, \mathbf{u}_Q)(x) = Q(\nabla \cdot \boldsymbol{\sigma}(\theta, \mathbf{u}))(Q^{-1}x).$$

The term $\partial_t \boldsymbol{\varepsilon}$ is similar, since $\mathrm{Tr}\,(\partial_t \boldsymbol{\varepsilon}) = \nabla \cdot \mathbf{v}$. The scalar Laplacian commutes with isometries, so this term transforms properly. Hence, the transformed fields satisfy the same PDE system. The geometry is clearly preserved by Q. The outward normal transforms as a vector, and so the clamped, traction-free, and uniform radial Robin conditions are preserved. Uniqueness gives the result. □

Instantiated Equivariant DeepONet. Let

$$X_{\mathrm{th}} = L^2(\Omega) \oplus L^2(\Omega; \mathbb{R}^2) \oplus L^2(\Omega; \mathbb{R}^2)$$

denote the thermoelastic state space carrying the $\mathrm{O}(2)$ action in (C.8), and let $\mathcal{P}$ convert Cartesian displacement and velocity to polar components. Under angular reflection, θ, u_r, v_r are even, while u_ϕ, v_ϕ are odd. Let $V_{m,+}$ and $V_{m,-}$ denote the corresponding even- and odd-parity $\mathrm{O}(2)$ Fourier representation spaces at angular frequency m.

Since the radial coordinate is fixed by the $\mathrm{O}(2)$ action, it acts as a multiplicity coordinate rather than introducing an additional representation type. In the numerical realization, retaining N_r radial locations gives N_r copies of each angular representation for each field component. So, for angular frequencies $m = 0, \ldots, M$, define

$$E_{\mathrm{th},M} = \bigoplus_{m=0}^{M} \left[\left(\mathbb{R}^{3N_r} \otimes V_{m,+}\right) \oplus \left(\mathbb{R}^{2N_r} \otimes V_{m,-}\right) \right].$$

Here the factors $\mathbb{R}^{3N_r}$ and $\mathbb{R}^{2N_r}$ index, respectively, the three even and two odd field components across the retained radii.

Let $\mathcal{F}_{\phi,M}$ denote the real angular Fourier transform truncated to frequencies $m = 0, \ldots, M$ and applied independently at each retained radius. Then $\mathcal{F}_{\phi,M} \circ \mathcal{P}$ maps the physical state into $E_{\mathrm{th},M}$, and the full-observation DeepONet is

$$\widehat{\Phi}_{\mathrm{th}} = \mathcal{P}^{-1} \circ \mathcal{F}_{\phi,M}^{-1} \circ B_\theta \circ \mathcal{F}_{\phi,M} \circ \mathcal{P}, \tag{C.9}$$

where $B_\theta : E_{\mathrm{th},M} \to E_{\mathrm{th},M}$ acts directly on the retained typed Fourier coefficients.

We may also see this as a fixed-basis branch–trunk DeepONet. If $b = B_\theta(a) \in E_{\mathrm{th},M}$, then the branch predicts the typed Fourier coefficients, while the real trigonometric functions up to angular frequency M provide the fixed trunk functions. For each polar field component $q \in \{\theta, u_r, u_\phi, v_r, v_\phi\}$,

$$\widehat{\Phi}_{\mathrm{th}}(u)_q(r, \phi) = b_{q,0}(r) + \sum_{m=1}^{M} \left[b_{q,m}^{(c)}(r) \cos(m\phi) + b_{q,m}^{(s)}(r) \sin(m\phi) \right], \tag{C.10}$$

with the coefficients carrying the even or odd reflection type associated with the corresponding polar field component. The inverse angular Fourier transform implements this fixed branch–trunk reconstruction, after which $\mathcal{P}^{-1}$ returns the displacement and velocity fields to Cartesian components.

The selected branch acts on $E_{\mathrm{th},M}$ using parity-aware $\mathrm{O}(2)$ intertwiners, global radial mixing, and, for the Gaunt variant, Fourier Gaunt interactions. At each retained angular frequency m, the linear intertwiners mix the radial and field multiplicity coordinates while preserving the corresponding $\mathrm{O}(2)$ representation factor. The mode-dependent low-rank radial map provides global mixing along the invariant radial coordinate, while the Fourier Gaunt product described in Appendix A.3 provides nonlinear equivariant coupling between angular frequencies.

At a fixed radius, the retained real trigonometric space for $m = 0, \ldots, 32$ has dimension $1+2\cdot 32 = 65$. The finite-sensor model observes 65 angles on each of 64 radii, verifies full rank of the angular evaluation matrix, converts vector samples into polar coordinates, and reconstructs all retained modes before the branch in (C.9).

Comparison Models. The sparse DeepONet uses fixed spatial samples. The symmetry-aware-trunk model uses polar harmonic output features but does not constrain the branch by the $\mathrm{O}(2)$ representation. The Polar FNO operates in polar coordinates and has width 48, four layers, 16 radial modes, and 33 angular modes.

Data Generation. The material constants are $\rho = 2700$, $c = 900$, $\kappa = 205$, $E = 70 \times 10^9$, $\nu = 0.33$, $\alpha = 2.3 \times 10^{-5}$, and $T_0 = 293.15$ in SI units. Initial conditions are sampled from Matérn-$5/2$ random fields. The temperature field has correlation length 0.008 m and is scaled to an RMS amplitude of 5 K. The displacement and velocity fields have correlation length 0.012 m; displacement is scaled to an equivalent-strain RMS of 8×10^{-5}, and velocity is scaled so that the ratio of the kinetic to elastic energy is 0.5, which are defined via

$$\mathcal{E}_{kinetic} = \frac{1}{2}\int_{\Omega} \rho|\mathbf{v}|^2\, dx, \qquad \mathcal{E}_{elastic} = \int_{\Omega} \left(\mu\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) + \frac{\lambda}{2}\left|\nabla\cdot\mathbf{u}\right|^2\right) dx,$$

respectively.

The finite-element generator uses an annular mesh with 32 radial and 256 angular cells. Since the linear thermoelastic problem is $\mathrm{O}(2)$-invariant, the finite-element system decouples across angular Fourier modes. Modes through $m = 32$ are retained and propagated independently using matrix exponentials. The resulting finite-element fields are interpolated onto a 64×128 polar grid using piecewise-linear interpolation. Five stored steps cover $0 \le t \le 0.5$, hence $\Delta t_{\mathrm{operator}} = 0.1$.

C.4 Two-Component Gross–Pitaevskii System

Equations and Symmetry Groups. Let $\Omega = (-6, 6)^2$, and suppose that $\Psi = (\psi_1, \psi_2)^\top : \Omega \to \mathbb{C}^2$ solves

$$i\partial_t \Psi = \left[-\kappa\Delta + V(x,y) + g\Psi^\dagger\Psi\right]\Psi, \qquad \kappa = 0.5, \quad g = 10, \tag{C.11}$$

with a centered (anisotropic harmonic trapping) potential

$$V(x,y) = \tfrac{1}{2}\left[(\omega_x x)^2 + (\omega_y y)^2\right], \qquad \omega_x = 0.4, \quad \omega_y = 0.6$$

and homogeneous Dirichlet boundary conditions. The scalar Gross–Pitaevskii equation originates in Gross (1961) and Pitaevskii (1961); (C.11) is the component-symmetric two-field version.

Let $D_2 = \langle r_x, r_y\rangle$ be generated by the two coordinate reflections. The symmetry group is

$$G_{\mathrm{GPE}} = \mathrm{U}(2) \times D_2.$$

For $(U, r) \in G_{\mathrm{GPE}}$, define

$$[\rho_{\mathrm{GPE}}(U,r)\Psi](\mathbf{x}) = U\Psi(r^{-1}\mathbf{x}). \tag{C.12}$$

Proposition C.4 (Gross–Pitaevskii equivariance)**.** *Let Φ_t^{GPE} denote the solution operator for (C.11). Then,*

$$\Phi_t^{\mathrm{GPE}} \circ \rho_{\mathrm{GPE}}(U,r) = \rho_{\mathrm{GPE}}(U,r) \circ \Phi_t^{\mathrm{GPE}}$$

for every $(U, r) \in \mathrm{U}(2) \times D_2$ for which the solution is defined.

Proof. We first consider the internal action. The Laplacian and potential act component-wise, so they commute with any fixed $U \in \mathrm{U}(2)$. Since U is unitary, $(U\Psi)^\dagger(U\Psi) = \Psi^\dagger\Psi$. Thus, every term on the right-hand side of (C.11) transforms by left multiplication with U. Next, consider $r \in D_2 \subset \mathrm{O}(2)$. Since r is orthogonal, the Laplacian commutes with the spatial action. The trapping potential considered is invariant under the coordinate reflections by design (it is even in each coordinate). Finally, each reflection preserves Ω (as it is a square) and its boundary, so the homogeneous Dirichlet condition is preserved. Thus, the equation is equivariant under D_2, as well. The internal $\mathrm{U}(2)$ action and spatial D_2 action commute, so the equation is equivariant under their product $\mathrm{U}(2) \times D_2$. Uniqueness of the initial-value problem gives the stated identity for the solution operator. □

Instantiated Equivariant DeepONet. Let

$$X_{\mathrm{GPE}} = L^2(\Omega; \mathbb{C}^2)$$

denote the state space carrying the $\mathrm{U}(2) \times D_2$ action in (C.12). The $\mathrm{U}(2)$ factor acts pointwise on the two complex components of Ψ, while D_2 acts on its spatial dependence through reflections of the two coordinate axes. To obtain features with explicit transformation laws under both factors, we decompose the field into the four reflection-parity types of D_2.

For each

$$q = (p_x, p_y) \in \{+1, -1\}^2,$$

let $\chi_q : D_2 \to \{+1, -1\}$ denote the corresponding character, defined on the generating reflections by

$$\chi_q(r_x) = p_x, \qquad \chi_q(r_y) = p_y,$$

and let $\mathbb{C}_{\chi_q}$ denote the corresponding one-dimensional representation space. The projection onto parity type q is

$$\mathcal{P}_q\Psi = \frac{1}{4}\left(\Psi + p_x\,\rho(r_x)\Psi + p_y\,\rho(r_y)\Psi + p_x p_y\,\rho(r_x r_y)\Psi\right), \tag{C.13}$$

so that

$$\rho(r)\mathcal{P}_q\Psi = \chi_q(r)\mathcal{P}_q\Psi, \qquad r \in D_2,$$

and

$$\Psi = \sum_{q \in \widehat{D}_2} \mathcal{P}_q\Psi.$$

A fixed-parity function is determined by its restriction to a chosen fundamental quadrant Ω_+, and so the numerical realization retains each parity component at N_+ points in Ω_+. Thus, define

$$E_{\mathrm{GPE}} = \bigoplus_{q \in \widehat{D}_2} \left[\mathbb{R}^{N_+} \otimes \left(\mathbb{C}^2 \boxtimes \mathbb{C}_{\chi_q}\right)\right].$$

Here, $\mathbb{C}^2$ carries the fundamental representation of $\mathrm{U}(2)$, $\mathbb{C}_{\chi_q}$ records the spatial reflection parity, and $\mathbb{R}^{N_+}$ indexes the retained spatial multiplicity coordinates in the fundamental quadrant.

Let $\mathcal{H}_{\mathrm{GPE}} : X_{\mathrm{GPE}} \to E_{\mathrm{GPE}}$ denote the numerical observation map obtained by projecting onto the four parity components and retaining their values on Ω_+. If $x_1, \ldots, x_{N_+} \in \Omega_+$ are the retained locations, then

$$\mathcal{H}_{\mathrm{GPE}}(\Psi) = \left\{\left((\mathcal{P}_q\Psi)(x_j)\right)_{j=1}^{N_+}\right\}_{q \in \widehat{D}_2}.$$

Let

$$E_{\mathrm{GPE}}^{(K)} = \bigoplus_{q \in \widehat{D}_2} \left[\mathbb{R}^{K} \otimes \left(\mathbb{C}^2 \boxtimes \mathbb{C}_{\chi_q}\right)\right]$$

denote the latent typed representation, and let

$$\mathcal{S}_{\mathrm{GPE},\phi} : E_{\mathrm{GPE}}^{(K)} \to X_{\mathrm{GPE}}$$

denote the learned typed reconstruction determined by the trunk parameters ϕ. Then, the equivariant DeepONet is given by the composition

$$\widehat{\Phi}_{\mathrm{GPE}} = \mathcal{S}_{\mathrm{GPE},\phi} \circ B_\theta \circ \mathcal{H}_{\mathrm{GPE}}, \tag{C.14}$$

where $B_\theta : E_{\mathrm{GPE}} \to E_{\mathrm{GPE}}^{(K)}$. In contrast to the fixed harmonic and Fourier reconstructions used above, $\mathcal{S}_{\mathrm{GPE},\phi}$ is parameterized by a learned symmetry-typed trunk

$$\tau_\phi : \Omega \to \bigoplus_{q \in \widehat{D}_2} \left(\mathbb{C}^K \otimes \mathbb{C}_{\chi_q}\right).$$

The trunk is invariant under the $\mathrm{U}(2)$ factor and D_2-equivariant, so

$$\tau_\phi(r \cdot x) = \rho_T(r)\tau_\phi(x), \qquad r \in D_2,$$

where $\rho_T(r)$ acts on the qth parity block by multiplication by $\chi_q(r)$.

Equivalently, this may be written in branch–trunk form. If

$$b = B_\theta\big(\mathcal{H}_{\mathrm{GPE}}(\Psi)\big) \in E_{\mathrm{GPE}}^{(K)},$$

then, for each parity type q and latent index k,

$$b_{q,k}(\Psi) \in \mathbb{C}^2 \boxtimes \mathbb{C}_{\chi_q}.$$

Writing the corresponding component of the learned trunk as

$$\tau_{q,k} : \Omega \to \mathbb{C},$$

we have

$$\tau_{q,k}(r \cdot x) = \chi_q(r)\tau_{q,k}(x),$$

while $\tau_{q,k}$ is invariant under $\mathrm{U}(2)$. The DeepONet output is then

$$\widehat{\Phi}_{\mathrm{GPE}}(\Psi)(x) = \sum_{q \in \widehat{D}_2} \sum_{k=1}^{K} b_{q,k}(\Psi)\, \tau_{q,k}(x). \tag{C.15}$$

For each parity type, multiplication gives the bilinear intertwiner

$$C_q : \left(\mathbb{C}^2 \boxtimes \mathbb{C}_{\chi_q}\right) \times \mathbb{C}_{\chi_q} \to \mathbb{C}^2, \qquad C_q(b, \tau) = b\tau.$$

Since $\chi_q(r)^2 = 1$, the matching branch and trunk parity factors contract to the trivial fiber action under D_2, while the spatial dependence of the trunk supplies the required transformation of the output argument. The fundamental $\mathrm{U}(2)$ action carried by the branch persists. Consequently, the output transforms with the required $\mathrm{U}(2)$ fiber action and D_2 spatial action.

The trunk parity is enforced directly in its parameterization. The learned scalar dependence is a function only of the reflection-invariant coordinates (x^2, y^2). If $g_{q,k}$ denotes this learned invariant dependence, then

$$\tau_{q,k}(x, y) = s_q(x, y)\, g_{q,k}(x^2, y^2),$$

where

$$s_{(+,+)}(x, y) = 1, \qquad s_{(-,+)}(x, y) = x, \qquad s_{(+,-)}(x, y) = y, \qquad s_{(-,-)}(x, y) = xy.$$

Thus, the factors $1, x, y, xy$ impose the four desired reflection parities on otherwise reflection-invariant learned trunk functions.

The selected branch acts on E_{GPE} using equivariant multiplicity maps together with invariant gating or the cubic Gram interaction described in Appendix A.3. Linear layers mix spatial and hidden multiplicity coordinates while preserving the corresponding $\mathrm{U}(2) \times D_2$ representation type. For hidden doublets $z_a, z_b, z_c \in \mathbb{C}^2$, the cubic interaction forms the $\mathrm{U}(2)$-invariant Hermitian product $z_a^\dagger z_b$ and uses it to construct the covariant doublet $(z_a^\dagger z_b)z_c$. Indeed, for $U \in \mathrm{U}(2)$, $(Uz_a)^\dagger(Uz_b) = z_a^\dagger z_b$, so the resulting cubic feature again transforms in the fundamental representation. The associated D_2 parity types combine multiplicatively. Numerically, each complex doublet is stored using four real coordinates according to the realification convention in Appendix A.4.

The finite-sensor model observes a reflection-closed 64×64 subset of the 128×128 interior grid. A 64-mode tensor-product DST-I reconstruction is applied before the symmetry-adapted map $\mathcal{H}_{\mathrm{GPE}}$, after which the branch, trunk, and fusion in (C.14) are unchanged. Reflection closure gives the D_2 sensor representation in Theorem B.2. The same real reconstruction is applied to each of the four stored real channels, so it also commutes with the realified $\mathrm{U}(2)$ action.

Comparison Models. The sparse DeepONet uses fixed point sensors and unconstrained branch and trunk MLPs. No separate symmetry-aware-trunk baseline is reported for this problem: the dominant $\mathrm{U}(2)$ action is internal and does not supply a natural harmonic output basis analogous to those in the other three benchmarks. The FNO reference acts on the four real channels with standard learned Fourier multipliers.

Data Generation. The numerical method uses an orthonormal two-dimensional DST-I for the kinetic half steps and an exact pointwise nonlinear/potential phase for the full step, giving a second-order Strang splitting (Strang, 1968). The grid contains 128×128 interior points, the internal solver time step is 0.002, and five operator steps cover $0 \le t \le 0.5$, so $\Delta t_{\mathrm{operator}} = 0.1$.

Initial states are drawn as two independent proper-complex Matérn fields with smoothness 2.5 and correlation length 2.5, followed by rescaling to total mass one. The component covariance is proportional to I_2, so the initial-state law is $\mathrm{U}(2)$-invariant. The covariance is defined through the Dirichlet-Laplacian spectrum, which is preserved by the coordinate reflections in D_2, giving reflection invariance. Mass normalization is likewise invariant under both actions.

C.5 Dataset and Observation Summaries

Table C.1: Dataset summary. Every trajectory contains six frames, giving five learned timesteps.

Problem	Grid Size	Spectral Cutoff	$\Delta t_{\mathrm{operator}}$	Train/Val/Test Trajectory Split
Allen–Cahn	64×128	$L = 22$	0.2	4096/256/512
Shallow water	64×128	$L = 31$	1800 s	4096/256/512
Thermoelasticity	64×128	$\lvert m\rvert \le 32$	0.1	4096/256/512
Gross–Pitaevskii	128×128	full DST-I grid	0.1	4096/256/512

C.6 Model Parameterization Summary

For completeness, Table C.2 summarizes the core parameters used in the experiments. Here, p denotes the conventional DeepONet latent dimension, while multiplicity denotes the principal representation multiplicity of the equivariant model. We will use L for the spherical-harmonic bandlimit and M for the maximum retained angular Fourier frequency. The number of equivariant blocks refers to the depth of the representation-valued branch. The finite-sensor equivariant models use the same equivariant cores as their full-field counterparts and differ only in the observation and reconstruction procedure.

D Training and Evaluation

This appendix records the optimization protocol and the exact scope of the reported metrics. The experiment-specific representation and model constructions are given with the corresponding PDEs in Appendix C.

D.1 Training Setup

All models are trained with an AdamW optimizer and a cosine-annealing learning-rate scheduler. The initial learning rate is 3×10^{-4}, the final learning rate is 10^{-7}, and the weight decay is 10^{-4}. Each model is trained for 200 epochs. Allen–Cahn and Gross–Pitaevskii use batch sizes of 16, whereas shallow water and thermoelasticity use batch sizes of 8. Models chosen for inference were selected by validation loss without reference to the held-out test set trajectories. The training objective is to predict one timestep forward within a given trajectory. The loss objective is the relative L^2 error, defined via appropriate spatial quadrature weights and averaged over the channel and batch dimensions; this is defined in the subsequent subsection.

The FNO baselines are implemented using the `neuralop` library (Kossaifi et al., 2026), with the spherical transforms supplied by `torch_harmonics` where applicable (Bonev et al., 2023).

Problem / model	Architecture	Resolution
Allen–Cahn		
Sparse vanilla DeepONet	$p = 512$, MLP 4×512	1024 sensors
Symmetry-aware trunk DeepONet	$p = 529 = (22+1)^2$, MLP 4×512	$L = 22$
Full-field equivariant DeepONet	multiplicity 4, 2 blocks	$L = 22$
Finite-sensor equivariant DeepONet	same core	640 sensors, $L = 22$
SFNO	width 64, 4 layers	23×23 modes
Shallow water		
Sparse vanilla DeepONet	$p = 512$, MLP 4×512	1024 sensors
Symmetry-aware trunk DeepONet	$p = 1024 = (31+1)^2$, MLP 4×1280	$L = 31$
Full-field equivariant DeepONet	multiplicity 28, 3 blocks	$L = 31$
Finite-sensor equivariant DeepONet	same core	4096 sensors, $L = 31$
SFNO	width 64, 4 layers	32×32 modes
Thermoelasticity		
Sparse vanilla DeepONet	$p = 512$, MLP 4×512	1024 sensors
Symmetry-aware trunk DeepONet	$p = 512$, MLP 4×512	full field
Full-field equivariant DeepONet	multiplicity 10, 3 blocks	$M = 32$, rank 8
Finite-sensor equivariant DeepONet	same core	4160 sensors, $M = 32$
Polar FNO	width 48, 4 layers	$16/33$ modes
Gross–Pitaevskii		
Sparse vanilla DeepONet	$p = 512$, MLP 4×512	1024 sensors
Full-field equivariant DeepONet	multiplicity 256, 3 blocks	trunk 4×512
Finite-sensor equivariant DeepONet	same core	4096 sensors, 64 DST
FNO	width 64, 4 layers	24×24 modes

Table C.2: Model parameters for all experiments.

Table D.1: Common optimization settings.

Problem	Epochs	Batch	Initial LR	Minimum LR	Weight decay
Allen–Cahn	200	16	3×10^{-4}	10^{-7}	10^{-4}
Shallow water	200	8	3×10^{-4}	10^{-7}	10^{-4}
Thermoelasticity	200	8	3×10^{-4}	10^{-7}	10^{-4}
Gross–Pitaevskii	200	16	3×10^{-4}	10^{-7}	10^{-4}

The total training time, average per-sample inferencing time, and peak memory usage are contained in Table D.2. The models were trained on an L40S GPU for Allen–Cahn, shallow water, and Gross–Pitaevskii, but an H100 for thermoelasticity. Inference was performed on H100 GPUs.

D.2 Prediction Metrics

Let z be a target field, $\widehat{z}$ its prediction, and $w_q > 0$ the spatial quadrature weights. For a physical field type c consisting of one or more components, define the relative L^2 error via

$$e_c(\widehat{z}, z) = \left[\frac{\sum_q w_q \|\widehat{z}_c(q) - z_c(q)\|_2^2}{\max\{\sum_q w_q \|z_c(q)\|_2^2, \varepsilon_{\text{num}}\}} \right]^{1/2}, \qquad \varepsilon_{\text{num}} = 10^{-12}.$$

For vector-valued fields, c denotes the full vector rather than its individual coordinate components.

This is then averaged over physical types: one scalar field for Allen–Cahn; height and tangent velocity for shallow water; temperature, displacement, and velocity for thermoelasticity; and the full complex doublet for Gross–Pitaevskii. The one-step score uses all $N_1 = 512 \times 5 = 2560$ held-out adjacent transitions. The five-step score uses $N_5 = 512$ autoregressive endpoint predictions from frame 0 to frame 5. The reported standard error is the sample standard deviation divided by $\sqrt{N}$ and quantifies held-out-example variability for a fixed checkpoint.

Table D.2: Training and inference cost for the models used in publication inference. Training time is rounded to the nearest hour, one-step inference time to the nearest millisecond, and peak memory to the nearest 0.1 GB.

Model	Training (h)	Inference (ms)	Peak memory (GB)
Allen–Cahn			
Sparse DeepONet	1	1	0.2
Symmetry-aware trunk DeepONet	1	1	0.1
Equivariant DeepONet (ours)	2	6	0.3
Finite-sensor equivariant DeepONet (ours)	2	6	0.3
SFNO	2	3	1.0
Shallow water			
Sparse DeepONet	3	1	0.4
Symmetry-aware trunk DeepONet	4	2	0.3
Equivariant DeepONet (ours)	8	5	2.0
Finite-sensor equivariant DeepONet (ours)	8	6	2.0
SFNO	5	3	0.6
Thermoelasticity			
Sparse DeepONet	5	1	0.3
Symmetry-aware trunk DeepONet	6	1	0.5
Equivariant DeepONet (ours)	8	2	0.2
Finite-sensor equivariant DeepONet (ours)	8	2	0.3
Polar FNO	6	1	0.5
Gross–Pitaevskii			
Vanilla DeepONet	8	2	0.7
Equivariant DeepONet (ours)	8	2	0.9
Finite-sensor equivariant DeepONet (ours)	8	3	1.0
FNO	12	2	2.1

D.3 Equivariance Metrics

For a sampled group element g, the generic relative defect is

$$d_g(z) = \frac{\|\widehat{\mathcal{F}}(\rho(g)z) - \rho(g)\widehat{\mathcal{F}}(z)\|}{\max\{\|\rho(g)\widehat{\mathcal{F}}(z)\|, \varepsilon_{\text{num}}\}}.$$

The norm is chosen consistently with the geometry and field typing (e.g., spherical scalar functions have metrics computed via spherical quadrature).

Equivariance is evaluated on the t_0 state of each of 512 held-out trajectories, with paired one-step and five-step inference. For each continuous symmetry factor, the defect is averaged over multiple group elements sampled independently of the held-out trajectories. Specifically, Allen–Cahn and shallow water use 64 Haar-sampled $\mathrm{SO}(3)$ rotations. Thermoelasticity separately averages over 64 non-identity rotations and 64 reflected elements of the discrete realization of the $\mathrm{O}(2)$ action. Gross–Pitaevskii averages over 64 Haar-sampled $\mathrm{U}(2)$ matrices and, for the spatial D_2 factor, all three non-identity group elements. The $\mathbb{Z}_2$ factor of Allen–Cahn has only one non-identity element and is evaluated using the sign action directly.

For each rollout horizon h and each separately tested group action j, let

$$d_{h,j} = \frac{1}{NK_j} \sum_{i=1}^{N} \sum_{k=1}^{K_j} d_{g_{j,k}}(z_i), \qquad N = 512,$$

where K_j is the number of sampled non-identity group elements for factor j. In Table 4.1, the aggregate value is

$$\mathrm{Eq}_h = \max_{j \in \mathcal{G}} d_{h,j},$$

where $\mathcal{G}$ indexes the separately tested group actions: rotation and sign for Allen–Cahn, rotation for shallow water, rotation and reflection for thermoelasticity, and internal $\mathrm{U}(2)$ and spatial D_2 for Gross–Pitaevskii. Table D.3 reports the individual components.

Table D.3: Factorwise equivariance defects across each experiment; these are the components whose maximum is reported as Eq_h in Table 4.1; lower is better throughout; bold denotes the best value within an experiment and factor.

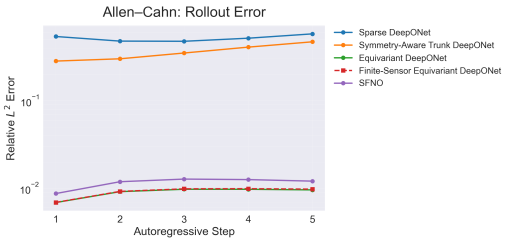
Model	Factor	$h = 1$	$h = 5$
Allen–Cahn			
Vanilla DeepONet	$\mathrm{SO}(3)$	4.0×10^{-1}	7.3×10^{-1}
	$\mathbb{Z}_2$	2.9×10^{-1}	7.9×10^{-1}
Symmetry-aware trunk DeepONet	$\mathrm{SO}(3)$	3.0×10^{-1}	5.6×10^{-1}
	$\mathbb{Z}_2$	2.5×10^{-1}	7.3×10^{-1}
Equivariant DeepONet (ours)	$\mathrm{SO}(3)$	$\mathbf{8.9\times10^{-7}}$	$\mathbf{1.3\times10^{-6}}$
	$\mathbb{Z}_2$	$\mathbf{1.2\times10^{-7}}$	$\mathbf{3.4\times10^{-7}}$
Finite-sensor equivariant DeepONet (ours)	$\mathrm{SO}(3)$	2.8×10^{-6}	9.8×10^{-6}
	$\mathbb{Z}_2$	$\mathbf{1.2\times10^{-7}}$	7.0×10^{-7}
SFNO	$\mathrm{SO}(3)$	1.0×10^{-2}	1.6×10^{-2}
	$\mathbb{Z}_2$	7.8×10^{-3}	1.2×10^{-2}
Shallow water			
Vanilla DeepONet	$\mathrm{SO}(3)$	6.9×10^{-1}	1.1×10^{0}
Symmetry-aware trunk DeepONet	$\mathrm{SO}(3)$	1.6×10^{-1}	4.5×10^{-1}
Equivariant DeepONet (ours)	$\mathrm{SO}(3)$	$\mathbf{1.8\times10^{-6}}$	$\mathbf{2.8\times10^{-6}}$
Finite-sensor equivariant DeepONet (ours)	$\mathrm{SO}(3)$	$\mathbf{1.8\times10^{-6}}$	2.9×10^{-6}

Model	Factor	$h = 1$	$h = 5$
SFNO	$\mathrm{SO}(3)$	4.1×10^{-2}	1.3×10^{-1}
Thermoelasticity			
Vanilla DeepONet	rotation	8.1×10^{-2}	1.1×10^{-1}
	reflection	8.5×10^{-2}	1.2×10^{-1}
Symmetry-aware trunk DeepONet	rotation	6.9×10^{-2}	9.4×10^{-2}
	reflection	7.0×10^{-2}	9.7×10^{-2}
Equivariant DeepONet (ours)	rotation	$\mathbf{3.4\times10^{-7}}$	$\mathbf{1.4\times10^{-6}}$
	reflection	$\mathbf{3.7\times10^{-7}}$	$\mathbf{1.4\times10^{-6}}$
Finite-sensor equivariant DeepONet (ours)	rotation	4.3×10^{-7}	1.9×10^{-6}
	reflection	4.5×10^{-7}	1.9×10^{-6}
Polar FNO	rotation	9.4×10^{-3}	2.6×10^{-2}
	reflection	1.0×10^{-2}	2.9×10^{-2}
Gross–Pitaevskii			
Vanilla DeepONet	$\mathrm{U}(2)$	1.8×10^{-1}	2.6×10^{-1}
	D_2	2.0×10^{-1}	2.8×10^{-1}
Equivariant DeepONet (ours)	$\mathrm{U}(2)$	$\mathbf{2.3\times10^{-6}}$	$\mathbf{5.0\times10^{-6}}$
	D_2	$\mathbf{9.1\times10^{-7}}$	$\mathbf{1.9\times10^{-6}}$
Finite-sensor equivariant DeepONet (ours)	$\mathrm{U}(2)$	2.4×10^{-6}	5.1×10^{-6}
	D_2	9.4×10^{-7}	2.0×10^{-6}
FNO	$\mathrm{U}(2)$	1.0×10^{-2}	1.8×10^{-2}
	D_2	1.1×10^{-2}	1.6×10^{-2}

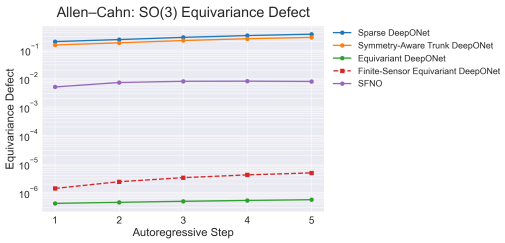
E ADDITIONAL FIGURES

This appendix collects held-out rollout, endpoint-comparison, and equivariance diagnostics for the four benchmark problems. Line graphs show the rollout of the relative L^2 error and the equivariance metrics. The qualitative graphs show the ground truth and prediction after either 1 or 5 steps in the first row. The second row shows the pointwise absolute errors on a logarithmic color scale dictated by the joint minima and maxima of the models featured in the figure.

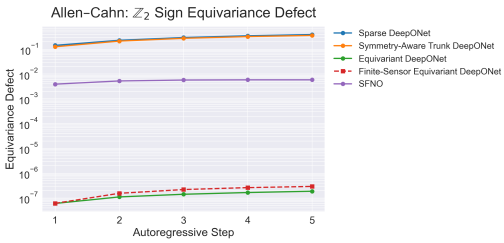
The qualitative panels show the following physical or derived scalar fields. For Allen–Cahn, we display the phase field u. For shallow water, we display the stored height perturbation $\eta = h - H$ and the speed $\|v\|_2$. For thermoelasticity, we display the temperature perturbation θ, displacement magnitude $\|u\|_2$, and velocity magnitude $\|v\|_2$. For Gross–Pitaevskii, writing $\Psi = (\psi_1, \psi_2)^\top$, we plot the total density $n = |\psi_1|^2 + |\psi_2|^2$ and the z-component of the pseudospin $S_z = |\psi_1|^2 - |\psi_2|^2$.



(a) Prediction error across rollout

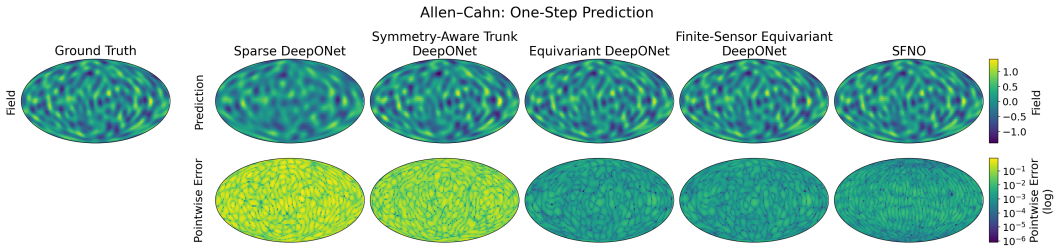


(b) $\mathrm{SO}(3)$ equivariance

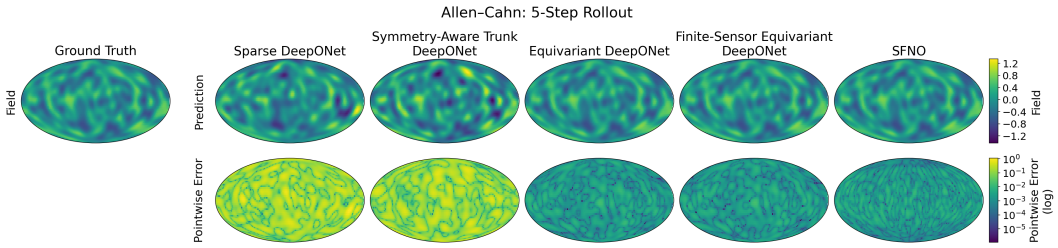


(c) $\mathbb{Z}_2$ sign equivariance

Figure E.1: Allen–Cahn rollout and equivariance diagnostics

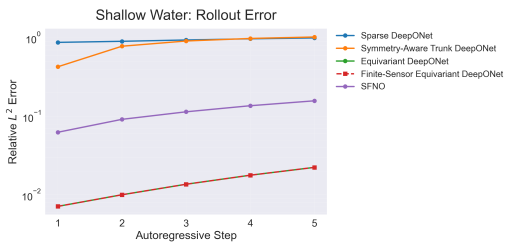


(a) One-step comparison

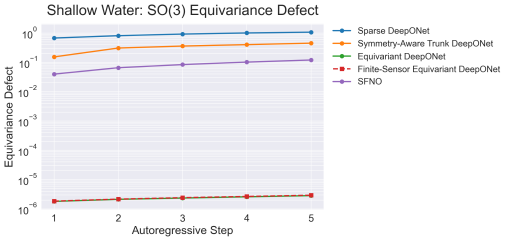


(b) Five-step comparison

Figure E.2: Allen–Cahn held-out field comparisons at one- and five-step horizons

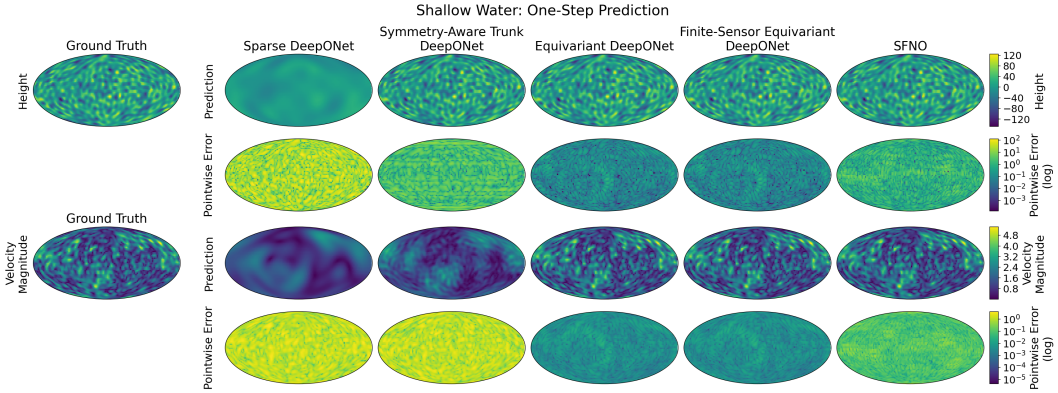


(a) Prediction error across rollout

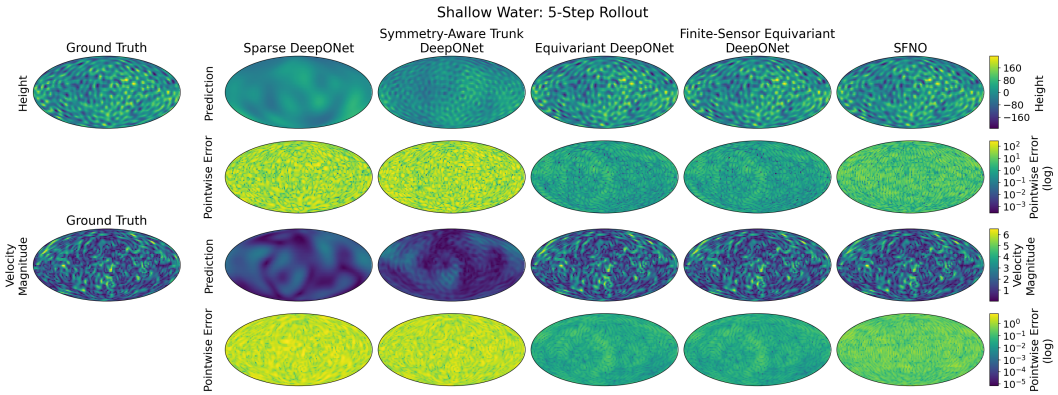


(b) $\mathrm{SO}(3)$ equivariance

Figure E.3: Shallow-water rollout and equivariance diagnostics

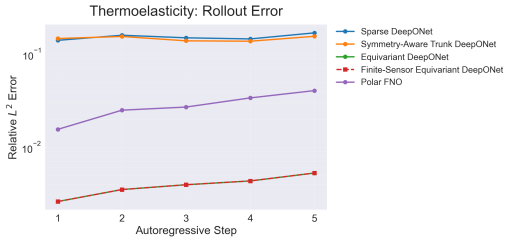


(a) One-step comparison

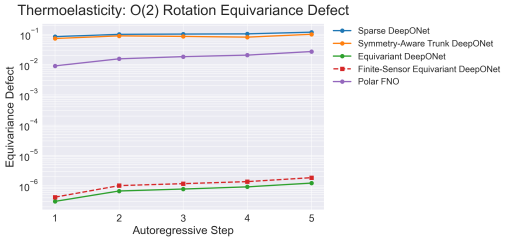


(b) Five-step comparison

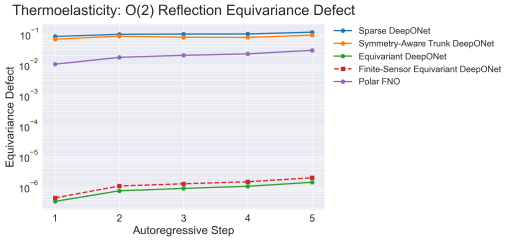
Figure E.4: Shallow-water held-out field comparisons at one- and five-step horizons



(a) Prediction error across rollout

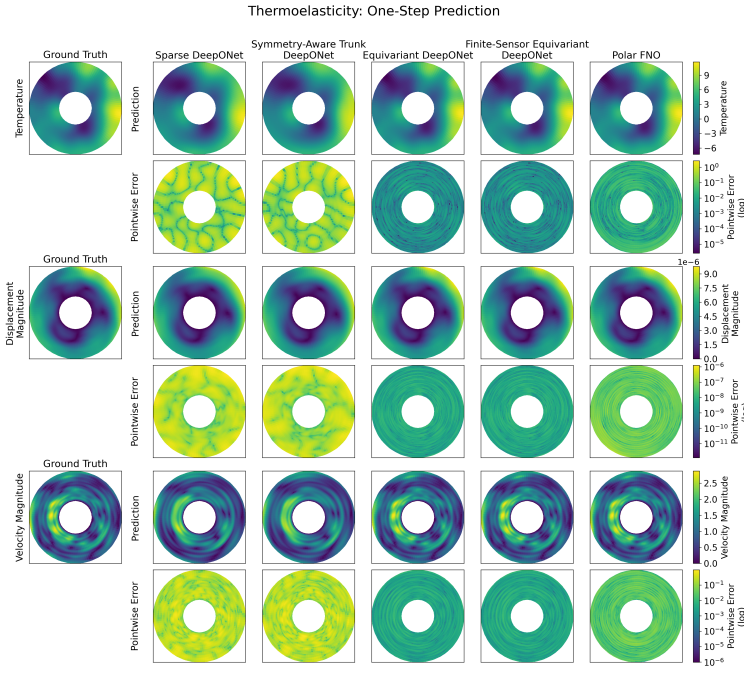


(b) Rotational component

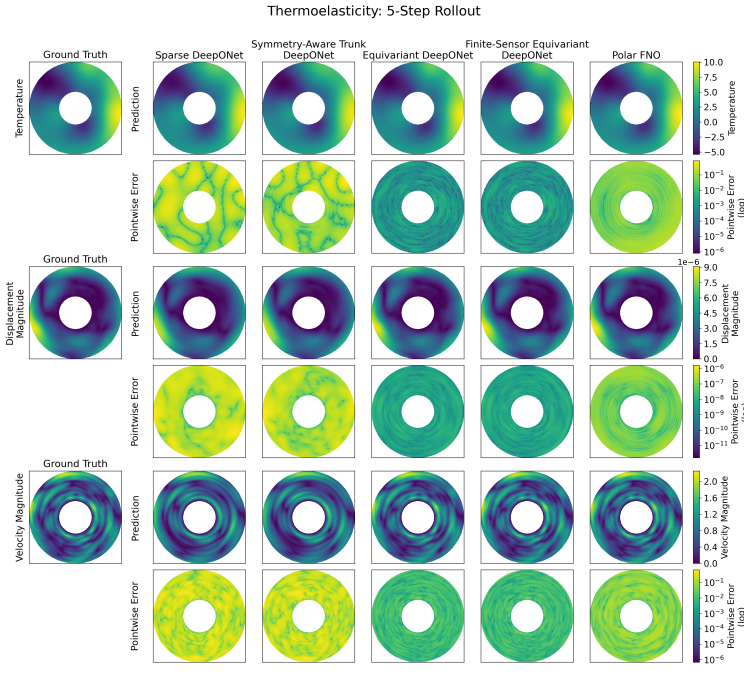


(c) Reflection component

Figure E.5: Thermoelastic rollout and $\mathrm{O}(2)$ equivariance diagnostics

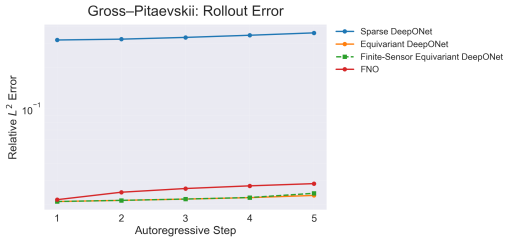


(a) One-step comparison

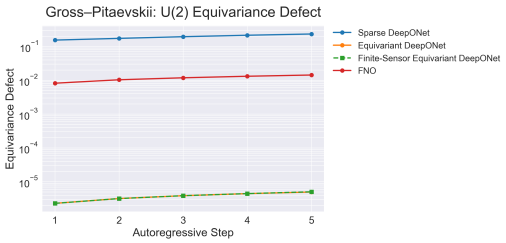


(b) Five-step comparison

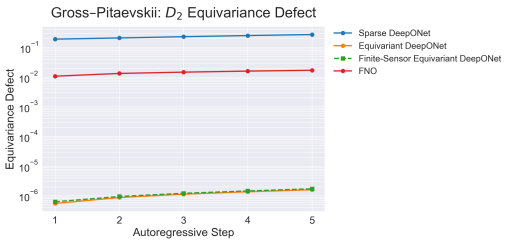
Figure E.6: Thermoelastic held-out field comparisons at one- and five-step horizons



(a) Prediction error across rollout

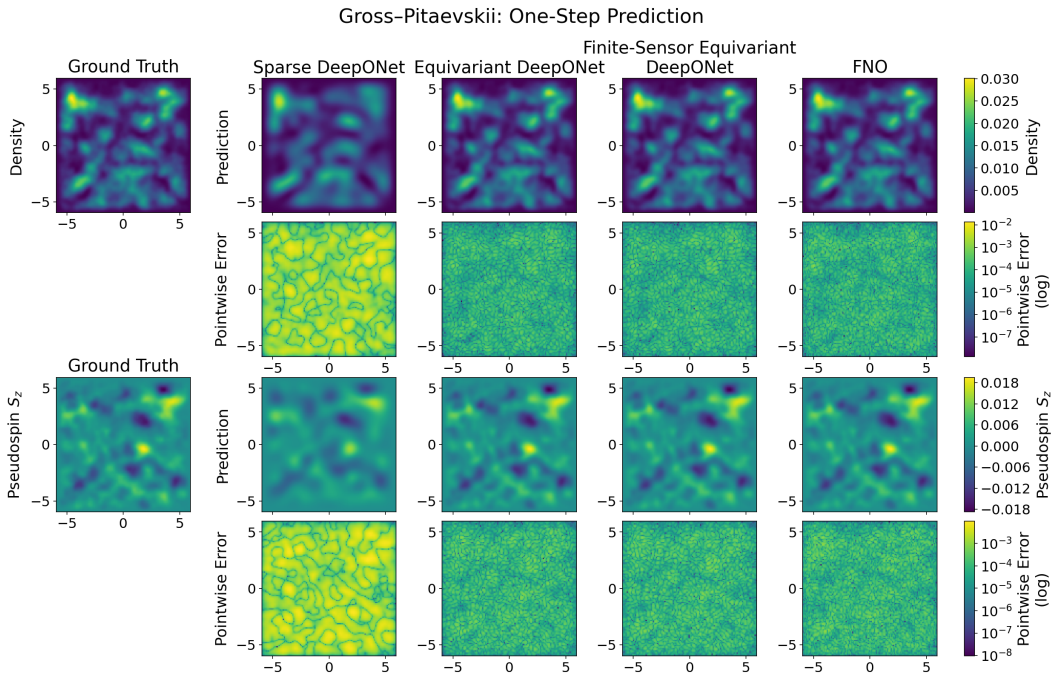


(b) Internal $\mathrm{U}(2)$ equivariance

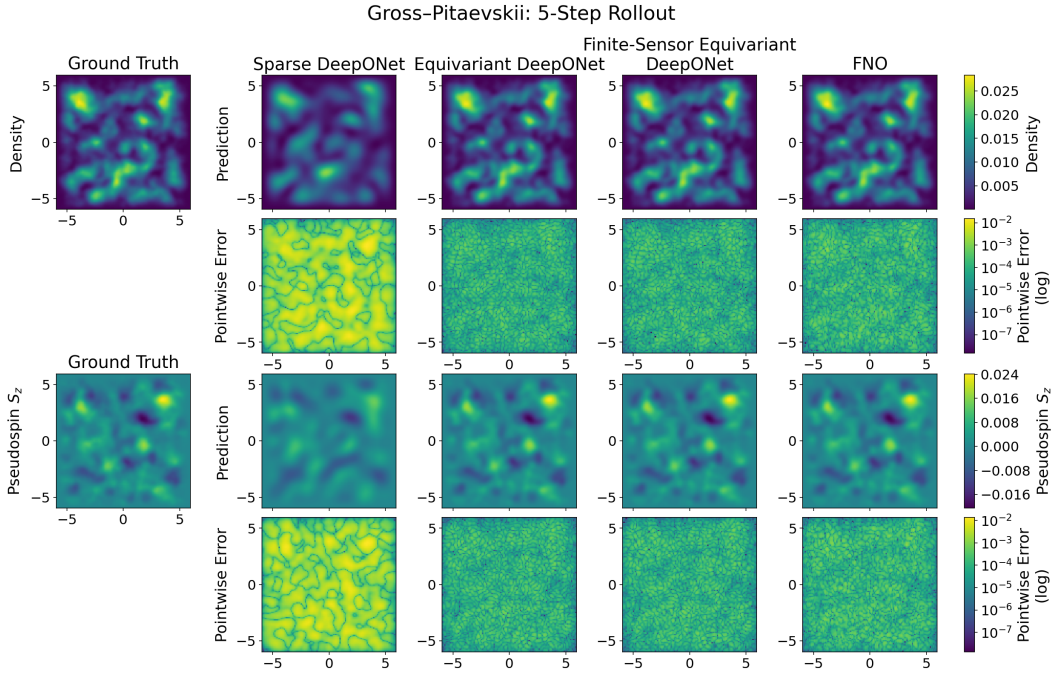


(c) Spatial D_2 equivariance

Figure E.7: Gross–Pitaevskii rollout and equivariance diagnostics for the internal $\mathrm{U}(2)$ and spatial D_2 actions



(a) One-step comparison



(b) Five-step comparison

Figure E.8: Gross–Pitaevskii held-out field comparisons at one- and five-step horizons