MESH-ADMISSIBLE STATE UPDATES IN GRAPH NEURAL SIMULATORS VIA MESH–GENERIC COMPATIBILITY*

Anonymous authors
Paper under double-blind review

ABSTRACT

Graph neural simulators (GNSs) learn state evolution on discretized physical domains, but trajectory accuracy alone does not certify that a learned update is physically admissible. Structure-preserving methods already incorporate physical and geometric structure into graph- and mesh-based simulators. We ask when the GENERIC reversible–dissipative algebraic contract can hold on a prescribed mesh state space, with its metric, admissible directions, and boundary ports governing the executed dynamics. We call this condition Mesh–GENERIC Compatibility (MGC) and give a constructive sufficient condition that enforces skew-adjointness, positive semidefiniteness, and energy–entropy degeneracy for every admissible neural readout, independent of training. MAGNS realizes this construction through neural dynamical assembly, which maps learned constitutive and operator quantities into the MGC-compatible operator family before numerical integration. Controlled interventions test integration, boundary exchange, operator assembly, and Hodge-space admissibility. On long-horizon acoustics, MAGNS matches the strongest structured baseline in median rollout accuracy while reducing mean rollout MSE by 6.3% and attaining a smaller common-chart energy gap. Darcy interventions show that changing the declared Hodge space controls harmonic circulation with fixed neural weights, while errors in the remaining flux components are nearly unchanged.

1 INTRODUCTION

Graph neural simulators (GNSs) learn physical evolution on graph- and mesh-structured domains (Sanchez-Gonzalez et al., 2020; Pfaff et al., 2021). A typical GNS encodes the current physical state, propagates information by message passing, and decodes a state change that is fed back during rollout. This is flexible, but trajectory supervision usually determines the update without specifying which state changes are physically admissible. Accurate prediction alone does not establish whether the executed update respects the mesh geometry or the reversible–dissipative structure of the dynamics.

Several lines of work add structure to learned dynamics. Thermodynamics-informed graph models and neural metriplectic/GENERIC systems constrain reversible and irreversible evolution through energy–entropy structure (Hernández et al., 2024; Tierz et al., 2025; Bermejo et al., 2025; Gruber et al., 2025; Sulskis & Ravi, 2026). IGNS and MeshFT incorporate physical structure directly into geometric graph or mesh-based simulation (Hoang et al., 2026; Noguchi & Kawahara, 2026). MeshFT, in particular, derives a port-Hamiltonian factorization that separates topology-fixed interconnection from metric- and constitutive-dependent structure. We ask: when can the GENERIC reversible–dissipative algebraic contract be satisfied by dynamics executed on the same prescribed mesh state space that determines the update's metric, admissible directions, and boundary ports?

We call this condition Mesh–GENERIC Compatibility (MGC). It requires both mesh admissibility and GENERIC's reversible–dissipative algebra on the same continuous-time dynamics. Neither requirement implies the other: a GENERIC operator constructed in an ambient state space may leave the prescribed mesh subspace, while a geometrically admissible mesh evolution need not satisfy

*Code: `https://anonymous.4open.science/r/MAGNS-CB68/`.

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.

the GENERIC degeneracy relations. Building on GENERIC/metriplectic formulations (Grmela & Öttinger, 1997; Gruber et al., 2025; Sulskis & Ravi, 2026) and discrete Hodge geometry (Hirani, 2003; Yang et al., 2024; Zheng et al., 2026), we derive a mesh-restricted operator construction sufficient for MGC. The construction enforces skew-adjointness, positive semidefiniteness, and energy–entropy degeneracy under the declared mesh metric and admissible directions for every admissible neural readout, independently of training (Theorem 1); open-system exchange uses the declared boundary ports.

To turn this construction into a learnable simulator, we introduce the mesh-admissible graph neural simulator (MAGNS). In MAGNS, neural dynamical assembly combines learned constitutive and operator quantities with the fixed mesh declaration to produce an MGC-compatible continuous-time vector field before numerical integration. The mesh metric, admissible directions, and boundary exchange specify the state space and ports; the network learns only the constitutive and operator degrees of freedom left open by the compatibility conditions.

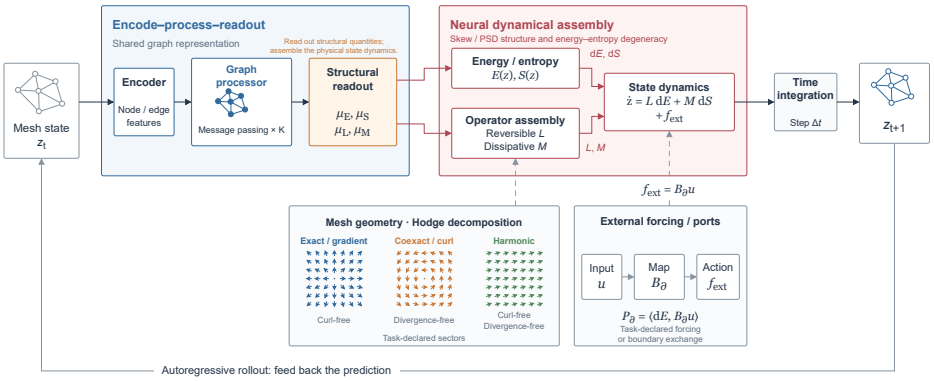

Figure 1: Mesh–GENERIC compatibility in MAGNS. The mesh declaration fixes admissible bulk directions and boundary ports. Neural dynamical assembly combines it with learned structural quantities to form an MGC-compatible continuous-time generator, which the numerical integrator advances from z_t to z_{t+1}.

Our contributions are threefold. **First, we formulate MGC as a joint condition on mesh dynamics** and derive a sufficient construction that enforces it. We also characterize the full pointwise operator class under full-range lifts and unrestricted coefficient blocks (Proposition 1). **Second, we realize this construction in MAGNS**, separating structural quantities prescribed by the discretization from constitutive and operator degrees of freedom learned from data. **Third, we test the distinctions implied by MGC through controlled interventions**. Conservative waves, open acoustics, Beam3D, and Darcy flow examine integration, boundary exchange, operator assembly, and Hodge-space admissibility, respectively. Prediction error, algebraic compatibility residuals, and finite-step balance defects distinguish the corresponding properties. In Darcy flow, changing only the declaration admits or suppresses global circulation without retraining, demonstrating control that trajectory error alone would not reveal.

2 Related Work

Graph neural simulation on meshes. Interaction Networks and graph neural simulators established the encode–process–decode paradigm for learning physical evolution on graph-structured domains (Battaglia et al., 2016; Sanchez-Gonzalez et al., 2020), with MeshGraphNets extending this interface to unstructured meshes (Pfaff et al., 2021). Later work develops message-passing PDE solvers and long-range communication (Brandstetter et al., 2022; Yu et al., 2025), mesh movement and masked training (Hu et al., 2024; Garnier et al., 2025), rollout refinement (Lippe et al., 2023), and hierarchical or generative formulations (Würth et al., 2025). This line of work mainly changes how state evolution is represented, propagated, and learned.

Reversible–dissipative structure in learned dynamics. Hamiltonian, Lagrangian, symplectic, and port-Hamiltonian neural models constrain learned evolution through prescribed physical structure, conservation laws, or power balance (Greydanus et al., 2019; Cranmer et al., 2020;

Zhong et al., 2020; Desai et al., 2021). GENERIC couples reversible and irreversible evolution through skew and positive-semidefinite operators with energy–entropy degeneracy conditions (Grmela & Öttinger, 1997; Öttinger, 2006). These conditions have been incorporated into learned models on dense states, graphs, and operators, including structure-preserving neural networks, GFINNs, thermodynamics-informed graph networks, neural metriplectic systems, and GENERIC-based operator learning (Hernández et al., 2021; Zhang et al., 2022; Hernández et al., 2023; 2024; Tierz et al., 2025; Bermejo et al., 2025; Gruber et al., 2025; Sulskis & Ravi, 2026). Projected metriplectic parameterizations enforce degeneracy by restricting operator action to gradient-orthogonal spaces (Gruber et al., 2025; Sulskis & Ravi, 2026). Related structure-preserving operator-learning methods preserve discretization-dependent mathematical structure through function-space or finite-element representations (Bouziani & Boullé, 2024). IGNS brings a Hamiltonian perspective to graph simulation through geometric encoding and port-Hamiltonian latent evolution (Hoang et al., 2026).

Mesh geometry and structured field dynamics. Discrete exterior calculus represents mesh topology through incidence operators and metric information through discrete Hodge stars (Hirani, 2003), while discrete Hodge decomposition resolves fields into exact, coexact, and harmonic components (Lim, 2020). Recent models use these structures on graphs, meshes, and discrete complexes, including Hodge-based graph models and Hodge-aware learning of physical fields (Roddenberry & Segarra, 2019; Yang et al., 2024; Xu & Pan, 2024; Zheng et al., 2026). MeshFT is the closest structured-mesh reference: it derives from mesh-level physical principles a local port-Hamiltonian factorization in which conservative interconnection is fixed by signed incidence, while metric, constitutive, and dissipative structure retain phenomenon-dependent freedom (Noguchi & Kawahara, 2026). MGC asks for GENERIC-compatible dynamics on a prescribed mesh state space. MAGNS combines the declared metric, admissible directions, and ports with learned quantities to assemble compatible operators.

Continuous-time structure and numerical realization. Structural properties of a continuous-time vector field do not automatically carry over to a finite time step. Geometric, symplectic, discrete-gradient, and GENERIC integrators address this numerical layer directly (Hairer et al., 2006; Gonzalez, 1996; McLachlan et al., 1999; Öttinger, 2018; Shang & Öttinger, 2020). MeshFT likewise separates the continuous dynamical representation from its integrator (Noguchi & Kawahara, 2026). We keep the same distinction between continuous-time compatibility and finite-step realization.

3 Formulation: Mesh–GENERIC Compatibility

3.1 Learning updates on a prescribed mesh

Let $\mathcal{K}$ be an oriented mesh and z_t the physical state carried by its nodes, edges, or cells. Graph neural simulation predicts successive states from an initial observation (Sanchez-Gonzalez et al., 2020; Pfaff et al., 2021). Rather than learning a free increment, we learn a continuous-time generator f_θ and integrate it:

$$\hat{z}_{t+1} = \Phi_{\Delta t}[f_\theta](\hat{z}_t; \mathfrak{D}, u_t), \qquad \hat{z}_0 = z_0. \tag{1}$$

Here $\Phi_{\Delta t}$ is a numerical integration map, u_t specifies external input, and $\mathfrak{D}$ declares the mesh constraints defined below. The modeling question is which generators are admissible before the integrator converts them into state updates.

3.2 Mesh geometry and admissible motion

The mesh supplies signed incidences D_0, D_1 with $D_1 D_0 = 0$ and positive-definite Hodge stars $\star_k$. These determine the inner products and adjoints used throughout the construction. For a map $A : C^k(\mathcal{K}) \to C^\ell(\mathcal{K})$ between mesh-state blocks,

$$\langle u, v \rangle_{\star_k} = u^\top \star_k v, \qquad A^\dagger = \star_k^{-1} A^\top \star_\ell . \tag{2}$$

For edge fields, the incidences separate exact, coexact, and harmonic directions: $\operatorname{im} D_0$, $\operatorname{im} D_1^\dagger$, and $\ker D_1 \cap \ker D_0^\dagger$. Their metric-orthogonal projectors satisfy (Hirani, 2003; Lim, 2020)

$$I = \Pi_{\mathrm{ex}} + \Pi_{\mathrm{co}} + \Pi_{\mathrm{har}}, \qquad \Pi_s^\dagger = \Pi_s, \qquad \Pi_s \Pi_r = \delta_{sr} \Pi_s. \tag{3}$$

A task selects which of these directions may carry bulk motion. For the selected sectors $\mathcal{I}_{\mathfrak{D}}$, define

$$P_{\mathfrak{D}} = \sum_{s \in \mathcal{I}_{\mathfrak{D}}} \Pi_s, \qquad \mathcal{D} = \operatorname{im} P_{\mathfrak{D}}, \qquad \mathfrak{D} = (\langle \cdot, \cdot \rangle_\star, P_{\mathfrak{D}}, B_\partial). \tag{4}$$

The projector restricts bulk directions, while B_∂ specifies boundary exchange. Unrestricted blocks use $P_{\mathfrak{D}} = I$; excluding harmonic flow uses $I - \Pi_{\mathrm{har}}$. Multiple state blocks use the corresponding cochain degrees and a block declaration (Appendix A.2). The metric, projector, and port are fixed independently of neural parameters.

This restriction concerns motion, not the information available to the model. In a closed system with fixed $P_{\mathfrak{D}}$, admissible trajectories stay on the affine space $z_0 + \mathcal{D}$; coordinates outside $\mathcal{D}$ can still affect constitutive laws. A boundary port may exchange components outside these bulk directions.

3.3 Compatibility of geometry and thermodynamics

The declaration determines where the state can move. GENERIC assigns reversible and dissipative roles to that motion through energy $E(z)$, entropy $S(z)$, and operators L, M:

$$\dot{z} = f(z; \mathfrak{D}, u) = L\nabla E + M\nabla S + B_\partial u. \tag{5}$$

Gradients are taken in the metric of Eq. (2). The algebraic conditions

$$L^\dagger = -L, \qquad M = M^\dagger \succeq 0, \qquad L\nabla S = 0, \qquad M\nabla E = 0 \tag{6}$$

make the reversible part entropy-neutral and the dissipative part energy-neutral. With no port input they imply $\dot{E} = 0$ and $\dot{S} = \langle \nabla S, M\nabla S \rangle \geq 0$.

Definition 1 (Mesh–GENERIC Compatibility)**.** *Given the declaration in Eq.* (4)*, a continuous-time generator is Mesh–GENERIC compatible when its bulk action lies in $\mathcal{D}$, its reversible and dissipative operators satisfy Eq.* (6) *under the declared metric, and external exchange enters through $B_\partial u$.*

Section 4 constructs f_θ for Eq. (1) by enforcing both requirements on the same learned operators. MGC is an algebraic condition; it does not impose the Jacobi identity on L or identify a unique physical decomposition into E, S, L, M.

4 MAGNS: Neural dynamical assembly for MGC

The declaration in Eq. (4) is prescribed; the constitutive laws and operators in Eq. (5) are learned. MAGNS connects the two by constructing operators whose images lie in $\mathcal{D}$ and whose action respects the energy–entropy conditions. The construction proceeds from scalar functionals to admissible channels, then to the generator used for rollout.

4.1 Learning constitutive quantities

A geometry-aware shared graph processor maps the observed mesh state to a typed structural state $\chi_\theta(z; \mathfrak{D})$. Independent readouts provide scalar functionals and operator coefficients, rather than a next-state increment. The functional heads have the form

$$E_\theta(z) = E_{\mathrm{decl}}(z) + e_\theta(\chi_\theta(z; \mathfrak{D})), \qquad S_\theta(z) = S_{\mathrm{decl}}(z) + s_\theta(\chi_\theta(z; \mathfrak{D})). \tag{7}$$

The declared terms retain known constitutive content; a conservative wave, for example, keeps analytic kinetic energy while learning its potential. A constant entropy needs no readout. Differentiating these scalars in the declared metric supplies the gradients in Eq. (5), including state dependence through the representation and readouts. Restricting bulk motion to $\mathcal{D}$ does not restrict the coordinates on which the functionals depend.

Operator readouts supply skew blocks $K_a = \widetilde{K}_a - \widetilde{K}_a^\dagger$ and positive-semidefinite blocks $R_b = F_b F_b^\dagger$. These coefficients control interactions between channels. The channels themselves must combine the geometric and thermodynamic restrictions: reversible channels lie in $\mathcal{D} \cap (\nabla S)^\perp$, and dissipative channels in $\mathcal{D} \cap (\nabla E)^\perp$.

4.2 From declared directions to compatible operators

Let V be the state-direction space with the declared metric, and let U_a be a finite-dimensional coefficient space with a prescribed inner product. Channel lifts $A_{S,a} : U_a \to V$ and $A_{E,b} : U_b \to V$ map coefficients into physical directions. The projector from Eq. (4) and the gradients from Eq. (7) give their defining conditions:

$$\begin{aligned} P_{\mathfrak{D}} A_{S,a} &= A_{S,a}, \qquad A_{S,a}^{\dagger} \nabla S = 0, \\ P_{\mathfrak{D}} A_{E,b} &= A_{E,b}, \qquad A_{E,b}^{\dagger} \nabla E = 0. \end{aligned} \tag{8}$$

The first condition places every channel image in the chosen mesh directions; the second makes it orthogonal to the relevant thermodynamic gradient. Orthogonality uses Eq. (2), so the same metric controls geometric admissibility and operator degeneracy.

A gradient-orthogonal completion enforces both conditions. Split the declared directions into orthogonal blocks $\mathcal{D}_x \oplus \mathcal{D}_y$. For $F = S$ or E, write $g = \nabla F = (g_x, g_y)$ and let $C_a : U_a \to V_x$ be a chosen mesh direction map. Project it into $\mathcal{D}_x$ and add a component in $\mathcal{D}_y$ that cancels its pairing with g:

$$A_{F,a} u = \begin{pmatrix} P_x C_a u \\ -r_a \langle g_x, P_x C_a u \rangle \end{pmatrix}, \qquad \langle g, A_{F,a} u \rangle = 0. \tag{9}$$

Here P_x projects onto $\mathcal{D}_x$ and $r_a \in \mathcal{D}_y$ satisfies $\langle g_y, r_a \rangle = 1$. One choice is $r_a = Q_a g_y / \|Q_a g_y\|^2$, where Q_a is a declared orthogonal projector into $\mathcal{D}_y$ and $Q_a g_y \neq 0$. Channels already orthogonal to g need no completion; domain conditions and specializations are given in Appendix A.5.

The completed channels turn coefficient blocks into mesh operators:

$$L = \sum_a A_{S,a} K_a A_{S,a}^{\dagger}, \quad K_a^{\dagger} = -K_a, \tag{10}$$

$$M = \sum_b A_{E,b} R_b A_{E,b}^{\dagger}, \quad R_b = R_b^{\dagger} \succeq 0. \tag{11}$$

Each term has its image in $\mathcal{D}$. The coefficient blocks impose skew-adjointness or positive semidefiniteness, while the channel orthogonality gives $L\nabla S = 0$ and $M\nabla E = 0$. These properties survive summation even when channel images overlap. Substituting the learned quantities into Eq. (5) produces

$$f_\theta(z; \mathfrak{D}, u) = L_\theta \nabla E_\theta + M_\theta \nabla S_\theta + B_\partial u. \tag{12}$$

The boundary port is the one prescribed in Section 3.2, not a learned correction to the bulk dynamics.

4.3 Compatibility and expressive freedom

The construction makes compatibility a property of the operator family rather than an outcome of trajectory fitting.

Theorem 1 (Sufficient MGC certificate for channel assembly)**.** *Fix a mesh declaration $\mathfrak{D}$. On a state domain where E, S are continuously differentiable and all active channel lifts are defined, assume Eq.* (8) *and the coefficient conditions in Eqs.* (10)–(11)*. Then the assembled generator in Eq.* (12) *is MGC-compatible: its bulk action lies in $\mathcal{D}$, $L^{\dagger} = -L$, $M = M^{\dagger} \succeq 0$, $L\nabla S = 0$, and $M\nabla E = 0$. Along any differentiable solution remaining in this domain,*

$$\dot{E} = \langle \nabla E, B_\partial u \rangle, \qquad \dot{S} = \langle \nabla S, M \nabla S \rangle + \langle \nabla S, B_\partial u \rangle. \tag{13}$$

In particular, a closed system has $\dot{E} = 0$ and $\dot{S} \geq 0$.

The proof applies the metric adjoint to the channel compositions (Appendix A.3). It holds for state-dependent functionals, coefficients, and lifts on their common domain. A complementary question is whether the constraints themselves limit the construction beyond MGC. Fix a state and write $V_E = \mathcal{D} \cap (\nabla E)^{\perp}$ and $V_S = \mathcal{D} \cap (\nabla S)^{\perp}$.

Proposition 1 (Pointwise operator-class characterization)**.** *Let common lifts W and J have images V_E and V_S, respectively. With unrestricted PSD and skew-adjoint coefficient blocks, the families*

$$M = WRW^{\dagger}, \quad R = R^{\dagger} \succeq 0, \qquad L = JKJ^{\dagger}, \quad K^{\dagger} = -K \tag{14}$$

contain exactly the operators with range in $\mathcal{D}$ satisfying Eq. (6) *at that state. If $d_E = \dim V_E$ and $d_S = \dim V_S$, the dissipative cone has dimension $d_E(d_E+1)/2$ and the reversible space has dimension $d_S(d_S-1)/2$.*

For the converse, right inverses of the full-range lifts pull each compatible operator back to its coefficient space, preserving positivity or skew-adjointness (Appendix A.4). The declaration therefore determines the compatible operator class; channel rank, supports, and coefficient couplings select the family represented by a particular model. Changing $P_{\mathfrak{D}}$ changes this class without changing neural weights, a distinction tested by the Darcy intervention in Section 5.4.

4.4 From the generator to a learned rollout

The assembled f_θ closes the update in Eq. (1). Training differentiates supervised rollout losses through this update, while assembly enforces compatibility at every generator evaluation. Task declarations select the active physics: wave and acoustics have $M = 0$ and constant S, Darcy has $L = 0$, and Beam3D uses both operators. The acoustic declaration also includes a radiating boundary port.

The choice of $\Phi_{\Delta t}$ determines how the continuous balances in Eq. (13) are realized at finite step size. A discrete-gradient method additionally imposes (Gonzalez, 1996; McLachlan et al., 1999)

$$\langle \nabla_d E, z_{t+1} - z_t \rangle = E(z_{t+1}) - E(z_t). \tag{15}$$

Together with compatible two-point assembly and a solved step equation, this gives finite-step balance (Appendix B). Explicit stepping and acoustic KDK instead introduce their own integration defects. The wave comparison below holds the learned generator fixed to isolate this numerical effect from operator compatibility.

5 Experiments

MGC constrains the dynamics that produce a prediction, not just the prediction error. We test whether this construction supports accurate long rollouts and whether its structural choices have observable effects. Conservative waves separate the learned generator from its numerical realization; acoustics and Beam3D examine open and coupled dynamics. We then use controlled ablations to distinguish the effects of operator assembly and admissible flow directions. Detailed protocols and structural measurements are in Appendix D and Appendix C.

5.1 Separating structure from time integration: conservative waves

A closed wave system makes the distinction at the end of Section 4.4 directly measurable. We predict 200 steps on a periodic 32×32 mesh using five seed groups from the MeshFT data protocol (Noguchi & Kawahara, 2026). The same learned weights are evaluated with Heun stepping and a discrete-gradient integrator. Table 1 compares trajectory error and energy drift, with published baseline results providing the reference scale.

Table 1: Analytic wave, five-seed means (VALID, $T = 200$). Baseline values are from Noguchi & Kawahara (2026). TSMSE measures trajectory error; energy drift measures the relative change from the initial energy.

Model	Integrator	One-step MSE	TSMSE	Energy drift
MAGNS	Heun (explicit RK2)	1.038×10^{-9}	3.019×10^{-5}	4.121×10^{-4}
MAGNS	discrete gradient	1.422×10^{-9}	4.042×10^{-5}	2.105×10^{-5}
MeshFT	kick–drift–kick (KDK)	1.3×10^{-9}	9.6×10^{-5}	1.3×10^{-4}
MGN	second-order stepping	1.6×10^{-7}	1.3×10^{-1}	25.9
MGN-HP	second-order stepping	5.7×10^{-4}	6.1×10^{-1}	16.0
HNN	second-order stepping	3.5×10^{-8}	3.0×10^{-3}	1.0×10^{-2}

MAGNS attains trajectory MSE 3.019×10^{-5} with explicit stepping, below the published structured baselines in this comparison. Switching only the integrator reduces energy drift by 94.9%, with trajectory MSE remaining of order 10^{-5}. The generator is unchanged: the improvement comes from the finite-step realization of the same compatible dynamics, rather than learning a different energy or operator. This separates the role of the assembly theorem from that of the discrete-gradient identity in Eq. (15).

5.2 PREDICTION WITH BOUNDARY EXCHANGE: ACOUSTIC SCATTERING

Acoustic scattering tests whether accurate propagation is retained when waves exchange energy through the boundary. We compare MAGNS with MeshFT (Noguchi & Kawahara, 2026), IGNS (Hoang et al., 2026), and MGN (Pfaff et al., 2021) on The Well's 256×256 scattering task (Ohana et al., 2024), using 200 common windows per split and $H = 80$. MAGNS, MeshFT, and MGN use $H = 1$ pretraining followed by $H = 4$ training; IGNS follows its native sequence-training scheme. Budgets, numerical settings, and the $H = 8$ comparison are in Appendix D.2. The common-chart energy is a quadratic diagnostic of the observed state, whose p channel is a scaled frame difference, not a calibrated measure of physical acoustic-energy fidelity.

Table 2: Acoustic scattering, $H = 80$, on 200 windows per split. One-step errors and chart-energy gaps are across-window medians; rollout MSE is reported as both median and mean. The chart-energy gap is $|E_{\text{chart}}(\hat{z}_{80}) - E_{\text{chart}}(z_{80})|/|E_{\text{chart}}(z_{80})|$. Lower is better; the best value in each split and column is bold.

Model	Split	One-step MSE Median	Rollout MSE Median	Rollout MSE Mean	Chart-energy gap Median
MAGNS	VALID	$\mathbf{7.77\times10^{-5}}$	$\mathbf{1.84\times10^{-2}}$	$\mathbf{2.04\times10^{-2}}$	**1.49**
MeshFT	VALID	1.24×10^{-3}	1.92×10^{-2}	2.29×10^{-2}	3.51
IGNS	VALID	5.72×10^{-3}	4.43×10^{-2}	5.07×10^{-2}	3.92×10^{3}
MGN	VALID	6.34×10^{-3}	9.58×10^{-2}	1.17×10^{-1}	2.11
MAGNS	TEST	$\mathbf{7.46\times10^{-5}}$	2.01×10^{-2}	$\mathbf{2.19\times10^{-2}}$	**1.28**
MeshFT	TEST	1.18×10^{-3}	$\mathbf{2.00\times10^{-2}}$	2.34×10^{-2}	3.27
IGNS	TEST	5.61×10^{-3}	4.69×10^{-2}	5.19×10^{-2}	3.83×10^{3}
MGN	TEST	6.43×10^{-3}	1.04×10^{-1}	1.13×10^{-1}	2.37

MAGNS reduces mean rollout MSE by 11.2% on VALID and 6.3% on TEST relative to MeshFT, while matching its TEST median accuracy (Table 2). The median final common-chart energy gap is 1.28 on TEST, a 60.9% smaller diagnostic discrepancy than MeshFT. Figure 2 shows how prediction and energy differences develop over the rollout: MAGNS and MeshFT retain substantially lower field errors than IGNS and MGN.

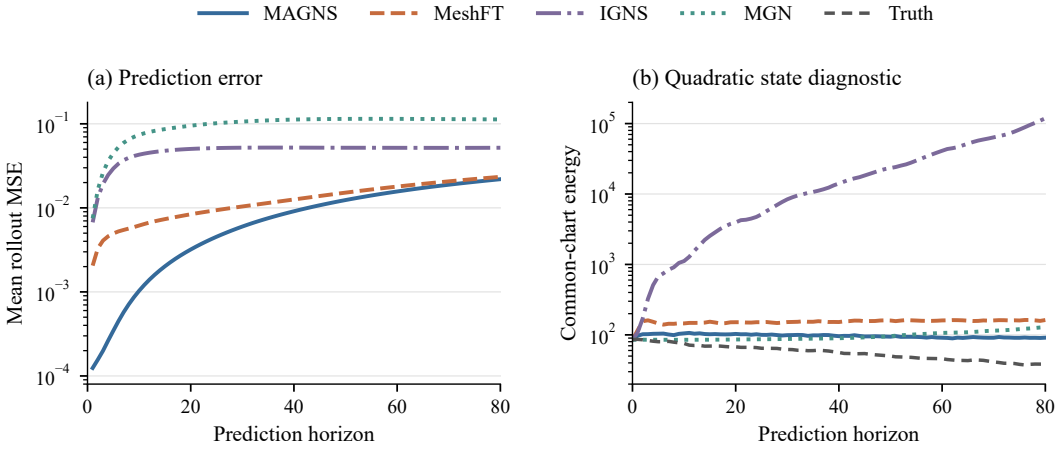

Figure 2: Acoustic TEST, 200 windows. Left: mean rollout MSE through each horizon. Right: median common-chart energy for all models and truth. Both vertical axes are logarithmic. The quadratic diagnostic is evaluated in the observed state coordinates.

The adapted IGNS accumulates decoded latent increments in the observed coordinates. Inverse-mass weighting amplifies errors in its small difference-derived p channel. The resulting energy gap cannot be attributed to the original IGNS formulation independently of these adaptations. MGC instead imposes compatibility on the mesh-state generator and its boundary exchange; field accuracy and the learned-energy port balance complement the common-chart diagnostic (Appendix E.3).

5.3 COUPLED REVERSIBLE–DISSIPATIVE DYNAMICS: BEAM3D

The forced viscoelastic beam couples deformation, velocity, and stress (Hernández et al., 2024), activating both operators in Eq. (5). This tests the construction beyond the conservative-wave setting. We compare MAGNS with MGN, IGNS, and Local-TIGNN (Tierz et al., 2025) over $H = 20$ rollouts. Table 3 reports three-seed results; training and selection details are in Appendix D.

Table 3: Beam3D, $H = 20$ closed-loop normalized MSE (VALID, three-seed mean $\pm$ sample s.d.). Best in bold.

Model	VALID
MAGNS	$\mathbf{0.03533 \pm 0.00537}$
IGNS	0.04518 ± 0.00420
Local-TIGNN	0.04679 ± 0.00466
MGN	0.06662 ± 0.00748

MAGNS has the lowest mean error on both VALID and TEST (Appendix D). Field plots show the associated differences in deformation and stress (Appendix F). Thus compatible assembly supports predictive accuracy in coupled reversible–dissipative dynamics, not only in the canonical wave specialization. The following ablation asks how much of this accuracy comes from assembly rather than network capacity.

5.4 ABLATIONS OF OPERATOR ASSEMBLY AND ADMISSIBLE DIRECTIONS

Operator assembly in Beam3D. We keep the training budget fixed and compare full assembly against a direct increment decoder, equal-rank unconstrained operators, and an equal-capacity control. These comparisons separate the effect of constructing the dynamics from that of simply allocating more parameters. Table 4 includes both the mean over all trajectories and the difficult trajectory that contributes most to the error.

Table 4: Beam3D assembly ablation: VALID $H = 20$ normalized MSE, seed 0. The mean includes all five trajectories; the hard case is also shown separately.

Configuration	Mean	Hard case
Full assembly	$\mathbf{0.03963}$	$\mathbf{0.1549}$
No assembly	0.1795	0.4931
Free operators	0.04022	0.1779
Equal-capacity control	0.07877	0.2463

Removing assembly raises mean error 4.53-fold. Matching parameter count recovers only part of the loss, leaving error nearly twice that of the full model. Full assembly and free operators have similar mean errors; in this seed-0 comparison, full assembly lowers error on the highest-error trajectory by 12.9%. The capacity-matched control shows that network size alone does not explain the gain over no assembly. Per-trajectory results are in Appendix D.

Hodge-space control in Darcy flow. The operator-class result in Section 4.3 predicts a different effect: changing the declaration changes which flows the model can realize, even with fixed neural weights. We test this on periodic heterogeneous Darcy flow. Case A contains physical harmonic circulation, whereas its reflection-symmetric counterpart B has none. For the same full-space-trained model, we switch the projector in Eq. (4) from I to $I - \Pi_{\mathrm{har}}$, retaining exact and coexact directions while excluding harmonic motion. Each switch rebuilds the compatible lifts and reassembles M as in Theorem 1; across the crossed evaluations, $\|M\, dE\|_\infty \leq 1.77 \times 10^{-15}$.

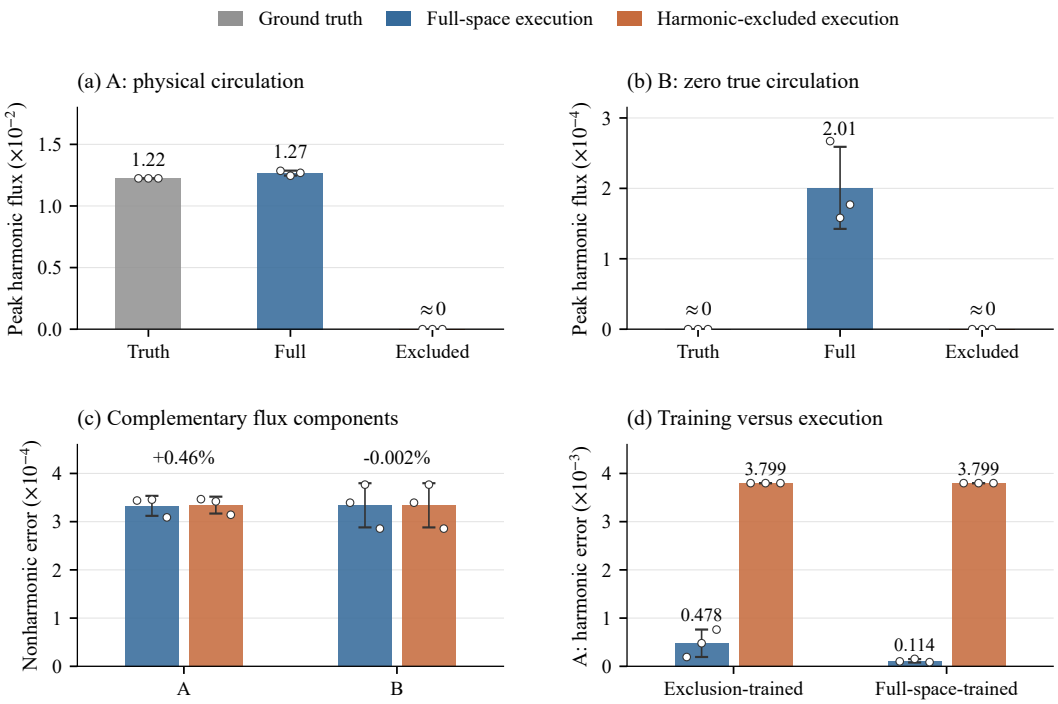

Figure 3: Fixed-weight Darcy ablation over 64-step TEST rollouts. (a,b) Peak harmonic-flux magnitude, $\max_t \|\Pi_{\mathrm{H}} q_t\|$, for truth, full-space execution, and harmonic-excluded execution. (c) Time-averaged nonharmonic flux error. (d) Harmonic error in A under both execution declarations, for exclusion-trained and full-space-trained models. Bars average three seeds; circles show seed means and error bars their sample standard deviations. Each seed averages eight trajectories.

Figure 3 shows that the switch removes harmonic circulation in both cases. In A it suppresses a physical component; in B it removes spurious circulation to numerical precision. Errors in the remaining flux components change by less than 0.5%. The declaration therefore controls a specific global mode rather than indiscriminately suppressing the learned flow.

The reverse intervention distinguishes this control from training. Allowing harmonic directions in an exclusion-trained A model reduces harmonic error from 3.80×10^{-3} to 4.78×10^{-4} without retraining. Training with those directions available lowers it further to 1.14×10^{-4}. Execution determines whether circulation can occur; training determines how accurately it is reproduced. Full crossed comparisons and energy diagnostics are in Appendix G.

6 Conclusion and Limitations

Mesh–GENERIC Compatibility places admissible flow directions and thermodynamic structure within the same learned dynamics. MAGNS realizes this condition through neural dynamical assembly. Competitive acoustic rollouts, the Beam3D assembly ablation, and fixed-weight Darcy interventions show how predictive accuracy and control over physical modes can be studied together.

The value of this control depends on the mesh declaration. Excluding a physically active direction removes it from the learned bulk operators, so optimization cannot repair a mismatched declaration; Darcy's circulation switch illustrates this trade-off. Transfer across discretizations therefore requires consistent choices of metric, direction space, and boundary ports. Within a declaration, finite-rank channels trade expressivity for computational cost: the full-range characterization describes a larger operator class than a fixed implementation represents. Compatibility also governs the continuous generator rather than its numerical trajectory. Integration error can accumulate even when the algebraic identities hold, making the numerical scheme part of the accuracy–structure trade-off in long rollouts.

AI USE STATEMENT

The theoretical formalization and the experimental protocol were specified by the authors prior to this manuscript. We did not use generative AI to develop the conceptual framework, to formulate the mathematical claims, to design the experiments, to generate synthetic datasets, or to clean training data. We used large language model assistants for required-disclosure tasks that did apply: writing proofs of already-stated identities, implementing and editing code, translating drafts. We additionally used them for recommended tasks: drafting and editing the manuscript (including bibliography), searching and organizing literature, minor layout edits to author-supplied figures, and code comments. All AI-assisted text, proofs, and code were checked against the authors' specification, the implementation, and the experiment logs. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work studies structure-preserving learning on public or synthetic simulation data (analytic waves, transient Darcy flow, The Well acoustic scattering, and the TIGNN Beam3D suite). It does not involve human subjects, private data, or a dual-use application beyond standard scientific computing. We have read and adhere to the ICLR Code of Ethics.

REPRODUCIBILITY STATEMENT

Theorem 1 and the pointwise operator-class characterization are proved in Appendix A. Integrator identities and the update algorithm are in Appendix B. Structural residuals and port ledgers are in Appendix C. Training budgets, seeds, splits, hyperparameters, and table provenance are in Appendix D. Acoustic rollout curves are in Appendix E. Beam3D field figures are in Appendix F. The synthetic Darcy generation protocol and full controls are in Appendix G. Datasets follow the cited MeshFT, The Well, and TIGNN releases. Code is linked from the title footnote.

REFERENCES

Peter W. Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, and Koray Kavukcuoglu. Interaction networks for learning about objects, relations and physics. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://proceedings.neurips.cc/paper_files/paper/2016/hash/3147da8ab4a0437c15ef51a5cc7f2dc4-Abstract.html.

Carlos Bermejo, Alberto Badías, David González, and Elías Cueto. MeshGraphNets informed locally by thermodynamics for the simulation of flows around arbitrarily shaped objects. *Advanced Modeling and Simulation in Engineering Sciences*, 12(1):27, 2025. doi: 10.1186/s40323-025-00311-8.

Nacime Bouziani and Nicolas Boullé. Structure-preserving operator learning. arXiv preprint arXiv:2410.01065, 2024. URL https://arxiv.org/abs/2410.01065.

Johannes Brandstetter, Daniel Worrall, and Max Welling. Message passing neural PDE solvers. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=vSix3HPYKSU.

Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. In *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*, 2020. URL https://arxiv.org/abs/2003.04630.

Shaan A. Desai, Marios Mattheakis, David Sondak, Pavlos Protopapas, and Stephen J. Roberts. Port-Hamiltonian neural networks for learning explicit time-dependent dynamical systems. *Physical Review E*, 104(3):034312, 2021. doi: 10.1103/PhysRevE.104.034312.

Paul Garnier, Vincent Lannelongue, Jonathan Viquerat, and Elie Hachem. MeshMask: Physics-based simulations with masked graph neural networks. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=bFHR8hNk4I.

Oscar Gonzalez. Time integration and discrete Hamiltonian systems. *Journal of Nonlinear Science*, 6(5):449–467, 1996. doi: 10.1007/BF02440162.

Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/26cd8ecadce0d4efd6cc8a8725cbd1f8-Abstract.html.

Miroslav Grmela and Hans Christian Öttinger. Dynamics and thermodynamics of complex fluids. i. development of a general formalism. *Physical Review E*, 56(6):6620–6632, 1997. doi: 10.1103/PhysRevE.56.6620.

Anthony Gruber, Kookjin Lee, Haksoo Lim, Noseong Park, and Nathaniel Trask. Efficiently parameterized neural metriplectic systems. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=uL1H29dM0c.

Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2 edition, 2006. doi: 10.1007/3-540-30666-8.

Quercus Hernández, Alberto Badías, David González, Francisco Chinesta, and Elías Cueto. Structure-preserving neural networks. *Journal of Computational Physics*, 426:109950, 2021. doi: 10.1016/j.jcp.2020.109950.

Quercus Hernández, Alberto Badías, Francisco Chinesta, and Elías Cueto. Thermodynamics-informed neural networks for physically realistic mixed reality. *Computer Methods in Applied Mechanics and Engineering*, 407:115912, 2023. doi: 10.1016/j.cma.2023.115912.

Quercus Hernández, Alberto Badías, Francisco Chinesta, and Elías Cueto. Thermodynamics-informed graph neural networks. *IEEE Transactions on Artificial Intelligence*, 5(3):967–976, 2024. doi: 10.1109/TAI.2022.3179681.

Anil Nirmal Hirani. *Discrete Exterior Calculus*. PhD thesis, California Institute of Technology, Pasadena, CA, 2003. URL https://thesis.caltech.edu/1885/.

Tai Hoang, Alessandro Trenta, Alessio Gravina, Niklas Freymuth, Philipp Becker, Davide Bacciu, and Gerhard Neumann. Improving long-range interactions in graph neural simulators via Hamiltonian dynamics. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=x66u6TEDUw.

Peiyan Hu, Yue Wang, and Zhi-Ming Ma. Better neural PDE solvers through data-free mesh movers. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=hj9ZuNimRl.

Lek-Heng Lim. Hodge laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020. doi: 10.1137/18M1223101.

Phillip Lippe, Bastiaan S. Veeling, Paris Perdikaris, Richard E. Turner, and Johannes Brandstetter. PDE-Refiner: Achieving accurate long rollouts with neural PDE solvers. In *Advances in Neural Information Processing Systems*, volume 36, pp. 67398–67433, 2023. doi: 10.52202/075280-2946. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/d529b943af3dba734f8a7d49efcb6d09-Abstract-Conference.html.

Robert I. McLachlan, G. R. W. Quispel, and Nicolas Robidoux. Geometric integration using discrete gradients. *Philosophical Transactions of the Royal Society A*, 357(1754):1021–1045, 1999. doi: 10.1098/rsta.1999.0363.

Satoshi Noguchi and Yoshinobu Kawahara. Mesh field theory: Port-Hamiltonian formulation of mesh-based physics. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=MgIvtHk5wm.

Ruben Ohana, Michael McCabe, Lucas Meyer, Rudy Morel, Fruzsina J. Agocs, Miguel Beneitez, Marsha Berger, Blakesley Burkhart, Stuart B. Dalziel, Drummond B. Fielding, Daniel Fortunato, Jared A. Goldberg, Keiya Hirashima, Yan-Fei Jiang, Rich R. Kerswell, Suryanarayana Maddu, Jonah Miller, Payel Mukhopadhyay, Stefan S. Nixon, Jeff Shen, Romain Watteaux, Bruno Régaldo-Saint Blancard, François Rozet, Liam H. Parker, Miles Cranmer, and Shirley Ho. The well: a large-scale collection of diverse physics simulations for machine learning. In *Advances in Neural Information Processing Systems*, volume 37, pp. 44989–45037, 2024. doi: 10.52202/079017-1430. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/4f9a5acd91ac76569f2fe291b1f4772b-Abstract-Datasets_and_Benchmarks_Track.html. Datasets and Benchmarks Track.

Hans Christian Öttinger. Nonequilibrium thermodynamics for open systems. *Physical Review E*, 73 (3):036126, 2006. doi: 10.1103/PhysRevE.73.036126.

Hans Christian Öttinger. GENERIC integrators: Structure preserving time integration for thermodynamic systems. *Journal of Non-Equilibrium Thermodynamics*, 43(2):89–100, 2018. doi: 10.1515/jnet-2017-0034.

Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W. Battaglia. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=roNqYL0_XP.

T. Mitchell Roddenberry and Santiago Segarra. HodgeNet: Graph neural networks for edge data. In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 220–224. IEEE, 2019. doi: 10.1109/IEEECONF44664.2019.9049000.

Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter W. Battaglia. Learning to simulate complex physics with graph networks. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 8459–8468. PMLR, 2020. URL https://proceedings.mlr.press/v119/sanchez-gonzalez20a.html.

Xiaocheng Shang and Hans Christian Öttinger. Structure-preserving integrators for dissipative systems based on reversible–irreversible splitting. *Proceedings of the Royal Society A*, 476(2234): 20190446, 2020. doi: 10.1098/rspa.2019.0446.

Jason Sulskis and Sathya Ravi. GENERIC-FNO: Embedding energy conservation and entropy production into Fourier neural operators. arXiv preprint arXiv:2606.08343, 2026. URL https://arxiv.org/abs/2606.08343.

Alicia Tierz, Icíar Alfaro, David González, Francisco Chinesta, and Elías Cueto. Graph neural networks informed locally by thermodynamics. *Engineering Applications of Artificial Intelligence*, 144:110108, 2025. doi: 10.1016/j.engappai.2025.110108.

Tobias Würth, Niklas Freymuth, Gerhard Neumann, and Luise Kärger. Diffusion-based hierarchical graph neural networks for simulating nonlinear solid mechanics. In *Advances in Neural Information Processing Systems*, volume 38, pp. 144338–144367, 2025. doi: 10.52202/085713-4336. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/bcd3c132bb19544815c5e660a2b044bd-Abstract-Conference.html.

Hao Xu and Jia Pan. HHD-GP: Incorporating helmholtz–Hodge decomposition into gaussian processes for learning dynamical systems. In *Advances in Neural Information Processing Systems*, volume 37, pp. 67282–67318, 2024. doi: 10.52202/079017-2148. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/7c35478c8829e5b8baa0e69f3240994a-Abstract-Conference.html.

Maosheng Yang, Viacheslav Borovitskiy, and Elvin Isufi. Hodge-compositional edge gaussian processes. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 3754–3762. PMLR, 2024. URL https://proceedings.mlr.press/v238/yang24e.html.

Youn-Yeol Yu, Jeongwhan Choi, Jaehyeon Park, Kookjin Lee, and Noseong Park. PIORF: Physics-informed Ollivier-Ricci flow for long-range interactions in mesh graph neural networks. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=qkBBHixPow.

Zhen Zhang, Yeonjong Shin, and George Em Karniadakis. GFINNs: GENERIC formalism informed neural networks for deterministic and stochastic dynamical systems. *Philosophical Transactions of the Royal Society A*, 380(2229):20210207, 2022. doi: 10.1098/rsta.2021.0207.

Dongzhe Zheng, Tao Zhong, and Christine Allen-Blanchette. Topology-preserving neural operator learning via Hodge decomposition. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=sLVaNS2DS3.

Yaofeng Desmond Zhong, Biswadip Dey, and Amit Chakraborty. Symplectic ODE-net: Learning Hamiltonian dynamics with control. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=ryxmb1rKDS.

A ADMISSIBLE OPERATORS: CONSTRUCTION AND EXPRESSIVITY

Channel assembly enforces the GENERIC identities (Grmela & Öttinger, 1997) on the declared mesh directions. We prove this compatibility certificate, characterize the pointwise operator class under full-range lifts, and give supported completion and thermal routing as concrete realizations.

Throughout, $\dagger$ denotes the declared metric adjoint, and $P = P_{\mathfrak{D}}$ is fixed independently of learned parameters. Lifts may depend on the state and the full functional derivatives; each result applies on the domain of its lift construction.

A.1 INNER PRODUCT, ADJOINTS, AND PAIRING CALCULUS

Let $C^k(\mathcal{K})$ be the space of k-cochains on the oriented mesh $\mathcal{K}$, of dimension n_k. The Hodge star $\star_k$ is represented by a symmetric positive-definite matrix $W_k \in \mathbb{R}^{n_k \times n_k}$. The declared inner product on C^k is

$$\langle u, v\rangle_{W_k} = u^\top W_k v. \tag{16}$$

The concatenated state z lives in a direct sum of such cochain spaces (and, when present, a thermal coordinate). The state inner product is the corresponding direct sum; we write $\langle\cdot,\cdot\rangle$ for that product and omit the subscript when the space is clear. The metric adjoint of a linear map $A : C^k \to C^\ell$ is the unique $A^\dagger$ satisfying $\langle Au, v\rangle_{W_\ell} = \langle u, A^\dagger v\rangle_{W_k}$ for all u, v, namely

$$A^\dagger = W_k^{-1} A^\top W_\ell. \tag{17}$$

In particular the coboundaries D_0 and D_1 of the main text have adjoints $D_k^\dagger = \star_k^{-1} D_k^\top \star_{k+1}$.

Lemma 1 (Adjoint calculus)**.** *For linear maps of compatible type, $(AB)^\dagger = B^\dagger A^\dagger$, $(A+B)^\dagger = A^\dagger + B^\dagger$, and $(A^\dagger)^\dagger = A$. If $A^\dagger = -A$, then $\langle u, Au\rangle = 0$ for every u. If $A = A^\dagger$ and $\langle u, Au\rangle \geq 0$ for every u, then $A \succeq 0$.*

Proof. The first three identities are the uniqueness of the metric adjoint applied to the defining pairings $\langle ABu, v\rangle = \langle Bu, A^\dagger v\rangle = \langle u, B^\dagger A^\dagger v\rangle$, $\langle (A+B)u, v\rangle = \langle Au, v\rangle + \langle Bu, v\rangle$, and $\langle A^\dagger u, v\rangle = \langle u, Av\rangle$. If $A^\dagger = -A$, then

$$\langle u, Au\rangle = \langle A^\dagger u, u\rangle = \langle -Au, u\rangle = -\langle Au, u\rangle = -\langle u, Au\rangle, \tag{18}$$

so $2\langle u, Au\rangle = 0$. The last claim is the definition of positive semidefiniteness on a real inner-product space. □

The equality $A^\top = A^\dagger$ holds exactly when $W_k A^\top = A^\top W_\ell$, in particular for identity metrics. All operator conditions below use $\dagger$.

Lemma 2 (Orthogonal projector sandwich)**.** *Let Q be self-adjoint and idempotent ($Q^\dagger = Q$ and $Q^2 = Q$), and let D be bounded. If $D^\dagger = -D$, then $(QDQ)^\dagger = -QDQ$. If $D = D^\dagger \succeq 0$, then QDQ is self-adjoint and positive semidefinite.*

Proof. Self-adjointness of Q gives $(QDQ)^\dagger = Q^\dagger D^\dagger Q^\dagger = QD^\dagger Q$. If $D^\dagger = -D$, this equals $-QDQ$. If $D^\dagger = D$, it equals QDQ. For any test vector v, $\langle v, QDQv\rangle = \langle Qv, D\,Qv\rangle$, which is nonnegative whenever $D \succeq 0$. □

A.2 HODGE DECOMPOSITION AND METRIC PROJECTORS

The cochain identity $D_1 D_0 = 0$ gives the orthogonal decomposition used by the declaration.

Lemma 3 (Hodge decomposition)**.** *Assume each W_k is symmetric positive definite and $D_1 D_0 = 0$. Let $\delta_{k+1} = W_k^{-1} D_k^\top W_{k+1}$ and $\mathcal{H}^k = \ker D_k \cap \ker \delta_k$. Then*

$$C^k(\mathcal{K}) = \operatorname{im} D_{k-1} \;\oplus_{W_k}\; \operatorname{im} \delta_{k+1} \;\oplus_{W_k}\; \mathcal{H}^k, \tag{19}$$

an orthogonal direct sum for $\langle\cdot,\cdot\rangle_{W_k}$. In degree $k = 1$, $\delta_1 = D_0^\dagger$, $\delta_2 = D_1^\dagger$, $\Delta_1 = D_0 D_0^\dagger + D_1^\dagger D_1$, and $\ker \Delta_1 = \mathcal{H}^1$.

Proof. Take $u = D_{k-1}a$ and $v = \delta_{k+1}b$. Then

$$\langle u, v\rangle_{W_k} = \langle D_{k-1}a, \delta_{k+1}b\rangle_{W_k} = \langle D_k D_{k-1}a,\, b\rangle_{W_{k+1}} = 0,$$

using $D_k D_{k-1} = 0$ (for $k = 1$, $D_1 D_0 = 0$). Thus the exact and coexact images are orthogonal. The adjoint relation also gives $\operatorname{im} D_{k-1}^{\perp_{W_k}} = \ker \delta_k$ and $\operatorname{im} \delta_{k+1}^{\perp_{W_k}} = \ker D_k$. In finite dimensions,

$$C^k = \operatorname{im} D_{k-1} \oplus_{W_k} \ker \delta_k, \qquad \ker \delta_k = \operatorname{im} \delta_{k+1} \oplus_{W_k} (\ker \delta_k \cap \ker D_k). \tag{20}$$

The last intersection is $\mathcal{H}^k$, which is (19).

For $k = 1$, $\Delta_1 u = 0$ if and only if $\langle u, \Delta_1 u\rangle_{W_1} = 0$. Expanding the quadratic form gives $\|D_0^\dagger u\|_{W_0}^2 + \|D_1 u\|_{W_2}^2 = 0$, hence $D_0^\dagger u = 0$ and $D_1 u = 0$, so $u \in \mathcal{H}^1$. The converse is immediate, and $\ker \Delta_1 = \mathcal{H}^1$. □

Lemma 4 (Metric projector)**.** *Let U be a full-column matrix whose columns span a subspace of C^k, and assume the Gram matrix $U^\top W_k U$ is positive definite. The map*

$$P_U = U(U^\top W_k U)^{-1} U^\top W_k \tag{21}$$

satisfies $P_U^2 = P_U$, $P_U^\dagger = P_U$, $P_U U = U$, and $(I - P_U)U = 0$. Thus P_U is the W_k-orthogonal projector onto $\operatorname{col}(U)$, and $I - P_U$ is the W_k-orthogonal projector onto its complement.

Proof. Let $G = U^\top W_k U$, which is invertible by assumption. Then $P_U U = U G^{-1}(U^\top W_k U) = U G^{-1} G = U$, so $(I - P_U)U = 0$. Next

$$P_U^2 = U G^{-1} U^\top W_k\, U G^{-1} U^\top W_k = U G^{-1} G G^{-1} U^\top W_k = P_U.$$

For self-adjointness, the defining pairing is

$$\langle P_U u, v\rangle_{W_k} = (P_U u)^\top W_k v = u^\top W_k U G^{-1} U^\top W_k v = \langle u, P_U v\rangle_{W_k},$$

using $W_k^\top = W_k$. Hence $P_U^\dagger = P_U$. The range of P_U is $\operatorname{col}(U)$ because $P_U = U(\cdot)$. The range of $I - P_U$ is the W_k-orthogonal complement because P_U is an orthogonal projector. □

The Euclidean projector $P_{\mathrm{E}} = U(U^\top U)^{-1} U^\top$ is metric self-adjoint exactly when $W_k P_{\mathrm{E}} = P_{\mathrm{E}} W_k$. Equation (21) uses an independent basis so its Gram matrix is invertible. The full-space declaration uses $P = I$; excluding a sector with basis U gives $P = I - P_U$.

A.3 Proof of the common channel-lift certificate

Each coefficient space has a fixed inner product; thus $A^\dagger$ is the adjoint from the state space back to its coefficient space. All statements below are pointwise at an admitted state. The finite sums in Eqs. (10)–(11) need not have disjoint supports.

Proof of Theorem 1. Adjoint composition gives

$$(A_{S,a} K_a A_{S,a}^\dagger)^\dagger = A_{S,a} K_a^\dagger A_{S,a}^\dagger = -A_{S,a} K_a A_{S,a}^\dagger.$$

Summing yields $L^\dagger = -L$. The same calculation using $R_b^\dagger = R_b$ gives $M^\dagger = M$, and for any state direction v,

$$\langle v, Mv\rangle = \sum_b \langle A_{E,b}^\dagger v, R_b A_{E,b}^\dagger v\rangle \geq 0.$$

The right-hand lift conditions in Eq. (8) give $L\nabla S = 0$ and $M\nabla E = 0$. The left-hand conditions give $PL = L$ and $PM = M$, so the bulk action lies in $\operatorname{im} P$. These are the algebraic and direction requirements of Definition 1; the only external term is the declared $B_\partial u$.

Along a differentiable solution, the chain rule gives

$$\dot{E} = \langle \nabla E, L\nabla E\rangle + \langle \nabla E, M\nabla S\rangle + \langle \nabla E, B_\partial u\rangle, \tag{22}$$

$$\dot{S} = \langle \nabla S, L\nabla E\rangle + \langle \nabla S, M\nabla S\rangle + \langle \nabla S, B_\partial u\rangle. \tag{23}$$

The first term in $\dot{E}$ is zero by skew-adjointness; its second term is $\langle M\nabla E, \nabla S\rangle = 0$. The reversible entropy term is $-\langle L\nabla S, \nabla E\rangle = 0$. The remaining terms prove Eq. (13), with nonnegative bulk entropy production. These instantaneous pairings use the full gradients of E, S and the current values of A, K, R. The trajectory condition in the theorem keeps every evaluation within the lift domain. □

Corollary 1 (Purely reversible and purely dissipative specializations)**.** *Under the same hypotheses, an empty dissipative family gives $M = 0$, and a closed system has $\dot{E} = \dot{S} = 0$. If S is constant, the reversible annihilation condition is automatic. An empty reversible family gives $L = 0$, and a closed system has $\dot{E} = 0$ and $\dot{S} \geq 0$. The reversible specialization needs no thermal coordinate.*

Proof. Set the corresponding channel sum to zero in Eqs. (22)–(23). □

Coordinate covectors versus metric gradients. Let the state inner product be represented by $G \succ 0$, and use Euclidean coordinates on each coefficient space. If dF is the column of ordinary coordinate derivatives, then $\nabla_G F = G^{-1}dF$ and $A^\dagger = A^\top G$. Consequently the implementation's covector-to-tangent maps are

$$\mathcal{L} = \sum_a A_{S,a} K_a A_{S,a}^\top, \qquad \mathcal{M} = \sum_b A_{E,b} R_b A_{E,b}^\top, \qquad L = \mathcal{L}G, \quad M = \mathcal{M}G. \tag{24}$$

Hence $L\nabla_G E = \mathcal{L}dE$ and $M\nabla_G S = \mathcal{M}dS$. The identities $A_S^\top dS = 0$ and $A_E^\top dE = 0$ express the same metric annihilation conditions; multiplication by G converts the covector maps to tangent operators.

A.4 Pointwise operator expressivity

Fix a state, the Hodge metric, the declaration, and the gradients $g_E = \nabla E$ and $g_S = \nabla S$. The spaces available to the two operator roles are

$$V_E = \mathcal{D} \cap g_E^\perp, \qquad V_S = \mathcal{D} \cap g_S^\perp. \tag{25}$$

The following result identifies when channel assembly fills the compatible operator class rather than a template-restricted subset.

Proposition 2 (Pointwise operator-class expressivity)**.** *Let U_E, U_S be finite-dimensional inner-product spaces, and let $W : U_E \to V$ and $J : U_S \to V$ be lifts with $\operatorname{im} W = V_E$ and $\operatorname{im} J = V_S$. If their coefficient families contain every self-adjoint PSD R on U_E and every skew-adjoint K on U_S, then*

$$\{WRW^\dagger : R = R^\dagger \succeq 0\} = \{M = M^\dagger \succeq 0 : PM = M,\ Mg_E = 0\}, \tag{26}$$

$$\{JKJ^\dagger : K^\dagger = -K\} = \{L : L^\dagger = -L,\ PL = L,\ Lg_S = 0\}. \tag{27}$$

Writing $d_F = \dim V_F$, the dissipative cone has dimension $d_E(d_E+1)/2$ and maximal rank d_E; the reversible space has dimension $d_S(d_S-1)/2$ and maximal rank $2\lfloor d_S/2 \rfloor$.

Proof. Since their images lie in $\mathcal{D}$ and are orthogonal to the target gradients, W, J satisfy the lift conditions, which give the forward inclusions. Conversely, adjointness and degeneracy make the image of any operator on the right orthogonal to its target gradient. The range restriction places that image in V_E or V_S. Adjointness also makes the operator vanish on the orthogonal complement of that space. For the orthogonal projectors Π_E, Π_S, therefore, $M = \Pi_E M \Pi_E$ and $L = \Pi_S L \Pi_S$.

Surjectivity gives right inverses $T_E : V_E \to U_E$ and $T_S : V_S \to U_S$. Extend them by zero on the orthogonal complements, so $WT_E = \Pi_E$ and $JT_S = \Pi_S$. Set

$$R = T_E M T_E^\dagger, \qquad K = T_S L T_S^\dagger.$$

Adjoint composition and the quadratic pairing give $R = R^\dagger \succeq 0$ and $K^\dagger = -K$. Hence $WRW^\dagger = \Pi_E M \Pi_E = M$ and $JKJ^\dagger = \Pi_S L \Pi_S = L$.

In orthonormal bases of V_E, V_S, these classes are the full PSD cone and skew-symmetric matrix space. The cone has nonempty interior in the $d_E(d_E+1)/2$-dimensional symmetric space, and the identity on V_E attains rank d_E. A skew matrix has one free entry per index pair and even rank; independent two-dimensional skew blocks attain $2\lfloor d_S/2 \rfloor$. □

Spanning and template richness. Concatenating fixed channel lifts writes Eqs. (10)–(11) with block-diagonal coefficient matrices. Spanning V_F by their images alone therefore does not supply the unrestricted blocks in Proposition 2. A sufficient choice is one full-range lift with all coefficient couplings. Alternatively, freely oriented rank-one production directions generate the PSD cone by spectral decomposition, and skew templates spanning $\bigwedge^2 V_S$ with unrestricted signed coefficients generate the reversible space, identifying $u \wedge v$ with $uv^\dagger - vu^\dagger$. A rank-r mobility has a decomposition into r such production directions; a fixed direction dictionary need not contain them.

For fixed channels, $\operatorname{rank} M \le \dim \sum_b \operatorname{im} A_{E,b} \le \sum_b \operatorname{rank} A_{E,b}$, with the corresponding bound for L and its even-rank restriction. The functional $v \mapsto \langle g_F, v\rangle$ on $\mathcal{D}$ has rank one exactly when $Pg_F \neq 0$, so $d_F = \dim \mathcal{D} - \mathbf{1}_{Pg_F \neq 0}$ for $F \in \{E, S\}$. The declaration and degeneracy thus determine the target dimension, while channel supports, ranks, and templates determine the represented subset. The proposition is pointwise; smooth choices across changing ranks and the Jacobi identity are separate requirements.

A.5 Gradient-orthogonal completion and task instantiations

Lemma 5 (Supported completion)**.** *Suppose $V = V_x \oplus V_y$ is an orthogonal block decomposition and $\mathcal{D} = \mathcal{D}_x \oplus \mathcal{D}_y$. Let $g = \nabla F$, let P_x project onto $\mathcal{D}_x$, and let Q_a be a declared orthogonal support projector with image in $\mathcal{D}_y$. If $\|Q_a g_y\|^2 > 0$, set*

$$r_a = \frac{Q_a g_y}{\|Q_a g_y\|^2}, \qquad A_{F,a} u = (P_x C_a u, -r_a \langle g_x, P_x C_a u\rangle).$$

Then $\operatorname{im} A_{F,a} \subseteq \mathcal{D}$ and $A_{F,a}^\dagger g = 0$ for every coefficient direction u.

Proof. The two components lie in their respective allowed spaces. Orthogonality and idempotence of Q_a give $\langle g_y, Q_a g_y\rangle = \|Q_a g_y\|^2$, so $\langle g_y, r_a\rangle = 1$. Pairing the two components with g cancels exactly, as in Eq. (9). Since cancellation holds for every u, the adjoint applied to g is zero. □

The completion block can be any declared state block. With a numerical threshold, the channel domain is $\|Q_a g_y\|^2 > \varepsilon$; normalization uses the unmodified denominator. Outside that domain the channel is rejected or explicitly omitted, while other compatible channels may remain available.

In covector coordinates, the same routing is obtained from the supported Riesz representative: if $g_y^{\mathrm{cov}} = d_y F$, then $r_a = Q_a G_y^{-1} g_y^{\mathrm{cov}} / \langle Q_a G_y^{-1} g_y^{\mathrm{cov}}, Q_a G_y^{-1} g_y^{\mathrm{cov}}\rangle_{G_y}$. The coefficient used for compensation is $(d_x F)^\top P_x C_a u$. The derivative includes the dependence of learned scalar densities on the structural state.

When completed lifts span the target space. Under the block decomposition of Lemma 5, let $r \in \mathcal{D}_y$ satisfy $\langle g_y, r\rangle = 1$, and suppose $P_x C$ covers $\mathcal{D}_x$. The completed image is the graph $G_r = \{(x, -r\langle g_x, x\rangle) : x \in \mathcal{D}_x\}$. Every $(x, y) \in \mathcal{D} \cap g^\perp$ has the unique decomposition

$$(x, y) = (x, -r\langle g_x, x\rangle) + (0, y + r\langle g_x, x\rangle),$$

whose second summand lies in $\{0\} \times (\mathcal{D}_y \cap g_y^\perp)$. Thus

$$\mathcal{D} \cap g^\perp = G_r \oplus \big(\{0\} \times (\mathcal{D}_y \cap g_y^\perp)\big). \tag{28}$$

The sum need not be orthogonal. This single routing has dimension $\dim \mathcal{D}_x$, leaving $\dim \mathcal{D}_y - 1$ transverse directions; it is full-range exactly when this number is zero. Adding lifts spanning that space gives a full-range common lift; Proposition 2 then applies if cross-channel coefficient couplings are available. In the thermal specialization, conduction supplies transverse directions to the extent that $\operatorname{im} B_\tau^\dagger$ spans them. Separate production and conduction terms do not alone provide the cross-couplings of an unrestricted common block.

Restriction precedes completion. A new declaration changes the allowed target directions and completion support. Rebuilding the lifts against the same full derivative preserves both requirements. Projecting a previously completed operator preserves symmetry and positivity but requires a separate annihilation condition: $PMPg = 0$ does not follow from $Mg = 0$.

Conservative wave and acoustic pairings. Wave and acoustics take constant S and omit dissipative channels. The wave model combines its canonical covector action with learned skew channels. The acoustic model learns a mixed quadratic energy and an assembled skew pairing with coefficients fixed over each rollout. Both use Eq. (24); the acoustic boundary port accounts for external exchange. Its observed p is a frame difference rather than a calibrated conjugate momentum (§5.2).

Darcy joint-sector mobility. The state is (a, e), cumulative edge flow and thermal energy density, with $L = 0$. The edge declaration is $P_x = I$ or $I - \Pi_{\mathrm{har}}$. Supported completion in e against dE gives A_E and mobility $\mathcal{M} = A_E \operatorname{diag}(\mu_M^2) A_E^\top$. The allowed sectors enter jointly before constitutive weighting and completion. The scalar functionals retain the full state as input when the direction declaration changes.

Beam3D full-state completion. Position, velocity, and stress use the full C0 declaration. Lemma 5 applies to the target/completion blocks with $F = S$ for reversible channels and $F = E$ for dissipative channels. The covector maps $A_S K A_S^\top$ and $A_E R R^\top A_E^\top$ satisfy the common certificate on their completion domains. Here E, S are learned scalar functionals of the observed state.

A.6 Thermal-coordinate specialization

For an explicit thermal coordinate, stronger functional choices recover a mechanical-to-thermal model. The scalar heads define energy and entropy; independent operator heads modulate the registered channels.

Energy. For this specialization, the scalar-functional family in Eq. (7) is restricted to

$$E(z) = E_{\mathrm{kin}}(p; G) + E_{\mathrm{nk},\theta}(q) + E_{\mathrm{th}}(\xi), \tag{29}$$

where the argument G in E_{kin} denotes the declared mass metric, E_{kin} is analytic in the momenta, $E_{\mathrm{nk},\theta}$ is a learned non-kinetic potential, and E_{th} is an analytic thermal law. The learned term leaves this kinetic metric fixed. Write $\nabla E = (g, \tau)$ with $g = \nabla_x E$ the mechanical block and τ the thermal block. On tasks that declare no thermal coordinate, E_{th} is omitted and $\nabla E = g$.

Entropy. μ_S is a thermal co-energy $\beta_\theta(\xi)$ with the componentwise gates

$$\beta_\theta(\xi) > 0, \qquad \partial_\xi \beta_\theta(\xi) \le 0. \tag{30}$$

Let $\omega_i > 0$ be the declared thermal quadrature weight at site i and $\xi_{\mathrm{ref},i}$ a fixed reference. Entropy is the integrable functional

$$S(\xi) = \sum_i \omega_i \int_{\xi_{\mathrm{ref},i}}^{\xi_i} \beta_{\theta,i}(r)\, \mathrm{d}r. \tag{31}$$

Differentiating under the integral gives the coordinate derivative $(d_\xi S)_i = \omega_i \beta_{\theta,i}(\xi_i)$; the metric gradient is $\nabla_\xi S = G_\xi^{-1} d_\xi S$. For $G_\xi = \operatorname{diag}(\omega_i)$ it equals β_θ. In particular S depends only on thermal coordinates, so the mechanical block of ∇S is identically zero:

$$\nabla E = (g, \tau), \qquad \nabla S = (0, s_\xi). \tag{32}$$

A positive nonincreasing parameterization enforces these co-energy conditions. Tasks with constant S omit the entropy readout.

Reversible intensity. μ_L holds coefficients of registered templates on declared *paired sectors*: task-chosen pairs of Hodge subspaces that a template is allowed to couple. A template $T_a[\mu_{L,a}]$ is admissible only if it is skew-adjoint under the declared inner product, is supported on its pair, and has a zero thermal block. Coefficient insertion and its metric-adjoint partner are defined together, so the assembled reversible operator has the block form

$$L = \begin{pmatrix} \sum_a T_a[\mu_{L,a}] & 0 \\ 0 & 0 \end{pmatrix} \tag{33}$$

whenever a thermal coordinate is present, and reduces to $\sum_a T_a[\mu_{L,a}]$ when there is none.

Dissipative intensity. μ_M holds production weights γ_a and a conduction mobility. Admissibility requires $\gamma_a \geq 0$ (square or softplus) and a factorization $R_\kappa = A_\kappa A_\kappa^\dagger \succeq 0$. A production channel is instantiated only if a thermal routing field r_a with $\langle r_a, \tau \rangle = 1$ exists; otherwise $\gamma_a = 0$ and that channel is omitted.

The scalar shape conditions and operator factorizations hold for every admitted readout.

A.7 Thermal assembly lemmas

For reference, the registered mechanical subprojectors satisfy

$$P_a^\dagger = P_a, \quad P_a^2 = P_a, \quad P_a P_x = P_x P_a = P_a, \quad K_a^\dagger = -K_a, \tag{34}$$

and the production directions obey

$$P_{D,a}^\dagger = P_{D,a}, \quad P_{D,a}^2 = P_{D,a}, \quad P_{D,a} P_x = P_x P_{D,a} = P_{D,a}, \quad \widetilde{\varphi}_a = P_{D,a} \varphi_a. \tag{35}$$

The following construction uses $\mathcal{D} = \mathcal{D}_x \oplus \mathcal{D}_\xi$; thermal routing and conduction images must lie in $\mathcal{D}_\xi$ as well. Embedding P_a into the mechanical block and putting zero in the thermal block gives reversible lifts in the common construction.

Lemma 6 (Reversible skew-adjointness)**.** *Let $L = \sum_a T_a[\mu_{L,a}]$ with each T_a admissible. Then $L^\dagger = -L$ under the declared inner product, independently of the numerical values of μ_L.*

Proof. Each admissible template has the form $T_a = P_a^\dagger K_a(\mu_{L,a}) P_a$ with $K_a(\mu_{L,a})^\dagger = -K_a(\mu_{L,a})$ for every coefficient and P_a the registered projector onto the declared pair of T_a. Lemma 2 with $Q = P_a$ (or the same argument with $P_a^\dagger(\cdot)P_a$) gives $T_a^\dagger = -T_a$. A finite sum of embedded skew-adjoint operators is skew-adjoint by Lemma 1. Conjugating by the global declaration P preserves the identity, again by Lemma 2. The coefficients $\mu_{L,a}$ scale or modulate K_a inside the admissible family; they do not change $K_a^\dagger = -K_a$. The zero thermal block in (33) is itself skew-adjoint. □

Lemma 7 (Reversible degeneracy)**.** *If S is thermal-only as in §A.6 and each T_a has a zero thermal block, then $L \nabla S = 0$.*

Proof. Let Π_{th} extract thermal coordinates, so $\Pi_{\mathrm{th}} z = \xi$ and $\Pi_{\mathrm{th}}^\dagger$ embeds a thermal cochain as $(0, \cdot)$ in the concatenated state. Thermal-only entropy means (32): $\nabla S = \Pi_{\mathrm{th}}^\dagger s_\xi$. A zero thermal block means $\Pi_{\mathrm{th}} T_a = 0$ and $T_a \Pi_{\mathrm{th}}^\dagger = 0$ for every a, hence $\Pi_{\mathrm{th}} L = 0$ and $L \Pi_{\mathrm{th}}^\dagger = 0$. Therefore

$$L \nabla S = L \Pi_{\mathrm{th}}^\dagger s_\xi = 0.$$

The identity does not use the values of s_ξ or of μ_L. □

Lemma 8 (Production-channel compensation)**.** *Let $\widetilde{\varphi}_a$ be a mechanical direction on the declared dissipative directions and let $a_a \in \mathcal{D}_\xi$ be a thermal routing field. If $\langle a_a, \tau \rangle$ is finite and nonzero, set*

$$r_a = \frac{a_a}{\langle a_a, \tau \rangle}, \qquad q_a = \langle \widetilde{\varphi}_a, g \rangle, \qquad w_a = \begin{pmatrix} \widetilde{\varphi}_a \\ -q_a r_a \end{pmatrix}. \tag{36}$$

Then $\langle r_a, \tau \rangle = 1$ and $\langle w_a, \nabla E \rangle = 0$. If no such r_a exists, the channel is not assembled.

Proof. Whenever the denominator is finite and nonzero, $\langle r_a, \tau \rangle = \langle a_a, \tau \rangle / \langle a_a, \tau \rangle = 1$ by construction. Expanding the declared pairing against $\nabla E = (g, \tau)$ then gives

$$\langle w_a, \nabla E \rangle = \langle \widetilde{\varphi}_a, g \rangle + \langle -q_a r_a, \tau \rangle = q_a - q_a \langle r_a, \tau \rangle = q_a - q_a = 0.$$

The normalization uses the thermal component τ of ∇E. A zero or nonfinite pairing omits the channel as specified in §A.6. The same cancellation holds for every allowed mechanical direction $\widetilde{\varphi}_a = P_{D,a} \varphi_a$. □

Lemma 9 (Conduction degeneracy)**.** *Let B_τ act on thermal cochains with $B_\tau \tau = 0$ and let $R_\kappa \succeq 0$. The conduction operator $M_{\mathrm{c}} = \Pi_{\mathrm{th}}^\dagger B_\tau^\dagger R_\kappa B_\tau \Pi_{\mathrm{th}}$ satisfies $M_{\mathrm{c}} \nabla E = 0$ and $M_{\mathrm{c}} = M_{\mathrm{c}}^\dagger \succeq 0$.*

Proof. Self-adjointness follows from Lemma 1:

$$M_{\mathrm{c}}^{\dagger} = \Pi_{\mathrm{th}}^{\dagger} B_{\tau}^{\dagger} R_{\kappa}^{\dagger} B_{\tau} \Pi_{\mathrm{th}} = M_{\mathrm{c}},$$

using $R_{\kappa}^{\dagger} = R_{\kappa}$ and $(\Pi_{\mathrm{th}}^{\dagger})^{\dagger} = \Pi_{\mathrm{th}}$. For any test vector v,

$$\langle v, M_{\mathrm{c}} v \rangle = \langle B_{\tau} \Pi_{\mathrm{th}} v,\ R_{\kappa} B_{\tau} \Pi_{\mathrm{th}} v \rangle \geq 0,$$

because $R_{\kappa} \succeq 0$. The thermal block of ∇E is τ, so

$$M_{\mathrm{c}} \nabla E = \Pi_{\mathrm{th}}^{\dagger} B_{\tau}^{\dagger} R_{\kappa} B_{\tau} \tau = 0.$$

The factorization $R_{\kappa} = A_{\kappa} A_{\kappa}^{\dagger}$ is one admissible way to enforce $R_{\kappa} \succeq 0$; any other PSD representative of the same mobility is equally admissible. □

Lemma 10 (Dissipative assembly)**.** *The thermal specialization*

$$M = \sum_{a} \gamma_a w_a w_a^{\dagger} + \Pi_{\mathrm{th}}^{\dagger} B_{\tau}^{\dagger} R_{\kappa} B_{\tau} \Pi_{\mathrm{th}}, \tag{37}$$

with $\operatorname{im} B_{\tau}^{\dagger} \subseteq \mathcal{D}_{\xi}$*, is an instance of Eq.* (11)*: each production lift is* $u \mapsto w_a u$ *and the conduction lift is the thermally embedded* $B_{\tau}^{\dagger}$*. It satisfies* $M = M^{\dagger} \succeq 0$ *and* $M \nabla E = 0$ *for every admissible* μ_M*.*

Proof. Each production term $\gamma_a w_a w_a^{\dagger}$ with $\gamma_a \geq 0$ is self-adjoint because $(\gamma_a w_a w_a^{\dagger})^{\dagger} = \gamma_a w_a w_a^{\dagger}$. For any test vector r,

$$\langle r, (\gamma_a w_a w_a^{\dagger}) r \rangle = \gamma_a \left| \langle w_a, r \rangle \right|^2 \geq 0, \tag{38}$$

so the term is positive semidefinite. Degeneracy is Lemma 8:

$$w_a w_a^{\dagger} \nabla E = w_a \langle w_a, \nabla E \rangle = 0.$$

The conduction term is Lemma 9. A sum of self-adjoint positive-semidefinite operators that each annihilate ∇E remains self-adjoint, positive semidefinite, and annihilating ∇E. If every production channel is omitted and conduction is absent, $M = 0$, which satisfies the same identities. □

Relation to the common certificate. These lemmas recover Theorem 1 with zero-thermal reversible lifts, compensated production, and energy-null conduction. The thermal co-energy shape conditions add physical assumptions to the common algebraic certificate.

A.8 When restriction binds

Lemma 11 (Harmonic sector)**.** *Under* (3) *and Lemma 3,* $\dim \ker \Delta_1 = \dim H^1(\mathcal{K})$*. The declaration* $P = I - P_{\ker \Delta_1}$ *is a proper subspace constraint if and only if* $\dim \ker \Delta_1 > 0$*. For a closed orientable surface this dimension is* $2g$ *with* g *the genus (Hirani, 2003; Lim, 2020).*

Proof. Lemma 3 identifies $\ker \Delta_1$ with the harmonic summand $\mathcal{H}^1 = \ker D_1 \cap \ker D_0^{\dagger}$, which is isomorphic to the first cohomology of the cochain complex and hence to $H^1(\mathcal{K})$. Let P_h be the metric projector of Lemma 4 onto a basis of $\ker \Delta_1$. Then P_h is the identity on that summand and zero on $\operatorname{im} D_0 \oplus \operatorname{im} D_1^{\dagger}$. Consequently $I - P_h$ fails to be the identity if and only if the harmonic summand is nontrivial. If $\dim \ker \Delta_1 = 0$, then $P_h = 0$ and $I - P_h = I$. □

Beam3D uses the full C0 declaration; Darcy tests a proper edge-space restriction. Thus the experiments separate operator compatibility from the additional effect of excluding an available Hodge sector.

B Neural dynamical assembly: algorithm and finite-step analysis

Algorithm 1 connects the learned constitutive quantities to an executed state update. The finite-step analysis then relates algebraic compatibility to the balance attained by the integrator.

The encoder and processor are those of a graph neural simulator (Sanchez-Gonzalez et al., 2020; Pfaff et al., 2021). Assembly maps the readouts deterministically to (E, S, L, M); rollout supervision trains the readouts through that map.

Algorithm 1 One MAGNS state update

Input: mesh state z, mesh $\mathcal{K}$ (incidence, Hodge stars, declared sectors, ports), integrator
Encode z on $G = (V, E)$ and run the shared message-passing processor
Read out task-active $\mu_E, \mu_S, \mu_L, \mu_M$ and differentiate the scalar functionals
Construct declared lifts against ∇S and ∇E as in Eq. (8)
Assemble $L = \sum_a A_{S,a} K_a A_{S,a}^\dagger$ and $M = \sum_b A_{E,b} R_b A_{E,b}^\dagger$ (empty families give zero)
Form $\dot{z} = L\nabla E + M\nabla S + B_\partial u$
Advance with the declared integrator, re-evaluating assembly at its stages or state-pair trials
Output: z_{t+1} and, if requested, algebraic residuals of (6)

Complexity. Let $n = |V|$, $m = |E|$, B the number of processor blocks, and d the hidden width. One processor pass costs $O(Bmd^2)$ for standard edge-wise MLPs. Assembly applies registered templates on declared sectors rather than forming dense $|z| \times |z|$ matrices. A factorized rank-r_a channel on n_a coordinates costs $O(n_a r_a + r_a^2)$ to apply with a dense coefficient block, or $O(n_a r_a)$ with a diagonal block, in addition to projector and functional-derivative costs. Dense state-space operators require $O(|z|^2)$ storage. An explicit integrator uses a fixed number of field evaluations per substep. An implicit discrete-gradient method repeats the evaluation during its nonlinear solve; its cost depends on the iteration count.

B.1 Finite-step identities and error bounds

Theorem 1 certifies the continuous generator. The results below quantify how a numerical step transfers that structure to finite changes in energy and entropy. Prediction accuracy is evaluated in §5.

B.2 Two-point discrete gradients

For $F \in \{E, S\}$, a state-pair discrete gradient $\nabla_d F = \overline{\nabla F}(z_n, z_{n+1})$ is any map obeying the exact directional identity

$$\langle \nabla_d F,\, z_{n+1} - z_n \rangle = F(z_{n+1}) - F(z_n) \tag{39}$$

(Gonzalez, 1996; McLachlan et al., 1999). When $z_{n+1} = z_n$, the convention is $\nabla_d F = \nabla F(z_n)$, which satisfies (39) as $0 = 0$. For F continuously differentiable on a neighborhood of the segment, the integral

$$\nabla_d F = \int_0^1 \nabla F\big((1-\theta)z_n + \theta z_{n+1}\big)\,\mathrm{d}\theta \tag{40}$$

is one such map, because

$$\begin{aligned}\langle \nabla_d F, z_{n+1} - z_n \rangle &= \int_0^1 \langle \nabla F(z_n + \theta\Delta z),\, \Delta z \rangle\,\mathrm{d}\theta \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} F(z_n + \theta\Delta z)\,\mathrm{d}\theta = F(z_{n+1}) - F(z_n),\end{aligned}$$

writing $\Delta z = z_{n+1} - z_n$. Gonzalez's midpoint correction provides another discrete gradient. Thermal-only entropy requires a discrete gradient with zero mechanical entries when paired with zero-thermal reversible blocks. Full-state functionals use lifts completed against the full two-point gradients. The same $\nabla_d E$ enters both the pairing identity and dissipative assembly.

B.3 Pairwise operators and the step residual

Write $\overline{L}, \overline{M}$ for operators using midpoint constitutive coefficients and lifts completed against $\nabla_d S, \nabla_d E$. At the state pair they satisfy

$$\overline{L}^\dagger = -\overline{L}, \qquad \overline{M} = \overline{M}^\dagger \succeq 0, \qquad \overline{L}\,\nabla_d S = 0, \qquad \overline{M}\,\nabla_d E = 0. \tag{41}$$

The nonlinear residual of one implicit step is

$$R = z_{n+1} - z_n - \Delta t\big(\overline{L}\,\nabla_d E + \overline{M}\,\nabla_d S\big) - \Delta t\, B_\partial u. \tag{42}$$

An exact root has $R = 0$. Numerically, the converged trial is accepted when its residual and structural conditions meet the prescribed tolerances.

Lemma 12 (Discrete energy and entropy defects)**.** *Let $\Delta t > 0$. Assume* (39) *for E and S and assume* (41) *at the pair. Let $W_{\mathrm{port}} = \Delta t\langle \nabla_d E, B_\partial u\rangle$, with $W_{\mathrm{port}} = 0$ on a closed system. Then*

$$\Delta E - W_{\mathrm{port}} = \langle \nabla_d E,\, R\rangle, \qquad \big|\Delta E - W_{\mathrm{port}}\big| \le \|\nabla_d E\|\,\|R\|, \tag{43}$$

and

$$\Delta S = \Delta t\langle \nabla_d S,\, \overline{M}\,\nabla_d S\rangle + \langle \nabla_d S, R\rangle + \Delta t\langle \nabla_d S, B_\partial u\rangle, \tag{44}$$

hence, on a closed system,

$$\Delta S \ge \Delta t\langle \nabla_d S,\, \overline{M}\,\nabla_d S\rangle - \|\nabla_d S\|\,\|R\|. \tag{45}$$

In particular, an exact root on a closed system satisfies $\Delta E = 0$ and $\Delta S = \Delta t\langle \nabla_d S, \overline{M}\nabla_d S\rangle \ge 0$.

Proof. The residual (42) rearranges to

$$z_{n+1} - z_n = \Delta t\big(\overline{L}\nabla_d E + \overline{M}\nabla_d S + B_\partial u\big) + R.$$

Pair with $\nabla_d E$ and apply (39):

$$\begin{aligned}\Delta E &= \langle \nabla_d E, z_{n+1} - z_n\rangle \\ &= \Delta t\langle \nabla_d E, \overline{L}\nabla_d E\rangle + \Delta t\langle \nabla_d E, \overline{M}\nabla_d S\rangle \\ &\quad + \Delta t\langle \nabla_d E, B_\partial u\rangle + \langle \nabla_d E, R\rangle.\end{aligned}$$

The first pairing is zero because $\overline{L}^\dagger = -\overline{L}$ (Lemma 1). The second is zero because $\overline{M}^\dagger = \overline{M}$ and $\overline{M}\nabla_d E = 0$:

$$\langle \nabla_d E, \overline{M}\nabla_d S\rangle = \langle \overline{M}\nabla_d E, \nabla_d S\rangle = 0.$$

The remaining terms are $W_{\mathrm{port}} + \langle \nabla_d E, R\rangle$, which is (43). Cauchy–Schwarz gives the absolute bound.

Pair instead with $\nabla_d S$:

$$\begin{aligned}\Delta S &= \Delta t\langle \nabla_d S, \overline{L}\nabla_d E\rangle + \Delta t\langle \nabla_d S, \overline{M}\nabla_d S\rangle \\ &\quad + \Delta t\langle \nabla_d S, B_\partial u\rangle + \langle \nabla_d S, R\rangle.\end{aligned}$$

The reversible pairing vanishes because $\langle \nabla_d S, \overline{L}\nabla_d E\rangle = -\langle \overline{L}\nabla_d S, \nabla_d E\rangle = 0$. The production pairing is nonnegative because $\overline{M} \succeq 0$. Dropping the port term on a closed system and applying Cauchy–Schwarz to $\langle \nabla_d S, R\rangle$ yields (45). If $R = 0$ and $B_\partial u = 0$, the identities collapse to $\Delta E = 0$ and $\Delta S = \Delta t\langle \nabla_d S, \overline{M}\nabla_d S\rangle \ge 0$. □

For residual tolerance $\|R\| \le \varepsilon$, Lemma 12 bounds the energy defect by $\|\nabla_d E\|\varepsilon$ and the entropy shortfall by $\|\nabla_d S\|\varepsilon$. Each nonlinear trial re-evaluates the endpoint functionals, midpoint readouts, two-point gradients, and assembled operators. Acceptance requires the residual tolerance and all structural and state-domain conditions; the accepted state is the converged trial. The wave results in §C report this finite-step balance alongside prediction error.

B.4 Local consistency of the two-point assembly

Proposition 3 (Diagonal consistency)**.** *Suppose E, S are C^1 near z_n and the chosen discrete gradients extend continuously to the diagonal with value $\nabla E(z_n), \nabla S(z_n)$. Assume continuous constitutive readouts, templates, and lift constructions near this pair, with every completion denominator bounded away from zero. If the evaluated port also converges to its value at z_n, then as $z_{n+1} \to z_n$,*

$$\overline{L}\nabla_d E + \overline{M}\nabla_d S + B_\partial u \to L(z_n)\nabla E(z_n) + M(z_n)\nabla S(z_n) + B_\partial u(z_n). \tag{46}$$

Proof. Continuity of the discrete gradients gives their diagonal limits; Eq. (40) has this property by continuity of ∇F. The readouts and nondegenerate completions then give $(\overline{L}, \overline{M}) \to (L(z_n), M(z_n))$. Taking the limit in the products and port term proves the claim. □

B.5 Explicit stepping, KDK, and Taylor drift

For a sufficiently smooth assembled field $F(z) = L\nabla E + M\nabla S + B_\partial u$, a second-order explicit method has local truncation error $O(\Delta t^3)$ and global error $O(\Delta t^2)$ on a fixed horizon under the usual stability assumptions (Hairer et al., 2006). KDK instead uses the declared splitting (Öttinger, 2018; Shang & Öttinger, 2020). Their energy defects follow from their update equations; the discrete-gradient identities above require the two-point assembly and residual in Eq. (42).

Proposition 4 (Explicit Euler energy drift)**.** *Let $z_{n+1} = z_n + \Delta t\, F(z_n)$ with $F = L\nabla E + M\nabla S$ on a closed system, and suppose E is C^2 with Hessian bounded by β along the segment. Then*

$$E(z_{n+1}) - E(z_n) = R_n, \qquad |R_n| \le \tfrac{1}{2}\beta\Delta t^2 \|F(z_n)\|^2. \tag{47}$$

The first-order term vanishes by Theorem 1; the remainder is the ordinary Taylor defect of an explicit step.

Proof. Taylor expansion with integral remainder gives

$$E(z_n + \Delta t F) = E(z_n) + \Delta t \langle \nabla E(z_n), F(z_n)\rangle + R_n,$$

where $|R_n| \le \frac{1}{2}\beta\Delta t^2\|F(z_n)\|^2$. The first-order pairing is $\dot E$ at z_n, which is zero on a closed system by Theorem 1. What remains is the second-order remainder. □

The wave comparison in §5.1 holds the generator fixed while changing the integrator. The acoustic experiment in §5.2 trains with Heun and evaluates with KDK, using the same learned mixed energy and assembled reversible operator. Its finite-step balance is recorded in Table 5.

B.6 Conditional state-error recursion

Let $z(t)$ solve the assembled continuous field and let $e_n = \|\widehat{z}_n - z(t_n)\|$. Suppose the exact-root numerical map $\Psi_{\Delta t}$ is stable with Lipschitz constant at most $1 + C_{\mathrm{st}}\Delta t$ for $C_{\mathrm{st}} \ge 0$, has local defect at most $C_{\mathrm{tr}}\Delta t^{p+1}$, and a residual R_n perturbs the root by at most $C_{\mathrm{sol}}\|R_n\|$. Comparing the accepted step with $\Psi_{\Delta t}(\widehat{z}_n)$ and then with $\Psi_{\Delta t}(z(t_n))$ gives

$$e_{n+1} \le (1 + C_{\mathrm{st}}\Delta t)\, e_n + C_{\mathrm{tr}}\Delta t^{p+1} + C_{\mathrm{sol}}\|R_n\|. \tag{48}$$

A nonsingular residual Jacobian gives such a local residual-to-state estimate when the accepted trial remains in its inverse-function neighborhood. A uniform bound over the steps supplies C_{sol}.

A discrete Gronwall bound on a horizon $T = N\Delta t$ gives

$$e_N \le \mathrm{e}^{C_{\mathrm{st}}T}\Big(e_0 + TC_{\mathrm{tr}}\Delta t^p + C_{\mathrm{sol}}\sum_{n=0}^{N-1}\|R_n\|\Big). \tag{49}$$

This bound compares numerical states with the assembled dynamics. Constitutive approximation and observation error enter the separate comparison with data.

B.7 Gradient path through executed steps

Training differentiates the executed solver iterations. For fixed observed z_0 and a closed-loop window of horizon H on a differentiable accepted branch, let Φ_j denote one step, including its unrolled nonlinear trials, and write $z_H = \Phi_{H-1}\circ\cdots\circ\Phi_0(z_0, \vartheta)$ with ϑ the differentiable readout parameters. For a differentiable observed loss $\ell(z_H)$,

$$\frac{\partial \ell}{\partial \vartheta} = \frac{\partial \ell}{\partial z_H}\sum_{j=0}^{H-1}\left(\frac{\partial \Phi_{H-1}}{\partial z_{H-1}}\cdots\frac{\partial \Phi_{j+1}}{\partial z_{j+1}}\right)\frac{\partial \Phi_j}{\partial \vartheta}. \tag{50}$$

The Jacobian product is the identity when $j = H-1$. Gradients pass through the predicted states, constitutive contexts, and solver trials used by the rollout loss. Structural residuals measure the resulting identities; compatibility itself comes from assembly.

Energy diagnostics. The balance defect in Eq. (43) uses the model's scalar E and declared port. A gap between predicted and reference chart energies measures a different quantity: disagreement between trajectories under a shared diagnostic. The acoustic results report both alongside field error.

C STRUCTURAL DIAGNOSTICS AND PORT LEDGERS

Theorem 1 concerns the assembled continuous-time generator. Here we measure its algebraic residuals and follow the associated energy balance through numerical integration. For open acoustics, the balance includes boundary work; for Beam3D and Darcy, it separates reversible pairing from dissipative production.

C.1 ANALYTIC WAVE

The continuous energy-pairing residual is 1.91×10^{-13}. With the same weights, discrete-gradient stepping lowers physical-energy drift by 94.9% relative to explicit stepping (Table 1). The solver completes all 40 closed-loop and 1280 one-step checks. The kinematic residual is 1.43×10^{-3}, and the learned-L contribution accounts for 0.145% of the operator action.

C.2 ACOUSTICS: DECLARED ENERGY AND BOUNDARY WORK

The acoustic model has $M = 0$ and a static learned energy E_θ. For each $H = 80$ rollout, the signed balance residual is

$$\mathcal{R}_E = E_\theta(z_{80}) - E_\theta(z_0) - W_\partial,$$

where W_∂ is the accumulated work of the declared boundary port. Table 5 records this balance under the KDK evaluation used in the main text. All quantities use the model's learned energy; the common-energy comparison with ground truth is given separately in Table 10.

Table 5: MAGNS declared-energy balance at $H = 80$, means over 200 windows per split. Negative boundary work denotes energy leaving the domain. The final column is the mean of the per-window absolute residual divided by initial declared energy.

Split	$E_\theta(z_0)$	$E_\theta(z_{80})$	W_∂	$\mathcal{R}_E$	$\lvert\mathcal{R}_E\rvert/\lvert E_\theta(z_0)\rvert$
VALID	155.472	133.449	−16.460	−5.563	2.57%
TEST	153.510	130.533	−17.707	−5.270	2.22%

The declared port removes energy on both splits. KDK leaves a mean relative absolute balance residual of 2.57% on VALID and 2.22% on TEST, These finite-step residuals complement the continuous balance identity and the field errors in Table 2.

C.3 BEAM3D

Bilinear probes on TEST rollouts measure skew-adjointness, symmetry, and degeneracy of the assembled MAGNS operators. Table 6 also gives the local-block quantities exposed by Local-TIGNN.

Table 6: Beam3D structural probes. MAGNS numbers are three-seed means on TEST rollouts. Local-TIGNN probes act on its local blocks.

Model	Quantity / scope	Reported value
MAGNS	skew probe (relative)	6.2×10^{-18}
MAGNS	symmetry probe (relative)	1.4×10^{-18}
MAGNS	energy-rate residual (relative)	1.6×10^{-21}
MAGNS	$\mathrm{RMS}(L\nabla S)$ / $\mathrm{RMS}(M\nabla E)$	2.4×10^{-11} / 1.2×10^{-12}
MAGNS	discrete $\lvert\Delta E\rvert$ / $\lvert\Delta S\rvert$ step defect (mean)	2.8×10^{-6} / 1.7×10^{-5}
Local-TIGNN	local-block skew / symmetry	0 / 0
Local-TIGNN	min. sampled local-block eigenvalue	-9.8×10^{-8}

Local-TIGNN matrices use float32, with eigenvalues computed in float64. Its minimum sampled eigenvalue is within float32 roundoff. For MAGNS, operator residuals remain near machine precision across the three seeds; the larger discrete-step defects reflect the separate effect of time integration.

D REPRODUCIBILITY AND EXPERIMENTAL PROTOCOLS

The experiments separate three questions: predictive accuracy, structural balance, and control over admissible directions. This section specifies their data and training protocols and gives the short-horizon and trajectory-level results underlying Section 5. Throughout, H denotes the number of closed-loop prediction steps; model selection uses VALID.

MAGNS uses width/block counts $64/4$ for waves, $48/3$ for acoustics, $80/9$ for Beam3D, and $32/2$ for Darcy. Wave training uses five seeds and batch 8; Beam3D uses three seeds and batch 8; Darcy uses three seeds and batch 2. Acoustic training uses seed 0 and batch 1. Task-specific horizons and budgets follow below.

D.1 ANALYTIC WAVE

Periodic 32×32 mesh, state (q, p), $\Delta t = 0.002$, $T = 200$, closed system, $c = 1$, $k_{\max} = 4$. Data follow MeshFT `PlaneWaveDataset` (Noguchi & Kawahara, 2026). Five seeds $\{0, \ldots, 4\}$ each control data, initialization, and training order, with 2000 training and 256 validation samples . Training uses Heun stepping for 10 epochs (batch 8, 2500 updates), with a fixed epoch-10 checkpoint. The same weights are then evaluated under Heun stepping and under discrete gradient. Rollout TSMSE uses $T = 200$. Table 1 compares these five-seed results with the published MeshFT, MGN, MGN-HP, and HNN aggregates.

D.2 ACOUSTICS

The Well `acoustic_scattering_discontinuous` dataset (Ohana et al., 2024) provides 256×256 fields at frame interval $\Delta t = 2/101$. We retain the causal observation chart (q_t, p_t), where $p_{t,i} = m_i(q_{t,i} - q_{t-1,i})/\Delta t$ with $m_i = \rho_i V_i$, and the unit-weight q/p prediction loss. Training draws from a frozen pool of 8000 windows, 2000 from each of four TRAIN shards. Evaluation uses 200 fixed VALID windows from shard 16 (86 trajectories) and 200 fixed TEST windows from shard 18 (90 trajectories), shared by all models.

MAGNS uses the geometry encoder, shared processor, modal attention, and typed structural state with width 48 and 3 blocks. Material-conditioned readouts define a positive mixed quadratic energy with learned q/p coupling and an assembled reversible operator; coefficients remain fixed during a rollout. The boundary exchange is represented by the declared port. Pretraining performs 1000 $H = 1$ updates; $H = 4$ training initializes from those weights and performs 1000 updates using AdamW. Both stages use batch 1 and seed 0. We evaluate the fixed final $H = 4$ checkpoint. The neural processor uses float32 and structural arithmetic uses float64.

Training uses Heun with 21 substeps per data frame. Evaluation uses KDK with 42 substeps and the same learned energy, reversible operator, and boundary port. The sparse GMRES solves use absolute tolerance 10^{-11}, relative tolerance 10^{-9}, restart 16, maximum 64 iterations, and damping 1. All 200 windows complete at $H = 80$ in each split.

The baselines share the causal (q_t, p_t) observation chart, evaluation windows, and metric definitions. The MeshFT adaptation (Noguchi & Kawahara, 2026) uses width 64, four graph blocks, positive learned node masses and edge weights, and KDK with 8 substeps per frame. Its $H = 1$ pretraining and $H = 4$ training each receive an 8000-update budget at batch 1. This retains mesh-based Hamiltonian evolution while adapting the constitutive readout to the common observations.

IGNS uses width 128 and four latent Hamiltonian blocks. It initializes from its $H = 8$ sequence model and receives 2000 $H = 80$ updates at batch 4: the smaller update count corresponds to 8000 training-window presentations, rather than a fourfold reduction in that count. The latent state evolves throughout each sequence, with decoded increments accumulated in the observed coordinates. MGN uses width 128 and 15 message-passing blocks, with $H = 1$ pretraining under an 8000-update budget followed by $H = 4$ training at batch 1. The latter is stopped at approximately 5000 updates after the training loss reaches a plateau. Its predicted state increments are fed back autoregressively. These settings align the prediction task while retaining mesh integration for MeshFT, latent evolution for IGNS, and state-feedback rollout for MGN.

D.2.1 SHORT-HORIZON EVALUATION

The recorded $H = 8$ comparison uses $H = 1$-trained MAGNS, MeshFT, and MGN models with an 8000-update budget. IGNS uses its native $H = 8$ sequence supervision under the same budget. The short-horizon MAGNS configuration has width 64, four blocks, and canonical-coupling KDK evaluation; the main $H = 80$ results use the width-48 mixed-energy model described above. Table 7 reports 200 windows per split, with each row's one-step and rollout errors taken from the same checkpoint.

Table 7: Acoustic $H = 8$ results. Values are medians over 200 windows per split; model-specific training horizons are stated above.

Model	Split	One-step MSE	Rollout MSE	Chart-energy gap
MAGNS	VALID	1.35×10^{-4}	3.45×10^{-3}	0.28
MeshFT	VALID	5.79×10^{-4}	5.08×10^{-3}	1.11
IGNS	VALID	1.69×10^{-3}	4.62×10^{-3}	14.1
MGN	VALID	3.83×10^{-3}	0.0328	2.54×10^{4}
MAGNS	TEST	1.14×10^{-4}	3.03×10^{-3}	0.255
MeshFT	TEST	4.55×10^{-4}	4.98×10^{-3}	1.06
IGNS	TEST	1.69×10^{-3}	4.91×10^{-3}	15.6
MGN	TEST	3.56×10^{-3}	0.0305	2.92×10^{4}

D.3 BEAM3D: ACCURACY ACROSS TRAJECTORIES

Forced viscoelastic beam (Hernández et al., 2024), $H = 20$ closed-loop normalized MSE. MAGNS uses hidden width 80, 9 processor blocks, batch 8, 22K updates, and three seeds $\{0, 1, 2\}$. MGN, IGNS, and Local-TIGNN also receive 22K updates; Local-TIGNN uses width 160 (Tierz et al., 2025). Table 8 extends the main-text VALID comparison to TEST. MAGNS has the lowest mean error on both splits. All models have lower TEST errors; VALID includes the difficult trajectory examined in Table 4.

Table 8: Beam3D $H = 20$ closed-loop normalized MSE, three-seed mean $\pm$ sample standard deviation. Checkpoints are selected on VALID.

Model	VALID	TEST
MAGNS	0.03533 ± 0.00537	0.00118 ± 0.00091
IGNS	0.04518 ± 0.00420	0.00693 ± 0.00167
Local-TIGNN (160)	0.04679 ± 0.00466	0.00781 ± 0.00093
MGN	0.06662 ± 0.00748	0.01525 ± 0.00504

D.4 WHERE OPERATOR ASSEMBLY IMPROVES PREDICTION

Table 4 of the main text is seed 0, 22K, VALID. The four rows compare full assembly with two broader relaxations and a free-increment decoder: full assembly, free operators (structured operator parameterization replaced by equal-rank free L, M), no assembly (direct Δz), and an equal-capacity control. All four variants retain the Hodge encoder and modal attention. The free-operator comparison isolates structured operator construction; Darcy separately tests Hodge-sector control.

Table 9: Seed-0 VALID $H = 20$ normalized MSE by trajectory (lexicographic ids). Hard case `dfb60d` is retained in the means.

Trajectory	Full	Free op.	No assem.	Equal-cap.
`1ea75a`	0.00194	0.00228	0.0483	0.0213
`2f721c`	0.00387	0.00792	0.177	0.0358
`4f188b`	0.00059	0.00081	0.0467	0.0199
`8e2ab6`	0.0368	0.0122	0.132	0.0706
`dfb60d`	0.155	0.178	0.493	0.246
Five-traj. mean	0.0396	0.0402	0.180	0.0788
Other four mean	0.0108	0.00580	0.101	0.0369
Hard-case share	78.2%	88.5%	54.9%	62.5%

Removing assembly raises the error on both the hard case and the other four trajectories (0.101 versus 0.011). Free operators approach full assembly in the aggregate, with a larger error on the hard case.

D.5 Sector restriction on this mesh

Beam3D uses the full node direction space. Darcy supplies the complementary test on a mesh with nontrivial harmonic directions: the same trained weights are executed with and without those directions in the dissipative assembly. Section G gives the paired results and the reverse switch for harmonic-excluded training.

E Acoustic propagation and energy accounting

E.1 Accuracy across the rollout

The acoustic results in Section 5.2 concern the final static mixed-energy model. Figure 4 follows all four models over the same 200 windows in each split. MAGNS and MeshFT retain substantially lower state errors throughout the rollout than IGNS and MGN. The similarity between VALID and TEST curves is consistent with the mean-MSE improvements in Table 2.

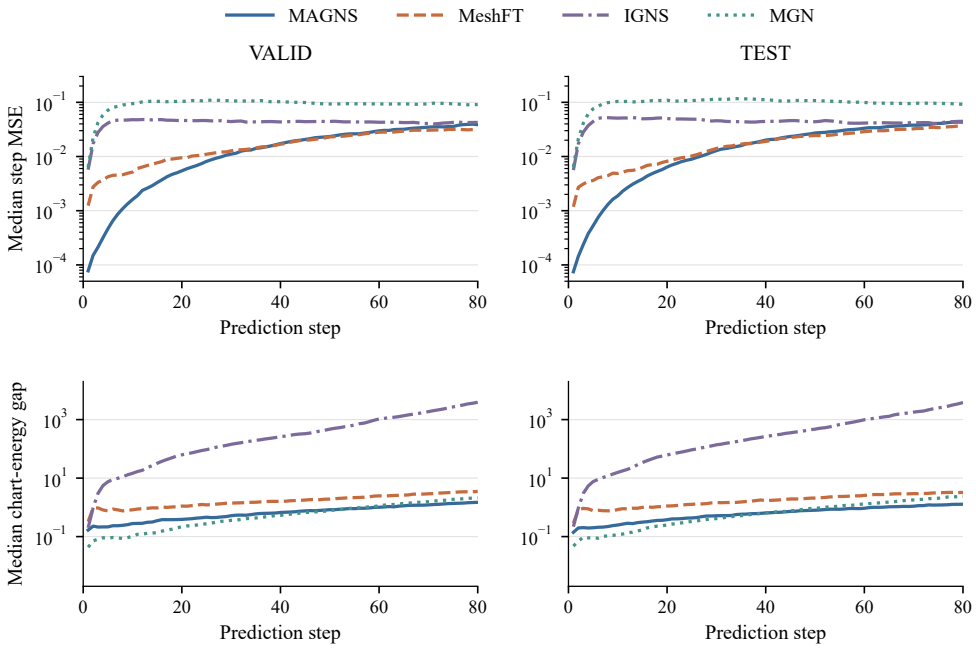

Figure 4: Acoustic errors over 200 windows per split. Top: median step-wise MSE. Bottom: median relative common-chart energy gap. All curves use the final checkpoints reported in Table 2.

MAGNS's mean prefix MSE on TEST rises from 2.63×10^{-4} at $H = 4$ to 3.20×10^{-3} at $H = 20$, 9.11×10^{-3} at $H = 40$, and 2.19×10^{-2} at $H = 80$. The corresponding VALID values are 2.32×10^{-4}, 2.68×10^{-3}, 7.85×10^{-3}, and 2.04×10^{-2}. These trajectories resolve how error accumulates beyond the training horizon.

E.2 Propagation in physical space

Figure 5 locates propagation errors on VALID window $(16, 0, 3)$. The reference column follows the outgoing wavefront; the remaining columns show absolute errors on one logarithmic scale. MAGNS and MeshFT retain lower errors across much of the domain, with discrepancies concentrated around the propagating and scattered fronts. Their rollout MSEs on this window are comparable (0.01984 and 0.01972), consistent with their close median errors in the population comparison.

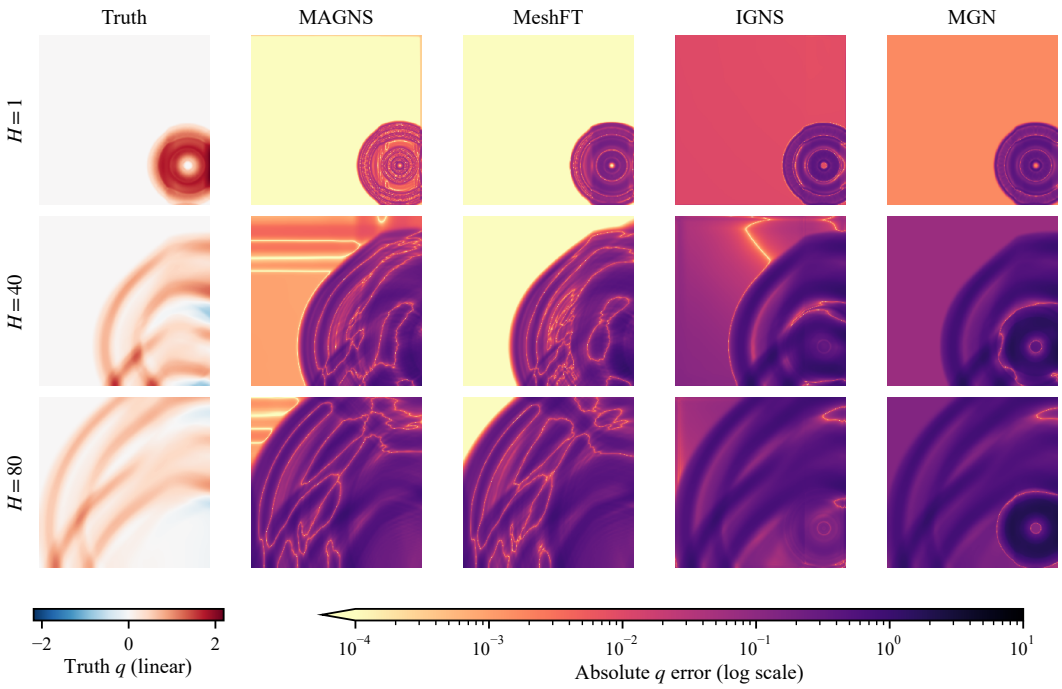

Figure 5: Acoustic propagation at horizons 1, 40, and 80. First column: reference q on a signed linear scale. Remaining columns: $|\hat{q} - q|$ on a common logarithmic scale across models and times; errors below 10^{-4} use the lightest color. MAGNS uses the final $H = 1/H = 4$ checkpoint; baselines match Table 2.

The second state channel helps explain why a field-error curve and an energy curve can tell different stories. Figure 6 separates propagation error in q from the RMS amplitude of the scaled frame-difference channel p. At $H = 80$, MAGNS and MeshFT have p RMS 3.60×10^{-4} and 3.50×10^{-4}, compared with 2.34×10^{-4} in truth. IGNS reaches 9.29×10^{-3}; inverse-mass weighting makes this discrepancy prominent in the quadratic diagnostic.

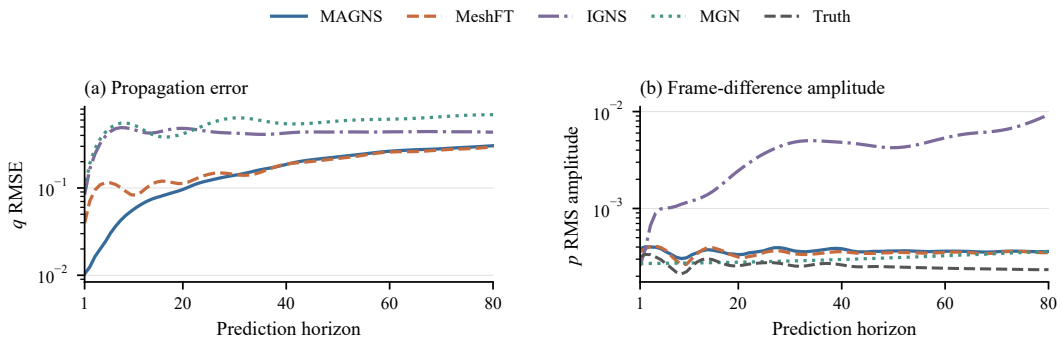

Figure 6: Channel-resolved evolution on the window in Figure 5: spatial RMSE in q and RMS amplitude of p. Separating the channels links the spatial propagation errors to the common-chart energy curves.

E.3 What the energy curves measure

The common chart evaluates all models and the reference fields with

$$E_{\text{chart}}(q, p) = \tfrac{1}{2} \sum_i \frac{p_i^2}{m_i} + \tfrac{1}{2} \sum_{e=(i,j)} w_e (q_j - q_i)^2, \qquad m_i = \rho_i V_i. \tag{51}$$

The observation $p_{t,i} = m_i(q_{t,i} - q_{t-1,i})/\Delta t$ is constructed from a backward frame difference; this mass scaling alone does not establish a calibrated physical acoustic-energy interpretation. Consequently, Eq. (51) measures a common quadratic state diagnostic. It is distinct from both physical

acoustic energy and MAGNS's learned E_θ used in the port ledger. The radiating system loses physical energy through its boundary, so we read this diagnostic together with field propagation and the learned-energy port balance.

The inverse-mass factor makes the diagnostic particularly sensitive to errors in p, whose true amplitude is much smaller than that of q. In the VALID window $(16, 0, 3)$ of Figure 5, IGNS has time-averaged channel MSEs 0.187 for q and 2.25×10^{-5} for p. At the final step, however, the RMS of p is 9.29×10^{-3} versus 2.34×10^{-4} in truth: its squared amplitude is approximately 1575 times larger. The common-chart energies are then 4.06×10^4 and 24.1. Thus the channel with little weight in unscaled state MSE can dominate this quadratic diagnostic.

The causal IGNS adapter evolves a latent Hamiltonian system and accumulates decoded increments of the observed state. This differs from the original initial-state-offset decoder and can change long-rollout error accumulation. The measured gap therefore concerns the adapted predictor, not the original formulation in isolation. Its readouts do not enforce a balance identity for Eq. (51) or couple that diagnostic to the observed boundary flux. MGC imposes compatibility on the declared mesh-state generator and its ports; this is a different property from balance in latent coordinates.

MGN provides a complementary example: its TEST chart-energy gap is smaller than MeshFT's (2.37 versus 3.27), yet its median field error is over five times larger. A weakly evolving prediction can stay near its initial energy without reproducing wave propagation or radiative loss. We therefore report field error, common-chart trajectories, and the learned-energy port ledger as separate quantities.

Table 10 evaluates each predicted state under the common kinetic-plus-Dirichlet energy used for the reference fields. The final-frame relative error is computed for each window before taking its median. The median of the predicted energy and the median of the truth energy are reported separately.

Table 10: Acoustic common-chart energy at $H = 80$. Each column reports a median over the same 200 windows per split.

Model	Split	Rel. truth	Pred. final E	Truth final E
MAGNS	VALID	1.49	95.6	39
MeshFT	VALID	3.51	164	39
IGNS	VALID	3.92×10^3	1.16×10^5	39
MGN	VALID	2.11	121	39
MAGNS	TEST	1.28	91.2	38.5
MeshFT	TEST	3.27	162	38.5
IGNS	TEST	3.83×10^3	1.19×10^5	38.5
MGN	TEST	2.37	129	38.5

MAGNS has a smaller final chart-energy gap than each baseline on both splits. Its TEST median predicted final energy is 91.2, compared with 162 for MeshFT and 38.5 for truth. Figure 4 shows how these gaps develop through the rollout.

The final MAGNS checkpoint also gives negative mean boundary work in both splits. Its learned-energy balance is reported in Table 5, using the same 200-window evaluations. Together, the field comparison, common-chart diagnostic, and port ledger distinguish propagation accuracy from the balance of the assembled generator.

F BEAM3D: FIELD ACCURACY UNDER STRUCTURED DYNAMICS

The three-seed rollout curves in Figure 7 complement the scalar errors in Table 3: prediction differences develop across the full 20-step trajectory. The field plots below locate these differences in displacement, velocity, and stress. They use seed 0 and the same first VALID trajectory as the earlier visualization.

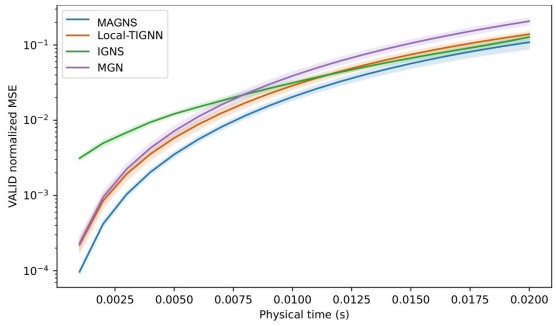

Figure 7: Beam3D VALID closed-loop normalized MSE over time, mean ± standard deviation across three seeds.

F.1 Deformation and stress

MAGNS has lower mean displacement-magnitude and von Mises stress errors than the three baselines at each displayed snapshot of this trajectory. At $t = 20$, its mean absolute displacement error is 0.0502, compared with 0.1527–0.2178 for the baselines; the corresponding stress errors are 392.8 and 1046.4–1904.1 in the recorded units. Figures 8 and 9 show where those differences occur. This field-level view complements the operator-assembly ablation: the predictive advantage is visible in both deformation and internal stress.

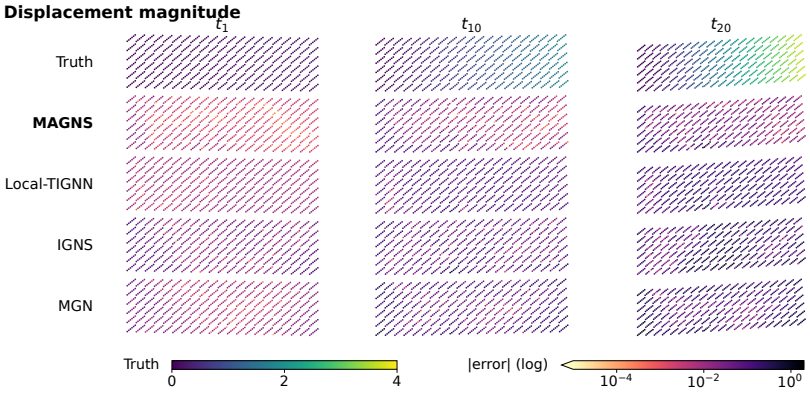

Figure 8: Beam3D displacement magnitude at steps 1, 10, and 20. Top row: reference field. Remaining rows: absolute scalar-field error on the reference-deformed nodes. Models and times share one logarithmic error scale.

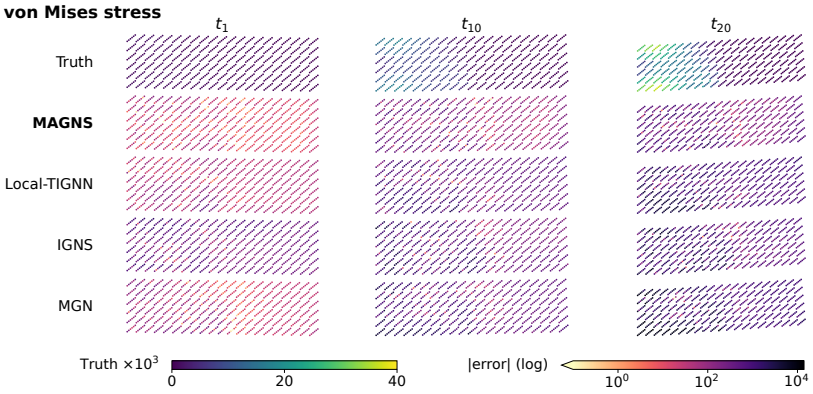

Figure 9: Von Mises stress on the same trajectory and snapshots. The common error scale resolves the spatial distribution of stress errors; the reference field has a separate linear scale.

F.2 Velocity as a complementary observable

Velocity tests a different part of the state from deformation and stress. Figure 10 shows its magnitude and spatial errors on the same trajectory. Local-TIGNN gives smaller speed errors here, while MAGNS's gains are concentrated in displacement and stress. The joint state metric in the main text summarizes these coupled observables across trajectories; the field plots resolve their distinct contributions.

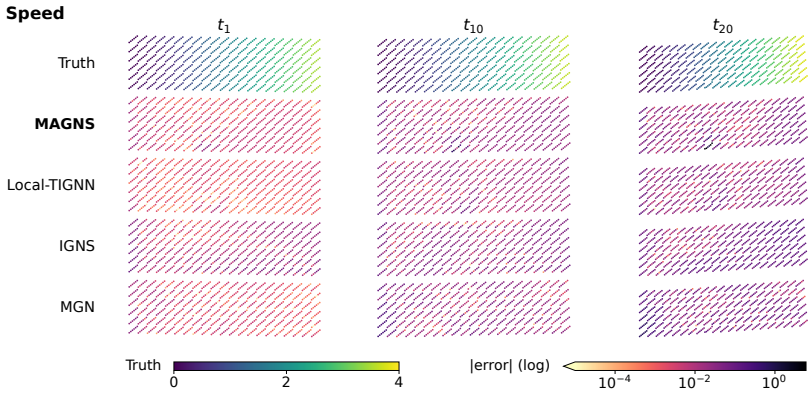

Figure 10: Speed magnitude and absolute scalar-field errors on the trajectory in Figures 8–9. The four models share the same error scale and spatial projection.

G Transient Darcy: control of dissipative flux directions

The Darcy ablation varies the admissible directions of the dissipative generator while holding its learned weights fixed. The crossed comparison separates what the network learns from which directions the assembled dynamics can realize. It extends the fixed-weight intervention in Figure 3.

G.1 Data and training

Both cases use a periodic 32×32 mesh and heterogeneous conductivity. Each A/B pair shares the medium and the RMS of its initial pressure. Their reflected initial conditions differ in whether they excite harmonic flux: A has a nonzero harmonic component, while B is reflection-symmetric with zero harmonic flux. Each case has $32/8/8$ TRAIN/VALID/TEST trajectories and 64 output intervals of length $\Delta t = 0.05$, giving a final time of 3.2.

The observed state is (a, e), where a is cumulative edge flow and e is vertex thermal energy. Pressure follows the storage relation, and $q = \dot{a}$ is the edge block of the assembled generator. MAGNS uses $L = 0$ with learned E, S, and M. The *full-space* declaration, retains $P_{\mathrm{ex}} + P_{\mathrm{coex}} + P_{\mathrm{har}}$. The *harmonic-excluded* declaration, retains $P_{\mathrm{ex}} + P_{\mathrm{coex}}$. This restriction is imposed during dissipative assembly, before integration.

For each case and training declaration, we train seeds $\{0, 1, 2\}$ for 5000 Adam updates with batch size 2, learning rate 10^{-4}, width 32, and two processor blocks in float64. The loss supervises two output intervals using TRAIN-RMS-normalized a, e, p, q errors. Training and evaluation use Heun with eight substeps per output interval. Every evaluation uses the fixed 5000-update checkpoint. Each trained checkpoint is evaluated under both declarations on all VALID and TEST trajectories, for 384 complete rollouts.

G.2 Execution controls circulation independently of training

Let $\varepsilon_q = q_{\mathrm{pred}} - q_{\mathrm{truth}}$. Harmonic and nonharmonic errors are the time averages of $\|P_{\mathrm{har}}\varepsilon_q\|_2$ and $\|(I - P_{\mathrm{har}})\varepsilon_q\|_2$, respectively, over output times $1, \ldots, 64$. The harmonic norm uses the orthonormal two-direction basis of the periodic mesh. We first average over the eight trajectories for each seed, then report the mean and sample standard deviation of the three seed means.

Table 11 separates training from execution. Holding full-space-trained weights fixed, harmonic exclusion changes A's nonharmonic error by 0.46% and B's by -0.002%. It removes B's spurious harmonic flow, but excludes a physical component in A. The reverse intervention recovers a direction excluded during training: allowing harmonic directions at execution reduces A's error for a harmonic-excluded-trained model from 3.80×10^{-3} to 4.78×10^{-4}. Training with those directions available reduces it further to 1.14×10^{-4}.

Table 11: Darcy TEST: crossed training and execution declarations, three-seed means. F denotes full space and X harmonic exclusion. The harmonic column includes sample s.d.; numerically identical seed means are shown without a negligible s.d.

Case	Train / execute	Harmonic error	Nonharmonic error	Four-field NMSE
A	F / F	$(1.14 \pm 0.35) \times 10^{-4}$	3.33×10^{-4}	7.36×10^{-4}
A	F / X	3.80×10^{-3}	3.34×10^{-4}	4.52×10^{-3}
A	X / F	$(4.78 \pm 2.85) \times 10^{-4}$	3.65×10^{-4}	6.14×10^{-4}
A	X / X	3.80×10^{-3}	3.57×10^{-4}	4.38×10^{-3}
B	F / F	$(3.33 \pm 0.59) \times 10^{-5}$	3.34×10^{-4}	8.03×10^{-4}
B	F / X	$(4.03 \pm 0.02) \times 10^{-18}$	3.34×10^{-4}	8.03×10^{-4}
B	X / F	$(2.26 \pm 1.02) \times 10^{-4}$	3.33×10^{-4}	8.15×10^{-4}
B	X / X	$(4.08 \pm 0.02) \times 10^{-18}$	3.31×10^{-4}	8.03×10^{-4}

The four-field NMSE averages the a, e, p, q MSEs after dividing each by its squared TRAIN RMS. Excluding the physical harmonic flow in A increases this score from 7.36×10^{-4} to 4.52×10^{-3} for the same full-space-trained weights. The thermal-field MSE changes by only 0.88%; the main error increase lies in flux and accumulated flow. B's four-field score remains 8.03×10^{-4} under both declarations. Table 12 gives the unnormalized component errors.

Table 12: Darcy TEST component MSEs, averaged over time, trajectories, and seeds. All rows use the checkpoints in Table 11.

Case	Train / execute	a MSE	e MSE	p MSE	q MSE
A	F / F	2.16×10^{-10}	8.49×10^{-4}	5.77×10^{-6}	1.04×10^{-10}
A	F / X	4.63×10^{-8}	8.57×10^{-4}	5.92×10^{-6}	1.21×10^{-8}
A	X / F	7.72×10^{-10}	6.40×10^{-4}	7.70×10^{-6}	2.72×10^{-10}
A	X / X	4.63×10^{-8}	6.68×10^{-4}	7.93×10^{-6}	1.22×10^{-8}
B	F / F	1.24×10^{-10}	9.29×10^{-4}	9.10×10^{-6}	1.10×10^{-10}
B	F / X	1.22×10^{-10}	9.29×10^{-4}	9.10×10^{-6}	1.09×10^{-10}
B	X / F	2.34×10^{-10}	9.32×10^{-4}	9.08×10^{-6}	1.39×10^{-10}
B	X / X	1.21×10^{-10}	9.29×10^{-4}	9.06×10^{-6}	1.07×10^{-10}

Table 13: Darcy VALID: the same crossed evaluation at the fixed training budget. Aggregation matches Table 11.

Case	Train / execute	Harmonic error	Nonharmonic error	Four-field NMSE
A	F / F	$(1.12 \pm 0.47) \times 10^{-4}$	3.00×10^{-4}	7.18×10^{-4}
A	F / X	4.14×10^{-3}	3.00×10^{-4}	5.50×10^{-3}
A	X / F	$(5.02 \pm 3.38) \times 10^{-4}$	3.31×10^{-4}	6.01×10^{-4}
A	X / X	4.14×10^{-3}	3.20×10^{-4}	5.35×10^{-3}
B	F / F	$(2.52 \pm 0.89) \times 10^{-5}$	2.97×10^{-4}	7.83×10^{-4}
B	F / X	$(3.65 \pm 0.05) \times 10^{-18}$	2.97×10^{-4}	7.83×10^{-4}
B	X / F	$(2.26 \pm 1.04) \times 10^{-4}$	2.97×10^{-4}	7.94×10^{-4}
B	X / X	$(3.56 \pm 0.04) \times 10^{-18}$	2.95×10^{-4}	7.83×10^{-4}

G.3 Direction control preserves the dissipative structure

The direction restriction remains active throughout the rollout: the largest harmonic flux norm under harmonic-excluded execution is 1.39×10^{-17}. Table 14 reports extrema across all 384 trajectories. The small $M\,dE$ residual and positive sampled entropy production agree with the dissipative construction. The measured finite-step drift concerns Heun integration of the learned energy. The separately prescribed physical energy is $E_{\mathrm{phys}} = \frac{1}{2}\sum_i h_{0,i} p_i^2 + \sum_i V_i e_i$; its maximum relative drift is 1.67%.

Table 14: Darcy structural and energy diagnostics over all evaluated seeds, cases, declarations, and splits. Maxima are not averages over trajectories.

Quantity	Extremum
$\|M\,dE\|_\infty$	1.77×10^{-15}
Absolute learned-energy rate	1.78×10^{-15}
Absolute mass drift	1.11×10^{-16}
Relative learned-energy drift	2.03×10^{-7}
Relative physical-energy drift	1.67×10^{-2}
Minimum sampled entropy production	6.05×10^{-4}
Harmonic flux norm under exclusion	1.39×10^{-17}

G.4 INCIDENCE-METRIC BASELINE

We also evaluate a MeshFT-style incidence-metric adaptation on the same resistance–heat data. It learns a positive edge metric and advances the pressure-storage system with incidence-based Crank–Nicolson steps, while accumulating flow and heat. It uses the same 5000-update, batch-2, two-interval training budget and TRAIN-RMS loss. This is the resistance–heat adaptation, distinct from the acoustic KDK baseline. Its six trained models complete all 96 VALID and TEST rollouts.

Table 15: Incidence-metric adaptation on the Darcy data, three-seed means. The corresponding MAGNS values are in Tables 11 and 13.

Case	Split	Harmonic error	Nonharmonic error	e MSE	Four-field NMSE
A	VALID	4.20×10^{-3}	4.14×10^{-2}	2.57×10^{-2}	5.35×10^{-1}
A	TEST	3.87×10^{-3}	4.09×10^{-2}	2.85×10^{-2}	5.52×10^{-1}
B	VALID	2.40×10^{-6}	4.09×10^{-2}	3.57×10^{-2}	5.20×10^{-1}
B	TEST	2.14×10^{-6}	4.05×10^{-2}	3.84×10^{-2}	5.38×10^{-1}

The incidence-metric baseline preserves its physical energy to 5.92×10^{-16} relative drift, but has substantially larger field errors. The comparison separates two capabilities: conserving a prescribed scalar and reproducing its associated spatial dynamics. The fixed-weight intervention in Figure 3 further isolates control over global circulation.