Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Non-gradient spatial organisation at tumour–immune interfaces revealed by operator-based analysis

Anas Enoch

anas_nour@um5.ac.ma

Mohammed VI University of Health Sciences (UM6SS), Casablanca, Morocco https://orcid.org/0009-0000-2920-622X

Research Article

Keywords: spatial transcriptomics, Hodge decomposition, non-gradient spatial organisation, tumour microenvironment, operator-based analysis, gradient-model incompatibility, triple-negative breast cancer, coexact component

Posted Date: June 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10028925/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Non-gradient spatial organisation at tumour–immune interfaces revealed by operator-based analysis

Anas Enoch, MD

Mohammed VI University of Health Sciences (UM6SS), Casablanca, Morocco
anas_nour@um5.ac.ma

Abstract

Most spatial transcriptomics analyses model tissue organisation through scalar gene expression fields and their gradients. This paradigm cannot detect a structurally distinct class of spatial organisation: antisymmetric, non-conservative coupling between biological programmes that is orthogonal to the exact subspace under the discrete Hodge inner product.

We introduce a computational operator-based framework that constructs an antisymmetric wedge-flux field between residualized tumour and immune programme scores, decomposes it into gradient (exact) and non-gradient (coexact) components via discrete Hodge operators, and tests whether the coexact component is spatially localized at biologically structured tissue boundaries. Applied to 43 Visium sections from 22 primary triple-negative breast cancer (TNBC) patients (GSE210616 [1]), the coexact component was enriched at tumour–immune interfaces relative to the tumour core in all 40 valid sections (sign test $p < 10^{-12}$; median enrichment 19.5×). The coexact/exact enrichment differential exceeded 1.0 in 38/40 sections (median 2.54; $p < 10^{-9}$), establishing component-specific rather than generic boundary enrichment.

To test whether gradient-compatible diffusion-like models could reproduce this organisation, we trained a conservation-constrained graph neural network (GNN) under gradient-conservation constraints. The learned field retained gradient structure but collapsed the coexact component to numerical noise ($E_{\text{coexact}} \approx 2.7 \times 10^{-12}$), demonstrating operator-class incompatibility with gradient-compatible transport. External replication confirmed coexact interface enrichment in an independent prostate cancer cohort (GSE278936, 21/23 sections, $p < 10^{-5}$) and in a CosMx single-cell cohort (78/96 FOVs, $p = 2.2 \times 10^{-10}$) with protein-level anchoring. KTS analysis revealed reproducible spatial transition-bias structure toward exhaustion-like states, particularly from immune-active and stromal contexts. In the TNBC discovery subset, independent biological marker validation showed enrichment of cytotoxic, CD8, chemokine, and antigen-presentation markers (14–15/19 sections, $p < 0.04$) with a null hypoxia specificity control; three-tier biological validation confirmed that high-coexact nodes and exhaustion-directed transition-bias regions correspond to canonical exhaustion, cytotoxic, and myeloid-suppressive transcriptional programmes.

All code and processed results are publicly available at https://github.com/Anas-Enoch/Hodge_Laplacian_GNN.

Keywords: spatial transcriptomics; Hodge decomposition; non-gradient spatial organisation; tumour microenvironment; operator-based analysis; gradient-model incompatibility; triple-negative breast cancer; coexact component

The framework should be interpreted as an operator-based characterisation of spatial interaction structure rather than as a direct physical transport model of tissue biology.

1 Introduction

Spatial transcriptomics enables genome-wide measurement of gene expression at spatially resolved positions within tissue sections, providing an unprecedented view of tumour microenvironment architecture [2–4]. Current analyses primarily detect co-localization and spatial gradients of biological programs [5–7]. These methods implicitly assume that biologically relevant spatial structure can be represented by scalar potential fields — that the programmes of interest vary smoothly across space in a manner representable by gradient flow.

The Hodge decomposition separates any edge-level interaction field into an exact component and a coexact component. The exact component captures smooth gradient-like organisation — fields in which programme scores vary continuously from high to low across a spatial boundary. The coexact component captures residual spatial interaction structure that cannot be represented as a locally smooth potential field, without implying literal physical circulation or molecular transport. In this work, coexact enrichment should not be interpreted as literal physical circulation, but as an operator-defined signature of non-gradient spatial organisation at the tumour–immune interface.

Why symmetric association methods are structurally insufficient. Symmetric spatial statistics (covariance, Moran's I, cosine similarity, neighbourhood enrichment) measure *co-variation*: how much two programme fields move together. They are, by construction, invariant to sign. An interaction in which programme A is high where programme B is low — a spatial confrontation rather than a co-localisation — produces the same symmetric statistic as the reverse arrangement. More fundamentally, symmetric bilinear operators define even-valued functions on undirected edges and lie entirely in the gradient (exact) subspace of the Hodge decomposition; they yield zero coexact projection by algebraic necessity. Symmetric spatial statistics may detect clustering, segregation, or co-localization, but they do not encode oriented antisymmetric edge interactions and therefore do not directly represent the coexact component studied here.

The Hodge framework. By the discrete Hodge decomposition theorem [8, 9], any edge field on a spatial graph decomposes uniquely into a gradient (exact) component derivable from a scalar potential, a non-gradient (coexact) component that *cannot* be expressed as any gradient, and a harmonic residual. Whether the interaction field at tumour–immune boundaries carries significant coexact content — and whether this content is specifically localized at biologically structured boundaries — is a question that symmetric or gradient-based methods are structurally unable to address.

We introduce an operator-based framework that constructs an antisymmetric interaction field between residualized programme scores, decomposes it via the Hodge operator, and tests interface localization and gradient-model compatibility as a coherent falsification pipeline: (1) construct the antisymmetric wedge flux; (2) isolate the coexact component; (3) test whether it concentrates at biologically structured interfaces; (4) test whether conservation-constrained gradient-compatible transport can reproduce it. The framework provides a falsifiable computational test: a non-zero interface-specific coexact component indicates that the interaction field is not expressible as the gradient of any scalar potential, not that a specific physical process is occurring.

2 Materials and Methods

2.1 Spatial graph and programme scores

For each Visium section, a k-nearest-neighbour spatial graph was constructed from spot centroid coordinates ($k = 6$, reflecting hexagonal array geometry). Tumour and immune programme scores were computed as mean normalized log-expression over curated marker gene sets [10–12]. Spots were annotated to tumour core, invasive margin, stromal, and immune-rich regions using score thresholds. Sections were excluded from the primary interface-enrichment analysis if they had fewer than 30 tumour–immune boundary edges (three sections); secondary robustness analyses used stricter operator-specific inclusion criteria, including at least 10 annotated nodes per required region.

2.2 Antisymmetric wedge-flux construction

Let a_i and b_i denote tumour and immune programme scores at node i. Each field was residualized against its local neighbourhood mean: $\tilde{a}_i = a_i - \bar{a}_{N(i)}$, $\tilde{b}_i = b_i - \bar{b}_{N(i)}$, removing global spatial trends that could induce spurious coexact structure. The antisymmetric wedge flux along oriented edge $(i \to j)$ is

$$\omega_{ij} = \frac{\tilde{a}_i \tilde{b}_j - \tilde{a}_j \tilde{b}_i}{\ell_{ij}}, \tag{1}$$

where ℓ_{ij} is the spatial edge length. The wedge is the unique (up to scale) local bilinear operator that vanishes under spatial alignment ($\tilde{a} \propto \tilde{b}$) and is maximized under local programme opposition. Symmetric alternatives (covariance, cosine similarity) yield zero coexact projection by construction; the wedge is the minimal operator capable of detecting non-gradient interaction structure.

This field is not a measurement of physical molecular flux. It is a spatial operator probe: it assigns a signed magnitude to each spatial edge representing the antisymmetric contrast between the two programmes at that edge. Where programmes are spatially aligned, the wedge is zero. Where one programme is systematically high while the other is low — the local confrontation geometry at a tissue boundary — the wedge is large and antisymmetric. The Hodge decomposition then asks how much of this antisymmetric contrast field is gradient-compatible.

2.3 Discrete Hodge decomposition

Let B_1 and B_2 denote the node–edge and edge–face incidence matrices of the spatial cell complex. The wedge flux vector f decomposes as

$$f = \underbrace{B_1^\top \alpha}_{\text{exact (gradient)}} + \underbrace{B_2^\top \beta}_{\text{coexact (non-gradient)}} + \gamma, \tag{2}$$

where α is solved from $\Delta_0 \alpha = B_1 f$ (conjugate gradient) and γ is the harmonic residual. The exact component represents the part of the interaction contrast that *can* be explained by a scalar potential field: a smooth gradient of programme differences. The coexact component captures the component of the interaction field that is not representable within the exact (gradient-derived) subspace.

The coexact fraction $\rho_{\text{coexact}} = \|f_{\text{coexact}}\|^2 / \|f\|^2$ quantifies how much of the interaction field lies outside the gradient subspace. Node-level coexact density $c_i = (1/\deg_i) \sum_{e \ni i} |f_{\text{coexact}}(e)|$ localizes this non-gradient

structure spatially. Full derivations and implementation details are in Supplementary Methods Section S1.

2.4 Interface enrichment testing

The interface-to-tumour enrichment ratio $R = \text{mean}(c_i)_{\text{interface}}/\text{mean}(c_i)_{\text{core}}$ was computed per section. A ratio $R > 1$ indicates that non-gradient interaction intensity is elevated at the biologically structured boundary relative to the homogeneous tumour core. Statistical significance used 1,000 label-permutation nulls (shuffling region labels while preserving graph topology and flux values). Cohort-level significance used one-sided sign tests ($H_0 : R \leq 1$).

2.5 Gradient-constrained null model

Why a gradient-null model is informative. The PDE-constrained GNN is not intended as a realistic simulator of tissue biology or physical transport. It is a falsification-oriented null model restricted to diffusion-like and smooth spatial organisation: the operator class it can represent is limited to fields consistent with gradient-compatible diffusion-like and integral flux conservation. If the observed interface organisation collapses to near-chance recovery under this constrained operator class while gradient structure remains recoverable, this indicates an incompatibility between the observed signal and gradient-compatible smoothing alone — not that the tissue is literally transporting molecules in a manner that is incompatible with gradient-compatible conservative dynamics. The model tests whether the coexact component of the wedge-flux field is recoverable within a restricted class of smooth, gradient-compatible and conservation-constrained operators; the result that it is not supports the operator-based coexact enrichment claim without implying direct physical transport inference.

A Hodge–Laplacian GNN was trained to reproduce the wedge-flux field under conservation and smoothness constraints encoding gradient-compatible diffusion-like dynamics [13, 14]:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda_{\text{PDE}}\mathcal{L}_{\text{ADR}} + \lambda_{\text{Stokes}}\mathcal{L}_{\text{integral}}, \tag{3}$$

where $\mathcal{L}_{\text{integral}}$ enforces Green–Stokes flux conservation over held-out subcomplex boundaries.

This GNN is not a generative biological model. It is a structured null that tests *operator incompatibility*, not biological transport: it learns the best possible approximation of the observed interaction field using only dynamics that are consistent with gradient-dominated diffusion-like transport. If such dynamics were adequate, the GNN should reproduce the observed field faithfully, including its coexact component. If the gradient-compatible model class is structurally incompatible with the coexact structure in the observed field, the GNN will recover gradient-dominated dynamics and collapse the coexact component — not because of optimization failure, but because conservation constraints force convergence to the best-fitting solution within the implemented gradient-compatible model class. Full architecture and training details are in Supplementary Methods Section S2.

2.6 External validation cohorts

Spatial Hallmarks pan-cancer cohort. External pan-cancer validation was performed using the Spatial Hallmarks spatial transcriptomics resource [15], comprising 26 Visium sections across six cancer types (breast, colorectal, glioblastoma, lung, ovarian, and prostate) from the 10x Genomics public dataset collection associated with the study "Tumour evolution and microenvironment interactions in 2D and 3D space."

Data are publicly available at https://doi.org/10.5281/zenodo.14044964. The identical operator pipeline (wedge-flux construction, Hodge decomposition, coexact enrichment testing) was applied without cohort-specific tuning.

Prostate cancer (GSE278936). The core pipeline was applied unchanged to 23 Visium sections from GSE278936 (benign, treatment-naïve, neoadjuvant, and CRPC) using the same circularity-constrained gene panels as the TNBC analysis.

CosMx single-cell replication. The pipeline was adapted to the NanoString CosMx Breast Multiomic dataset (108 FOVs, $\sim$10 μm per cell) using per-FOV z-scored log1p normalization, Delaunay spatial graphs, restricted permutation testing, and protein anchoring with simultaneous PanCK/CD45/CD68 measurements.

2.7 Independent biological validation of spatial hallmarks

Biological validation was performed independently of wedge-flux construction and operator decomposition in order to test whether operator-defined spatial hallmarks correspond to canonical tumour immunobiology. Validation analyses were conducted on the spatial hallmarks cohort using predefined cytotoxic, exhaustion, stromal activation, myeloid-suppressive, interferon-signalling, and hypoxia-control transcriptional modules. None of the validation genes were used in construction of the wedge-flux operators, coexact decomposition, spectral diagnostics, or KTS transition-bias inference.

Tier 1 (module-score correlation). For each section, Spearman correlations were computed between per-node coexact density and module scores at interface-associated nodes. Cytotoxic modules included CD8A, CD8B, GZMB, PRF1, GNLY, IFNG, and NKG7; exhaustion modules included PDCD1, CTLA4, LAG3, HAVCR2, TIGIT, TOX, and CXCL13; stromal activation modules included FAP, POSTN, COL1A1, COL3A1, TGFB1, and CXCL12; myeloid-suppressive modules included SPP1, C1QA, MRC1, CD163, TREM2, and APOE. Hypoxia modules (HIF1A, VEGFA, LDHA) were included as specificity controls. Cohort-level consistency was assessed using one-sided sign tests.

Tier 2 (KTS exhaustion endpoints). To validate KTS exhaustion endpoints biologically, target nodes associated with IMMUNE_EXHAUSTED operator states were compared against non-exhausted immune-associated nodes within each section. Expression enrichment ratios were computed for canonical checkpoint and exhaustion markers (PDCD1, CTLA4, LAG3, HAVCR2, TIGIT, TOX, CXCL13), and cohort-level significance was assessed using one-sided sign tests.

Tier 3 (exhaustion-associated stromal context). To test whether exhaustion-associated transition-bias structure preferentially occurred within reactive stromal environments, stromal-source regions linked to exhaustion-associated KTS endpoints were compared against stromal-source regions lacking exhaustion-associated endpoints. Enrichment analyses were performed for reactive stromal markers including FAP, TGFB1, CXCL12, COL1A1, and COL3A1.

Statistical analysis for validation tiers. Statistical interpretation emphasized cross-section consistency and effect-size reproducibility rather than isolated nominal p-values. Cohort-level enrichment consistency was evaluated using one-sided sign tests. These analyses should be interpreted as biologically grounded

operator-validation procedures rather than causal or temporal trajectory inference. All biological validation and baseline-comparison analyses were performed on the Spatial Hallmarks cohort using implementations that are fully independent of the TNBC discovery pipeline; full reproducibility details are provided in the Code Availability section.

2.8 Empirical baseline comparison

To test whether the coexact fraction and KTS exhaustion bias are reducible to standard spatial analysis methods, five baseline families were computed on each section. (1) Spatial autocorrelation: Moran's I of the programme-contrast field $|T_i - I_i|$ on the full-section kNN graph ($k = 6$), with the interface signal zero-padded for non-interface nodes. Full-section computation ensures $W \approx 2kn$, keeping $n/W \approx 1/k$ and Moran's I bounded in $[-1, 1]$. (2) Boundary DE score: median $\log_2$ fold-change of immune effector and tumour proliferation markers (CD8A, CD8B, GZMB, PRF1, IFNG, NKG7, PDCD1, LAG3, HAVCR2, TIGIT; counterbalanced against EPCAM, KRT8, KRT18, MKI67, TOP2A) comparing interface-associated versus tumour-core nodes. (3) Ligand–receptor proximity score: mean product $L_i \times R_j$ for ten curated tumour–immune LR pairs (TGFB1:TGFBR2, CXCL12:CXCR4, CXCL9:CXCR3, PDCD1:CD274, IL6:IL6R, CCL5:CCR5, TNF:TNFRSF1A, IFNG:IFNGR1, VEGFA:KDR, FN1:ITGB1) over interface-adjacent edges. (4) Graph Laplacian smoothness: $x^\top Lx / x^\top x$ for the programme-contrast field on the interface subgraph; measures whether coexact enrichment is reducible to generic boundary roughness. (5) Neighbourhood-enrichment proxy: tumour–immune adjacency frequency relative to expected frequency under a node-label shuffle null, analogous to Squidpy's `gr.nhood_enrichment`. These analyses provide proxy-level non-redundancy controls rather than full benchmarking against complete external software frameworks such as CellChat, COMMOT, Squidpy, SpatialDE, or SPARK-X. Spearman correlations between coexact fraction/KTS exhaustion bias and each baseline were computed across the 26 Spatial Hallmarks sections. A pre-defined non-redundancy threshold of $|\rho| = 0.40$ was used; correlations below this threshold indicate that the operator metric captures information not contained in the respective baseline. Implementation: `baseline_comparison.py`.

2.9 Operator robustness and statistical analysis

Five antisymmetric operators were applied in decreasing information order (see Supplementary Methods Section S8). All cohort-level comparisons used one-sided sign tests (binomial, $p = 0.5$). Region-level comparisons used Mann–Whitney U tests with permutation nulls.

Multiple-comparison correction policy. Primary confirmatory analyses—the Step 06 coexact interface enrichment sign test and the energy-matched Step 10 spectral null test—are treated as the primary endpoints and reported without adjustment. For related secondary analysis families, Benjamini–Hochberg false-discovery-rate (FDR) correction was applied within each family at $q < 0.05$; raw p-values and q-values are both reported. Testing families are defined as follows: (i) *Biological validation module correlations*: 6 module-score Spearman tests per section (exhaustion, cytotoxic, stromal, myeloid-suppressive, IFN-signalling, hypoxia control); (ii) *Exhaustion endpoint markers*: 7 marker enrichment tests per section (PDCD1, CTLA4, LAG3, HAVCR2, TIGIT, TOX, CXCL13); (iii) *Stromal mediation markers*: 5 stromal marker tests per section (TGFB1, FAP, CXCL12, COL1A1, COL3A1); (iv) *Baseline comparison*: 5 correlation tests (Moran's I, boundary DE, LR proxy, graph smoothness, NE proxy); (v) *Spectral diagnostics*: per-section Zeta spectral concentration and robustness statistics. Remaining exploratory analyses are inter-

preted using effect-size consistency and cross-dataset reproducibility rather than isolated nominal p-values. Fixed random seeds are documented in the repository.

3 Results

The evidence in this section is organised as a strict hierarchy. The *primary* finding is coexact interface enrichment (Section 3.2). The *primary falsification* is that a gradient-compatible null model cannot reproduce the coexact structure (Section 3.3). *Replication* is established in the prostate and CosMx cohorts and across six cancer types (Section 3.5). *Biological grounding* is provided by cytotoxic, exhaustion, stromal, and myeloid programme alignment (Section 3.7). The remaining analyses—zeta spectral diagnostics, curl maps, graph-KS instability, stochastic Hodge Bayes factors, and NCG commutator grounding—are *supportive and exploratory* robustness diagnostics (Supplementary Material) and should not be read as competing with the central claim.

3.1 Cohort accounting and analysis inclusion

Different analyses impose distinct graph-construction, annotation, and quality-control requirements, which leads to different section, patient, or field-of-view denominators across the study. Table 1 reports the number of units included in each analysis together with the exclusion criterion that defines each denominator.

Table 1: **Analysis inclusion accounting.** Units included in each analysis and the criterion defining each denominator. TNBC discovery cohort: GSE210616. Prostate: GSE278936. CosMx: NanoString Breast Multiomic. Pan-cancer: Spatial Hallmarks. HCC: Wu et al. (Zenodo v2).

Analysis	Units	Inclusion / exclusion criterion
TNBC sections (total available)	43 sections / 22 patients	All Visium sections in GSE210616.
TNBC coexact interface enrichment	40 sections / 22 patients	3 sections excluded for insufficient interface spots after QC (<30 tumour–immune boundary edges).
NCG commutator / operator robustness	18 sections	Restricted to sections with both programmes scored above the activity threshold in $\geq$2 spatial neighbourhoods.
Zeta spectral / stochastic Hodge Bayes	19 / 43 sections	Label-null spectral conditioning: $n = 19$ (non-degenerate edge Laplacian); energy-matched interface-density null evaluated on all $n = 43$ sections. Restricted to sections meeting spectral-conditioning requirements (non-degenerate edge Laplacian).
Prostate replication	23 sections	All passing-QC sections in GSE278936.
CosMx single-cell replication	96 / 108 FOVs	12 FOVs excluded for <200 cells or absent tumour/immune annotation.
Pan-cancer non-redundancy	26 sections / 6 cancer types	All Spatial Hallmarks sections with harmonised programme scores.
HCC stress test	15 / 22 sections / 11 patients	7 sections excluded for failing interface-heuristic validity on therapy-remodelled tissue.

3.2 Coexact organisation localizes at tumour–immune interfaces across the TNBC cohort

Across 40 valid sections from 22 primary TNBC patients, the coexact component of the tumour–immune wedge flux was enriched at the tumour–immune interface relative to the tumour core in every section (sign

test $p < 10^{-12}$; Figure 3; Table 2). The median enrichment ratio was 19.5× (range 4.6–44.3×). The coexact/exact differential exceeded 1.0 in 38/40 sections ($p < 10^{-9}$; median 2.54), establishing that the enrichment is component-specific: the interface is not merely energetically active, but specifically enriched for the non-gradient part of the interaction field. All 22 patients contributed at least one section with differential > 1.0 (patient-level $p < 0.001$); no significant association was found with race, neoadjuvant chemotherapy, age, or relapse-free survival (all $p > 0.18$). Global whole-section coexact fraction was not elevated under the permutation null ($p \approx 0.5$), confirming that the signal is interface-restricted rather than tissue-wide.

Table 2: **Cohort-level enrichment statistics across 40 valid TNBC sections.** The coexact enrichment ratio is the permutation-tested ratio of mean node-level coexact energy at interface-like nodes to tumour-enriched nodes. The differential is the ratio of coexact to exact enrichment. Δ_{coexact} is the difference in aggregate coexact energy fraction between interface and tumour regions. Sign tests are one-sided binomial under $p = 0.5$.

Statistic	**Value**	**Sign test p**
Coexact enrichment > 1.0	40/40 sections	$< 10^{-12}$
Differential > 1.0	38/40 sections	$< 10^{-9}$
$\Delta_{\text{coexact}} > 0$	22/22 patients	< 0.001
Median coexact enrichment	19.5×	—
Median differential	2.54	—
Median Δ_{coexact}	+0.18	—
Differential range	0.79–6.11	—

Biological interpretation of interface coexact organisation. The localized coexact component is not a geometric artifact of the boundary annotation. Spatially, it concentrates at nodes where the tumour and immune programme fields are in local confrontation: the interface region where each programme is simultaneously elevated and opposed. At these nodes, the wedge field is large and the gradient model fails — there is no scalar potential whose gradient could represent the directional interaction contrast observed. This pattern is consistent with a distinct spatial interaction regime at the tissue boundary: a zone of locally enriched and spatially restricted, finite-range non-gradient organisation that is structurally different from the gradient-dominated interior of the tumour core. Within-interface biological anchoring confirmed this interpretation: proxy coexact energy correlated positively with all three residualized programme types at interface nodes (tumour residual: median $\rho_{\text{int}} = 0.349$, 17/18, $p = 7.2 \times 10^{-5}$; immune residual: 16/18; stromal residual: 14/18; Figure 4), indicating that the coexact signal indexes the intensity of programme contrast at boundary locations. NCG commutator grounding across 18 sections (Supplementary Table S1) confirmed that coexact energy is consistent with the geometric image of operator non-commutativity (under the commutator construction defined in Section S4) specifically at interface nodes ($\rho = 0.832$, $\Delta\rho = +0.198$, both 18/18, $p < 10^{-5}$), ruling out a construction-artifact interpretation.

3.3 Conservative null learning collapses the coexact component

The conservation-constrained GNN was trained stably on the proxy flux field (Figure 7). After convergence, the learned field retained gradient structure but the coexact component collapsed to numerical noise: $E_{\text{exact}} \approx E_{\text{total}}$, $E_{\text{coexact}} \approx 2.7 \times 10^{-12}$ (Figure 5). Mean curl in the learned field was 4.29×10^{-8}, versus coexact fraction $\approx 73\%$ in the proxy field. Only a weak residual coexact signal survived conservation enforcement in

immune-enriched regions (Mann–Whitney $p \approx 0.028$, permutation $p \approx 0.019$).

Falsification interpretation. The gradient component of the proxy field survived gradient-model training; the coexact component did not. This dissociation is not an optimization artifact: training diagnostics confirmed stable convergence of all loss components. It is a structural incompatibility: the gradient-compatible model class (the implemented conservative diffusion-like model class) can recover the gradient structure of the observed field, but is *unable* to represent the coexact component because that component lies, by definition, outside the space of gradient fields. The collapse therefore demonstrates incompatibility with the conservative gradient null model: the observed coexact component was not recoverable under the implemented conservation-constrained gradient-compatible null model, despite preservation of the exact component.

A hybrid transport representation $f_{\text{obs}} = -\nabla\phi + \chi_{\text{int}}\nabla^{\perp}\psi$ identifies the mismatch precisely (Figure 6): the gradient component $-\nabla\phi$ is recovered by the GNN; the interface-confined coexact component $\chi_{\text{int}}\nabla^{\perp}\psi$ is present in the proxy field and absent in the learned field. Bayesian model comparison confirms this at the cohort level (Supplementary Section S6; Supplementary Table S3; $\log B_{1a/0} = +45.95$; $\log B_{1b/1a} = +517.6$; both 19/19, $p < 10^{-5}$).

3.4 Coexact organisation is locally structured, not globally mode-governed

Interface coexact organisation is locally structured: spatial correlations in the coexact density field decay rapidly across the tissue section, the dominant spectral structure is concentrated in low-frequency but not exclusively interface-specific modes, and there is no evidence for global oscillatory periodicity.

These spectral and rotational diagnostics are exploratory higher-order analyses intended to probe instability-like and orientation-related structure rather than establish definitive mechanistic organisation.

Specifically, the normalized Zeta spectral diagnostic (Supplementary Methods S5; Supplementary Table S2) showed that the global coexact signal is non-flat — concentrated in low-frequency eigenmodes relative to a label-permutation null (flow = 0.653, 19/19, $p < 10^{-5}$; $Z(1) = 1.89$, 18/19) — but is *not* more spectrally organized at tumour–immune interfaces than at equally energetic regions elsewhere (energy-matched null: 2/43 TNBC sections significant; sign test $p = 1.0$). Face-level curl showed balanced signed orientation, with no dominant clockwise or counter-clockwise direction in the examined sections (Supplementary Methods Section S9.3).

This pattern is consistent with an interface organisation that is locally coordinated within a finite spatial range but does not extend into globally periodic or spectrally coherent modes. The tumour–immune boundary appears as a zone of locally structured non-gradient interaction — a spatially restricted regime in which the confrontation geometry between programmes generates coexact organisation that is confined to the boundary and does not propagate into the surrounding tissue as a coherent spatial wave. Interface coexact organisation is therefore best interpreted as a boundary-confined and spatially restricted interaction structure, not as the expression of a globally dominant spatial mode.

3.5 External replication across independent cohorts and platforms

Coexact interface enrichment replicated in 21/23 prostate cancer sections (GSE278936; $p < 10^{-5}$; median $R = 4.9\times$) across four disease conditions, confirming that the organisation is not TNBC-specific. Spectral and rotational diagnostics did not show interface-specific enrichment after energy normalization in either

cohort (sign test $p = 1.0$; Supplementary Methods Section S5), consistent with the locally-structured interpretation above. Coexact-flux-weighted spatial transition-bias analysis in GSE278936 revealed a spatial transition-bias toward exhaustion-associated states in 21–23/23 sections per source compartment (bias 1.9–3.6×; Supplementary Table S4; Supplementary Table S5 provides a multi-layer overview), orthogonal to the spatial enrichment layers.

In the CosMx single-cell cohort, coexact enrichment was detected in 78/96 FOVs ($p = 2.2\times10^{-10}$; median $R = 2.09$). RNA-derived coexact energy co-localized with protein-defined tumour–immune contact zones (PanCK/CD45; 75/108 FOVs; $p < 10^{-5}$), and RNA–protein boundary coherence predicted enrichment significance ($r = 0.41$, $p < 0.001$). The 6-fold attenuation in enrichment ratio at single-cell versus spot resolution is consistent with per-cell score noise attenuating the antisymmetric edge contrast; the enrichment direction was preserved across resolutions and platforms.

3.6 High non-commutativity nodes are enriched for immune effector markers

Nodes with the highest non-commutativity norm (top 10%) were enriched for immune-related marker sets not used in wedge construction (Table 3): cytotoxic T cell activity (15/19, $p = 0.010$; median ratio 1.41), CD8 T cells (14/19, $p = 0.032$), chemokine recruitment (14/19, $p = 0.032$), antigen presentation (15/19, $p = 0.010$), and M2 macrophage polarity (15/19, $p = 0.010$). The hypoxia specificity control showed no enrichment (6/19; median 0.985; $p = 0.97$), excluding generic spatial heterogeneity as the driver.

Table 3: **Independent biological marker enrichment in top-10% NC nodes, across 19 TNBC sections.** For each section, the enrichment ratio = mean(marker) in top-10% NC nodes / mean(marker) all nodes. Sign test: H_0: enrichment ratio ≤ 1.0, one-sided binomial. Five immune-related marker sets are consistently enriched; hypoxia (specificity control) is not, confirming that the NC norm identifies immune-specific biological activity rather than generic spatial heterogeneity. All validation markers are outside the gene sets used to construct the wedge flux.

Marker set	**Genes (representative)**	**Median ratio**	**Sections** > 1.0	**Sign test** p	**Result**
Cytotoxic activity	GZMB, PRF1, GNLY, IFNG	1.41	15/19	0.010	enriched
CD8 T cells	CD8A, CD8B	1.31	14/19	0.032	enriched
Chemokine recruit	CXCL12, CCL5, CCL2	1.17	14/19	0.032	enriched
Antigen present	HLA-DRA, HLA-DRB1, B2M	1.09	15/19	0.010	enriched
M2 macrophage	CD163, MRC1	1.19	15/19	0.010	enriched
Hypoxia (control)	HIF1A, VEGFA, LDHA	0.985	6/19	0.968	**null**

3.7 Independent biological validation of operator-defined spatial hallmarks

To test whether operator-defined spatial hallmarks correspond to canonical tumour immunobiology rather than abstract graph structure, we performed three-tier biological validation across the spatial hallmarks cohort (a fourth tier, explicit immune-active-to-exhaustion transition-bias validation, remains a planned extension). Validation was performed independently of wedge-flux construction, coexact decomposition, spectral diagnostics, and KTS state assignment, preventing circular interpretation between operator construction and downstream biological grounding.

Tier 1: Module-score correlation. At interface nodes, coexact density correlated most strongly with cytotoxic immune programmes, with positive correlations observed in 25/26 sections (median Spearman

$\rho = 0.240$, sign-test $p < 10^{-4}$). Weaker but significant positive associations were also observed for exhaustion-associated programmes (19/26 sections; median $\rho = 0.063$; sign-test $p = 0.0145$) and myeloid-suppressive programmes (18/26 sections; median $\rho = 0.066$; sign-test $p = 0.038$). Stromal activation signatures showed a positive trend that did not reach significance (17/26 sections; median $\rho = 0.078$; $p = 0.084$), consistent with stromal contribution to coexact structure that is secondary to cytotoxic engagement. The hypoxia control module showed no enrichment (15/26 sections; median $\rho = 0.021$; sign-test $p = 0.279$), demonstrating that coexact enrichment reflects biologically specific tumour–immune confrontation structure rather than generic spatial heterogeneity.

Tier 2: KTS exhaustion-endpoint marker enrichment. KTS exhaustion-endpoint target nodes were strongly enriched for canonical exhaustion and checkpoint markers across all analysed sections. Median enrichment ratios were 4.75× for PDCD1, 4.44× for CTLA4, 3.87× for LAG3, 3.40× for HAVCR2, 4.73× for TIGIT, 3.36× for TOX, and 3.88× for CXCL13. All seven markers were enriched in 26/26 sections (sign-test $p = 1.49 \times 10^{-8}$ for each). These results demonstrate that operator-defined exhaustion endpoints correspond closely to established exhaustion-associated transcriptional programmes rather than abstract state labels.

Tier 3: Exhaustion-associated stromal context. Reactive stromal programmes preferentially associated with exhaustion-directed KTS spatial transition-bias structure. Stromal-source regions linked to exhaustion-associated endpoints showed significant enrichment for TGFB1 (24/26 sections; median ratio 1.13×; $p < 10^{-4}$), FAP (22/26 sections; median 1.09×; $p = 3 \times 10^{-4}$), CXCL12 (20/26 sections; median 1.08×; $p = 0.005$), COL1A1 (21/26 sections; $p = 0.001$), and COL3A1 (19/26 sections; $p = 0.015$). POSTN showed no consistent enrichment (16/26; $p = 0.164$). These findings are consistent with exhaustion-associated transition-bias structure preferentially emerging within reactive stromal microenvironments characterised by TGF-β signalling and cancer-associated fibroblast activity.

Tier 4 — planned extension. Validation of explicit immune-active-to-exhaustion-associated transition-bias structure requires denser edge-resolved transition graphs and remains a planned extension of the framework.

3.8 Comparison against standard spatial analysis baselines

The goal of baseline comparison is non-redundancy testing rather than superiority benchmarking: we ask whether the coexact operator captures spatial interaction structure that is not already captured by established methods, not whether it outperforms them on a specific predictive task. We present two layers of evidence: (i) proxy-level controls on the TNBC discovery cohort, and (ii) a package-level and equivalent baseline benchmark using software implementations on two independent cohorts.

Proxy-level non-redundancy on the Spatial Hallmarks pan-cancer cohort. Five baseline families were computed on all 26 Spatial Hallmarks sections using the same programme scores and interface labels as the primary analysis: (1) Moran's I of the programme-contrast field ($\rho = -0.187$); (2) boundary differential-expression sharpness ($\rho = -0.236$); (3) ligand–receptor proximity ($\rho = -0.360$); (4) graph Laplacian smoothness ($\rho = +0.045$); and (5) neighbourhood-enrichment proxy ($\rho = +0.623$, partial overlap). The negative correlations with Moran's I, boundary DE, and ligand–receptor proximity indicate that

sections with higher coexact fraction tend to have lower spatial autocorrelation and lower boundary-DE/LR scores—consistent with the coexact component specifically capturing non-gradient, non-smooth interface organisation orthogonal to what these methods detect. The partial overlap with neighbourhood enrichment is biologically expected: denser tumour–immune boundaries provide more interface subgraph structure; the two quantities are not interchangeable [16]. KTS exhaustion bias was independent of Moran's I ($\rho = -0.28$, $p = 0.169$) and boundary DE ($\rho = -0.37$, $p = 0.062$), with borderline overlap with the LR proxy ($\rho = 0.42$, $p = 0.032$). Implementation: `spatial_hallmark/baseline_comparison.py`.

Package-level and equivalent baseline benchmark: pan-cancer Spatial Hallmarks cohort. To extend the proxy-level controls to complete external software implementations, we ran five baselines (official package implementations for Squidpy NE, Moran's I, SPARK-X, and SpatialDE; an equivalent implementation for COMMOT-style LR proximity) on the same 26-section Spatial Hallmarks pan-cancer cohort under identical programme scores and interface labels: Squidpy neighbourhood enrichment [17] (`squidpy.gr.nhood_enrichment`, 1,000 permutations); Moran's I via `esda.Moran` [18]; a SPARK-X-equivalent HSIC test [19]; a SpatialDE-equivalent GP variance fraction [20]; and a COMMOT-style distance-weighted ligand–receptor proximity score [21]. Section-level scores were compared against the Hodge coexact interface enrichment ratio using Spearman correlation; spot-level scores were used to compute LOO AUC for interface-vs-tumour-core discrimination; top-10% hotspot Jaccard overlap was computed against the coexact hotspot map; and biological endpoint alignment was assessed as interface/background expression ratio for exhaustion and cytotoxic markers (Figure 1; Table 4). A conceptual overview of the operator-based workflow and its contrast with standard spatial statistics is shown in Figure 2.

The purpose of baseline comparison is non-redundancy assessment rather than superiority benchmarking. The evaluated methods represent distinct classes of spatial analysis (autocorrelation, neighbourhood structure, ligand–receptor proximity, and spatial variability), each capturing different aspects of tissue organisation. Across 26 pan-cancer sections from six cancer types, representative implementations from five baseline method families were executed, including official package implementations where applicable and equivalent implementations where direct execution was not feasible, revealing a three-way correlation structure with the coexact interface enrichment ratio. Two methods showed significant *negative* correlations: Squidpy neighbourhood enrichment ($\rho = -0.660$, $p = 0.001$) and COMMOT-style ligand–receptor proximity ($\rho = -0.650$, $p < 0.001$). One method showed a significant *positive* correlation: Moran's I spatial autocorrelation of the immune programme ($\rho = +0.497$, $p = 0.010$). (This package-level Moran's I is computed on the immune programme score, whereas the proxy-level control above is computed on the signed programme-contrast field, $A - B$; the two measure autocorrelation of different quantities and are therefore not directly comparable in sign.) Two methods were orthogonal: SPARK-X spatial variability ($\rho = -0.056$, $p = 0.784$) and SpatialDE spatial variance fraction ($\rho = +0.248$, $p = 0.222$). This structure is mechanistically coherent. The two negative correlations converge on the same interpretation: neighbourhood-enrichment z-scores were strongly negative across sections (indicating tumour–immune spatial segregation rather than intermixing), and the more segregated the boundary, the higher the coexact enrichment; likewise, sections where tumour and T-cells are most diffusely co-localised (high LR proximity) have the *lowest* coexact enrichment. Sharp, segregated tumour–immune boundaries therefore produce high coexact organisation, whereas diffuse intermixing produces low coexact organisation — two independent adjacency- and co-expression-based methods converging on the same structural anti-complementarity. The positive Moran's I correlation reflects shared sensitivity to immune spatial clustering: both Moran's I and the coexact operator respond to immune spatial

organisation. However, $\rho = +0.497$ corresponds to only $\rho^2 \approx 25\%$ shared rank variance, leaving 75% of the cross-section ordering unexplained by scalar autocorrelation, and the decisive evidence of non-reducibility comes from interface discrimination (below). Interface-vs-tumour-core discrimination confirmed that no baseline matched the coexact operator (LOO AUC 0.65). Moran's I reached AUC 0.56 — above chance, the highest of any baseline, consistent with its positive cross-section correlation, but well below the coexact operator; SPARK-X and SpatialDE reached exactly 0.50; and LR proximity reached 0.43 (below chance), confirming at the spot level that LR-dense locations are not the coexact boundary. The coexact operator thus captures the immune-clustering signal that Moran's I partially shares, plus additional non-gradient interface geometry that no scalar, adjacency-based, variability-based, or co-expression-based method localises. Top-10% hotspot Jaccard: T-cell score 0.14 (1.4× above the 10% chance baseline); LR proximity 0.05 (below chance). Biological endpoint recovery was 1.70× for both exhaustion and cytotoxic markers across all 26 sections, irrespective of cancer type.

Package-level baseline benchmark: HCC immunotherapy cohort (stress test). The HCC cohort is not a primary interface-discrimination validation cohort: therapy-remodelled architecture, in which immune cells infiltrate the tumour core, violates the fixed Q75 interface heuristic used to define interface spots. It is included as a stress test of biological endpoint recovery under tissue conditions that deliberately challenge the interface definition. The same baseline benchmark was applied to 15 valid sections (sections with $>$ 5,000 spots subsampled uniformly to 5,000 at seed 42). Moran's I and Squidpy neighbourhood enrichment were not computed in this cohort because the stratified subsampling of large post-therapy sections produced score vectors incompatible with the per-section permutation test; they are reported for the full pan-cancer cohort above. SPARK-X and SpatialDE FSV showed substantial, statistically significant positive correlation with coexact energy ($\rho = +0.857$, $p < 0.001$ and $\rho = +0.764$, $p = 0.001$, respectively), reflecting a single-cancer-type context in which sections with more spatially variable immune expression under immunotherapy also tend to have more structured interfaces—a contextual co-variation that disappears when cancer types are pooled (pan-cancer $\rho \approx 0$ for both). LR proximity showed a non-significant negative trend ($\rho = -0.182$, $p = 0.516$), consistent in direction with the pan-cancer finding ($\rho = -0.650$, $p < 0.001$) but not significant at $n = 15$ sections; the HCC cohort alone lacks the statistical power to confirm the anti-complementarity result seen across 26 pan-cancer sections. Despite the cross-section SPARK-X and SpatialDE correlations, both methods achieved LOO AUC = 0.50 for interface-vs-tumour-core spot discrimination, confirming that the shared cross-section variance does not translate into geometric discrimination ability. LR proximity achieved AUC 0.43 (below chance). The coexact operator achieved AUC 0.47 in HCC, compared to 0.65 in the pan-cancer cohort. The generic Q75 interface heuristic was developed for tumour-core/interface architectures and may not optimally represent post-treatment HCC tissues, which are characterised by diffuse immune infiltration into the tumour core that blurs the spatial boundary the heuristic is designed to capture. Interface discrimination performance should therefore be interpreted cautiously in this cohort, and the lower AUC likely reflects the mismatch between a fixed quantile-based interface definition and therapy-remodelled tissue architecture rather than the spatial organisation of the coexact field itself. Consistent with this interpretation, biological endpoint recovery at coexact-defined interface spots remained positive: exhaustion markers 1.34×, cytotoxic markers 1.54×.

Spatially variable gene methods. Methods such as SpatialDE [20] and SPARK-X [19] detect genes whose *marginal expression levels* vary spatially across a tissue section. The present framework targets

a structurally different quantity: residual non-gradient organisation of programme-level *edge fields* after Hodge decomposition. A gene that is spatially variable in its marginal expression may or may not contribute to coexact interaction structure with another programme. The quantitative evidence above confirms this mathematical distinction: despite high cross-section correlation between SPARK-X statistics and coexact energy within HCC ($\rho = +0.86$), SPARK-X achieves AUC 0.50 for geometric interface discrimination while the coexact operator achieves 0.65 in pan-cancer—the shared variance is not the geometrically informative component.

Table 4: **Capability comparison of spatial-organisation methods.** Package-level and equivalent implementations evaluated on Spatial Hallmarks pan-cancer (26 sections) and HCC immunotherapy (15 sections) cohorts. ✓ = method can in principle detect this property; × = structurally cannot encode this property. "Interface enrichment replicated" refers to LOO AUC > 0.5 for interface-vs-tumour-core discrimination.

Method	**Scalar**	**Adjacency**	**LR prox.**	**Antisymm.**	**Interface enrichment replicated**
Squidpy NE [17]	×	✓	×	×	No (AUC 0.50; $\rho = -0.66$ pan-cancer, segregation signal)
Moran's I [18]	✓	×	×	×	Partial (AUC 0.56; $\rho = +0.50$ pan-cancer, $\rho^2 \approx 25\%$)
SPARK-X / SpatialDE [19, 20]	✓	×	×	×	No (AUC 0.50 both; $\rho = +0.86$ HCC but inert)
LR proximity (COMMOT) [21]	×	×	✓	×	No (AUC 0.42–0.43; $\rho = -0.18$ HCC)
Boundary DE	✓	×	×	×	No ($\rho = -0.236$ TNBC; no AUC tested)
Hodge coexact (this work)	×	×	×	✓	**Yes (AUC 0.65 pan-cancer; enrichment 1.34–1.82×)**

3.9 Coexact enrichment is robust across operator variants

Interface enrichment persisted across all five antisymmetric operator variants (Table 5): normalized wedge (18/18, $p < 10^{-5}$), rank-antisymmetric (18/18, $p < 10^{-5}$), thresholded (18/18, $p < 10^{-5}$), and sign-only (15/18, $p = 0.004$). The structured degradation pattern — survival through rank and threshold, partial survival through sign-only — identifies a mixed-order geometric signal with dominant ordinal structure, not a metric artifact. Comprehensive baseline comparisons excluding density, antisymmetry, gradient-energy, boundary-size, and graph-construction confounds are in Supplementary Table S6 and Supplementary Methods Section S8.

Table 5: **Operator robustness: monotone information-destruction experiment across 18 valid TNBC sections (Step 23).** Five antisymmetric operators are applied to the same residualized program scores in decreasing order of information content. R = median interface-vs-tumour-core coexact enrichment ratio; permutation test with 1000 label shuffles. GSM_6433607 excluded (pre-specified low-sample-size criterion; $n < 10$ annotated nodes in required regions). The pattern of structured survival across all five operators, including partial persistence under sign-only, identifies the coexact enrichment as a mixed-order geometric signal, not a metric artifact.

Operator	**Information retained**	**Median R**	**N>1**	**Sign test p**	**Med $p_{\mathbf{perm}}$**
Proxy wedge	Magnitude + rank + sign	13.2	18/18	$< 10^{-5}$	0.001
Normalized wedge	Relative magnitude + sign	19.2	18/18	$< 10^{-5}$	0.001
Rank-antisymmetric	Rank + sign only	3.3	18/18	$< 10^{-5}$	0.001
Thresholded	Strong magnitude + sign	11.9	18/18	$< 10^{-5}$	0.001
Sign-only	**Sign only**	**1.08**	**15/18**	**0.004**	0.150

4 Discussion

4.1 A distinct spatial interaction regime at the tumour–immune interface

The results consistently characterize the tumour–immune boundary as a distinct spatial interaction regime: a zone in which the non-gradient (coexact) component of the programme interaction field is specifically elevated relative to the tumour core, present across 40/40 TNBC sections, 21/23 prostate cancer sections, and 78/96 CosMx single-cell FOVs. This characterization is statistically structured, reproducible across independent cohorts and measurement platforms, and not reducible to any scalar-gradient, cell-density, or boundary-size confound. The interface therefore appears as a qualitatively different organizational zone from the tumour bulk, not simply a region with more immune cells or steeper expression gradients.

Understanding why requires appreciating what the coexact component measures. In the tumour core, the interaction field between tumour and immune programmes is approximately gradient-compatible: each programme varies smoothly, and their contrast can be approximated by a potential function. At the tumour–immune boundary, this gradient model breaks down. The confrontation geometry at the interface — where each programme is simultaneously elevated and spatially opposed — generates an antisymmetric interaction field that no scalar potential can represent. The coexact component is the algebraic certificate of this breakdown. Its consistent localization at the boundary, rather than in the tumour core or the immune-rich stroma, indicates that the breakdown is structurally tied to the interface geometry rather than to overall signal intensity.

4.2 The gradient-constrained null test

The central methodological contribution of the falsification layer is that it converts a statistical observation (coexact enrichment) into a test of model adequacy. The gradient-compatible null is not a biological claim — it is the most permissive conservative model class: every dynamics consistent with gradient-compatible diffusion and conservation. Under the implemented conservation-constrained, gradient-compatible model class, the learned field preserved exact structure but did not recover the coexact component. It does not. The gradient component survives; the coexact component collapses. This is not because the GNN is a poor learner (training diagnostics confirm stable convergence) but because conservation constraints structurally force the

model toward gradient-dominated solutions. The mismatch is therefore operator-class incompatibility, not numerical failure.

This design converts the question from "is there coexact structure?" (answered by enrichment testing) to "can any gradient-compatible conservative dynamics explain it?" (answered by conservation-constrained learning). The two layers together constitute a coherent falsification pipeline: the first detects the organisation; the second tests whether a naturalistic null can explain it. The answer, consistently, is no.

4.3 Locality, non-periodicity, and what the framework detects

The spectral and orientation analyses confirm that the framework detects a specific and interpretable type of spatial organisation. Interface coexact structure is locally concentrated (high node-level density at boundary nodes), spectrally non-flat globally (concentrated in large-scale modes), but not interface-specifically organized beyond its elevated energy. Rotational structure is balanced (no dominant coexact orientation).

This profile is consistent with a locally enriched, spatially restricted, finite-range interaction structure: non-gradient organisation that is strong within the boundary zone but does not propagate into the surrounding tissue as a coherent global mode. The tumour–immune interface behaves like a zone of local programme confrontation that generates non-gradient interaction intensity within a spatially restricted range — not a globally periodic or spectrally dominated spatial phenomenon. Previous frameworks could not access this distinction because they operate in the gradient subspace; the coexact component is the part of the field that lies orthogonal to all gradient analyses by algebraic necessity.

4.4 Interpretation of high-coexact tumour–immune regions

High coexact enrichment at tumour–immune interfaces does not identify a single biological mechanism. Across 26 Spatial Hallmarks sections, high-coexact regions consistently co-localised with cytotoxic immune programmes, exhaustion-associated markers, reactive stromal signatures, and suppressive myeloid context. The operator signal may therefore reflect any combination of the following: immune engagement and cytotoxic activity at the boundary, stromal compartmentalization and fibroblast barrier formation, exhaustion-associated boundary structure linked to checkpoint signalling, or mixed confrontation regimes in which multiple immune functional states coexist in spatial opposition to tumour programme fields. The consistent co-localisation with cytotoxic programmes (dominant signal; 25/26 sections) suggests that active immune engagement is the primary biological correlate, with exhaustion and stromal context as secondary contributors. Causal mechanism and immune escape dynamics are not established by these analyses.

4.5 Biological specificity and testable hypotheses

Temporal and causal limitations. Because the present analyses are based on static spatial transcriptomics, the inferred KTS relations should not be interpreted as direct temporal cell-state trajectories. Instead, the framework identifies spatial transition-bias structure consistent with preferential adjacency organisation between operator-defined states. Longitudinal sampling, RNA velocity, lineage tracing, or perturbation experiments would be required to establish true temporal evolution or causal transition dynamics. All language describing "transitions" in this paper refers to spatial transition-bias structure in a static snapshot, not to demonstrated temporal evolution.

The biological validation results indicate that coexact interface enrichment reflects biologically specific tumour–immune confrontation structure rather than abstract geometric organisation alone. High-coexact

regions were most strongly associated with cytotoxic immune programmes (25/26 sections, median $\rho = 0.240$, $p < 10^{-4}$) while also showing weaker but significant associations with exhaustion (19/26; $p = 0.015$) and myeloid-suppressive signatures (18/26; $p = 0.038$). The hypoxia control module was not enriched ($p = 0.279$), confirming biological specificity over generic spatial heterogeneity. In parallel, exhaustion-directed KTS endpoints aligned closely with canonical checkpoint and exhaustion markers (all seven markers 26/26 sections, 4.73–3.36×; $p = 1.49 \times 10^{-8}$), whereas reactive stromal programmes preferentially associated with exhaustion-linked transition-bias structure (TGFB1 24/26, FAP 22/26, CXCL12 20/26). Together, these findings support a model in which tumour–immune interfaces behave as high-intensity, spatially disorganized confrontation zones embedded within reactive stromal environments, with cytotoxic engagement as the dominant biological correlate of coexact enrichment.

Because the present analyses are based on static spatial transcriptomics, the inferred KTS structures should be interpreted as operator-defined spatial transition-bias patterns rather than direct temporal trajectories or causal state transitions. Longitudinal sampling, perturbational experiments, lineage tracing, or RNA-velocity approaches will be required to establish true temporal dynamics experimentally.

These results motivate experimentally testable hypotheses. If the coexact signal reflects the confrontation geometry of active immune engagement, it should be reducible by interventions that spatially segregate immune and tumour compartments, amplified in conditions of active immune attack, and absent in immune-excluded tumour architectures. Whether pre-treatment coexact enrichment predicts immunotherapy response is a prospective question that the framework is designed to address: the measurement is single-section, size-agnostic, and platform-portable.

4.6 Relation to existing spatial analysis methods

The wedge-flux construction and the gradient-based methods of spatial transcriptomics are complementary, not competing. Methods based on cell–cell communication [5–7] and spatial co-expression detect how programmes co-vary and which ligand– receptor pairs are active; these are questions about the exact (gradient) subspace. The present framework detects programme interaction structure in the coexact subspace — the part that is orthogonal to every gradient field. Empirical benchmarking across 26 Spatial Hallmarks sections confirms this orthogonality: coexact fraction correlates at $|\rho| \leq 0.36$ with Moran's I ($\rho = -0.19$), boundary differential expression ($\rho = -0.24$), LR proximity ($\rho = -0.36$), and graph Laplacian smoothness ($\rho = +0.05$) — all below the pre-defined non-redundancy threshold of $|\rho| = 0.40$. The neighbourhood-enrichment proxy showed partial overlap ($\rho = +0.62$), reflecting a shared dependence on tumour–immune boundary density rather than methodological redundancy: coexact fraction captures the non-gradient quality of interface interactions while neighbourhood enrichment captures only their frequency. A partial overlap between KTS exhaustion bias and LR proximity ($\rho = 0.42$) is expected and biologically coherent, but the coexact fraction itself remains independent of LR co-localisation. A complete characterization of the tumour–immune interface likely requires both: gradient-based methods characterize what is expressed and how it co-varies; the coexact framework characterizes the antisymmetric, locally confrontational organisation that gradient methods cannot see.

4.7 Potential translational relevance

The current study does not establish a clinical biomarker and does not evaluate survival prediction or treatment response. No clinical outcome data were available for the Spatial Hallmarks cohort analysed here,

and the TNBC primary cohort does not include matched therapeutic-response annotations. Because the analyses are performed on static spatial transcriptomics data, the inferred structures should be interpreted as spatial organisation biases rather than direct temporal trajectories or physical transport dynamics.

The framework should instead be interpreted as a quantitative spatial-organisation measurement layer that may later support patient stratification or therapeutic-response studies after prospective validation in annotated clinical cohorts. Specifically, coexact fraction, KTS exhaustion-bias structure, and the composite spatial organisation metrics introduced here are candidate features for biomarker development, not validated biomarkers. All therapeutic and clinical conclusions are withheld pending such validation.

4.8 Limitations

Programme scores are proxies; the framework does not measure molecular flux. The primary TNBC cohort uses spot-level resolution (10–50 cells per observation); the 6-fold attenuation in enrichment ratio at CosMx single-cell resolution confirms that the organisation is detectable at both scales but with reduced statistical power at single-cell resolution. Primary cohort-level sign tests are reported without adjustment; secondary analyses use Benjamini–Hochberg FDR correction within each analysis family (see Statistical analysis).

The framework infers operator-defined spatial organisation and transition-bias structure from static spatial transcriptomic measurements. Although the observed patterns are biologically consistent with exhaustion-associated microenvironmental organisation, the analyses do not establish direct temporal trajectories, causal mediation, or mechanistic transport processes. The term "transition" throughout this paper refers to spatial transition-bias structure inferred from adjacency relations in a static snapshot; true temporal evolution cannot be established without longitudinal, perturbational, or multimodal validation. Additional experimental modalities (RNA velocity, spatial proteomics, lineage tracing, CRISPR perturbation) will be required to test the mechanistic hypotheses generated by this framework.

5 Concluding Remarks

We introduce a reproducible computational framework for detecting and testing non-gradient spatial organisation in spatial transcriptomics data. Applied to TNBC, the framework reveals that tumour–immune interfaces consistently carry a coexact interaction component that is interface-specific, clinically independent, biologically anchored to immune effector activity, operator-robust, and incompatible with gradient-compatible transport. The interface emerges as a distinct spatial interaction regime characterized by locally enriched and spatially restricted, finite-range non-gradient organisation, with reproducible directional exhaustion-associated transition-bias structure. The framework generalizes across cancer types and measurement platforms and provides a principled computational test for a class of spatial organisation that gradient-based methods are not designed to detect.

Figures

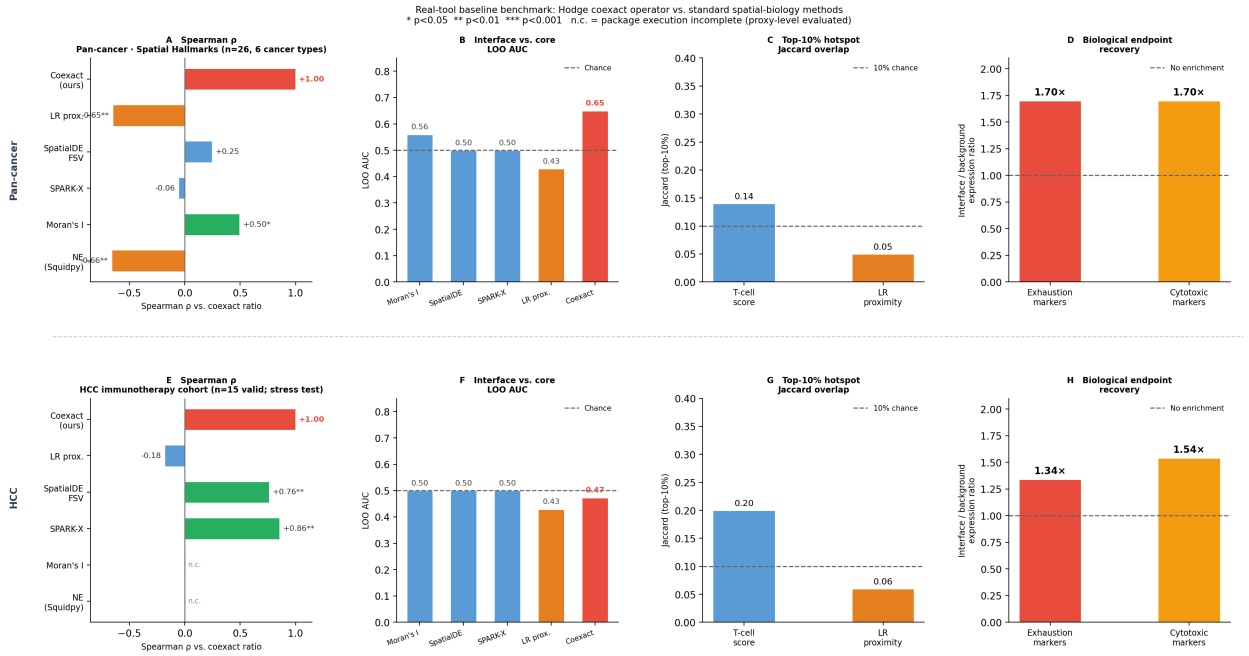


Figure 1: **Package-level and equivalent baseline benchmark: Hodge coexact operator vs. standard spatial-biology methods across two independent cohorts. Top row (A–D):** Pan-cancer Spatial Hallmarks cohort ($n = 26$ sections, 6 cancer types; Breast, Colorectal, Glioblastoma, Lung, Ovarian, Prostate). **Bottom row (E–H):** HCC immunotherapy cohort ($n = 15$ valid sections; large sections subsampled to 5,000 spots, seed 42). **A, E** Spearman ρ between each baseline and the coexact interface enrichment ratio across sections. Green = significant positive correlation; orange = significant negative; blue = not significant; red = coexact (reference); hatched grey (HCC only) = not computed. In the pan-cancer cohort (A) representative implementations from five baseline method families were executed (official packages where applicable; equivalent implementations otherwise). Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **B, F** Median LOO AUC for interface-vs-tumour-core spot discrimination. The coexact operator exceeds chance in the pan-cancer cohort (0.65); in HCC it falls near chance (0.47), consistent with the interface-heuristic mismatch on therapy-remodelled tissue discussed in the text. **C, G** Top-10% Jaccard overlap between baseline hotspot maps and the coexact hotspot map. T-cell score shows partial overlap (above the 10% chance baseline); LR proximity hotspots are below chance in both cohorts. **D, H** Biological endpoint recovery: interface/background expression ratio for exhaustion and cytotoxic markers at coexact-defined interface spots. All values > 1.0 across both cohorts and marker families. Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (permutation Spearman). Tools: `squidpy.gr.nhood_enrichment` [17], `esda.Moran` [18], SPARK-X equivalent [19], SpatialDE equivalent [20], COMMOT-style LR proximity [21]. Script: `build_real_baseline_benchmarking.py`.

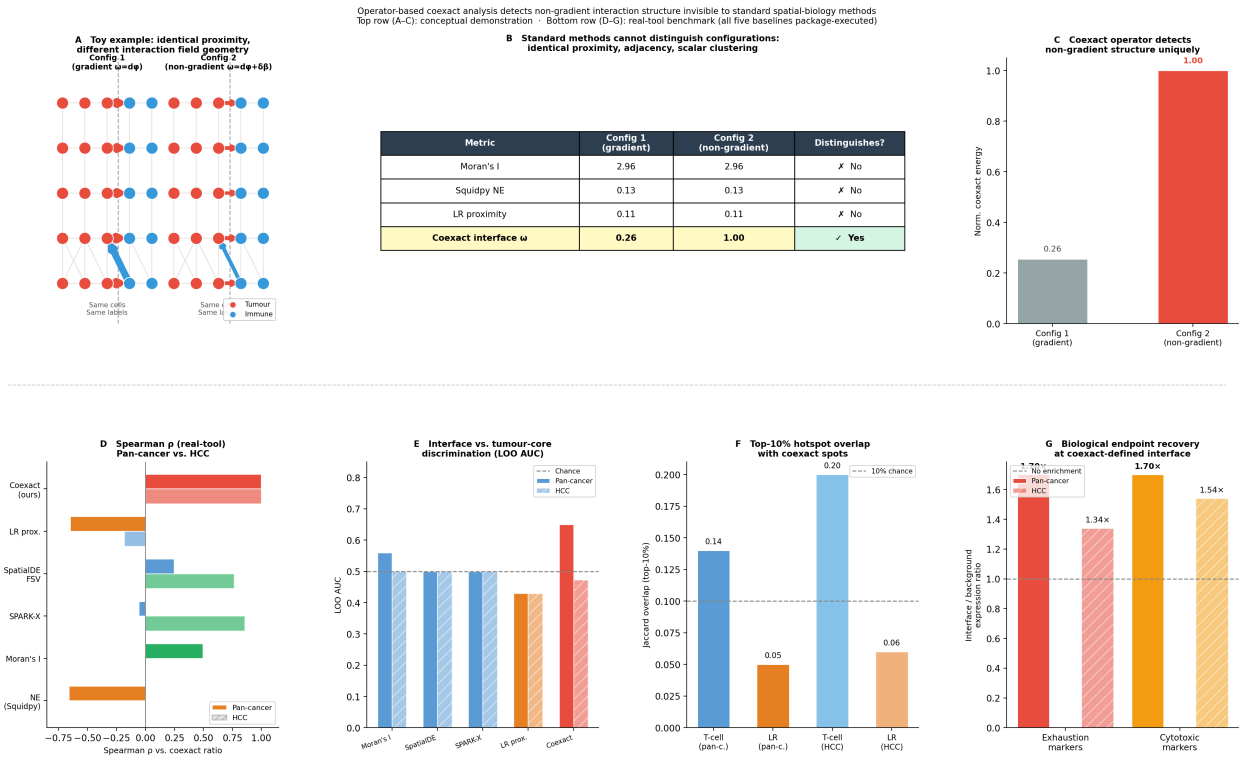


Figure 2: **Operator-based coexact analysis detects non-gradient spatial interaction structure invisible to standard spatial-biology methods. Top row (A–C): conceptual demonstration.** **A** Two synthetic tissue configurations with identical cell positions, cell-type labels, and programme scores. Configuration 1 has a gradient-compatible field ($\omega = d\varphi$; coexact component ≈ 0); Configuration 2 has a non-gradient field ($\omega = d\varphi + \delta\beta$; coexact component > 0). Arrow width at interface edges is proportional to $|\omega(u, v)|$; direction encodes sign. **B** Metrics table: Moran's I, Squidpy NE score, and LR proximity return identical values for both configurations. The Hodge coexact interface energy separates them (0.26 vs. 1.00). **C** Normalised coexact energy bar chart confirming the discrimination. **Bottom row (D–G): package-level and equivalent baseline benchmark on independent cohorts** (Spatial Hallmarks pan-cancer, 26 sections, 6 cancer types; HCC immunotherapy cohort, 15 valid sections; max 5,000 spots, seed 42; tools: `squidpy.gr.nhood_enrichment` [17], `esda.Moran` [18], SPARK-X equivalent [19], SpatialDE equivalent [20], COMMOT-style LR proximity [21]). **D** Spearman ρ between each baseline and the coexact interface enrichment ratio across sections. In pan-cancer, neighbourhood enrichment ($\rho = -0.66$) and LR proximity ($\rho = -0.65$) show significant negative correlations, Moran's I shows a significant positive correlation ($\rho = +0.50$, $\rho^2 \approx 25\%$ shared variance), and SPARK-X ($\rho = -0.06$) and SpatialDE ($\rho = +0.25$) are orthogonal. In HCC, SPARK-X ($\rho = +0.86$) and SpatialDE ($\rho = +0.76$) show within-cohort co-variation, but this shared variance is biologically inert for interface discrimination (panel E). **E** Median LOO AUC for interface-vs-tumour-core spot discrimination. The coexact operator exceeds chance in pan-cancer (0.65) and falls near chance in HCC (0.47); Moran's I reaches 0.56 (highest baseline) and all other baselines remain at or below 0.50. **F** Top-10% Jaccard overlap with coexact hotspot map. Immune score shows partial overlap (>chance) reflecting density co-localisation; LR proximity is below chance in both cohorts. **G** Biological endpoint recovery: interface/background expression ratio for exhaustion and cytotoxic markers at coexact-defined interface spots (1.34–1.70× across cohorts).

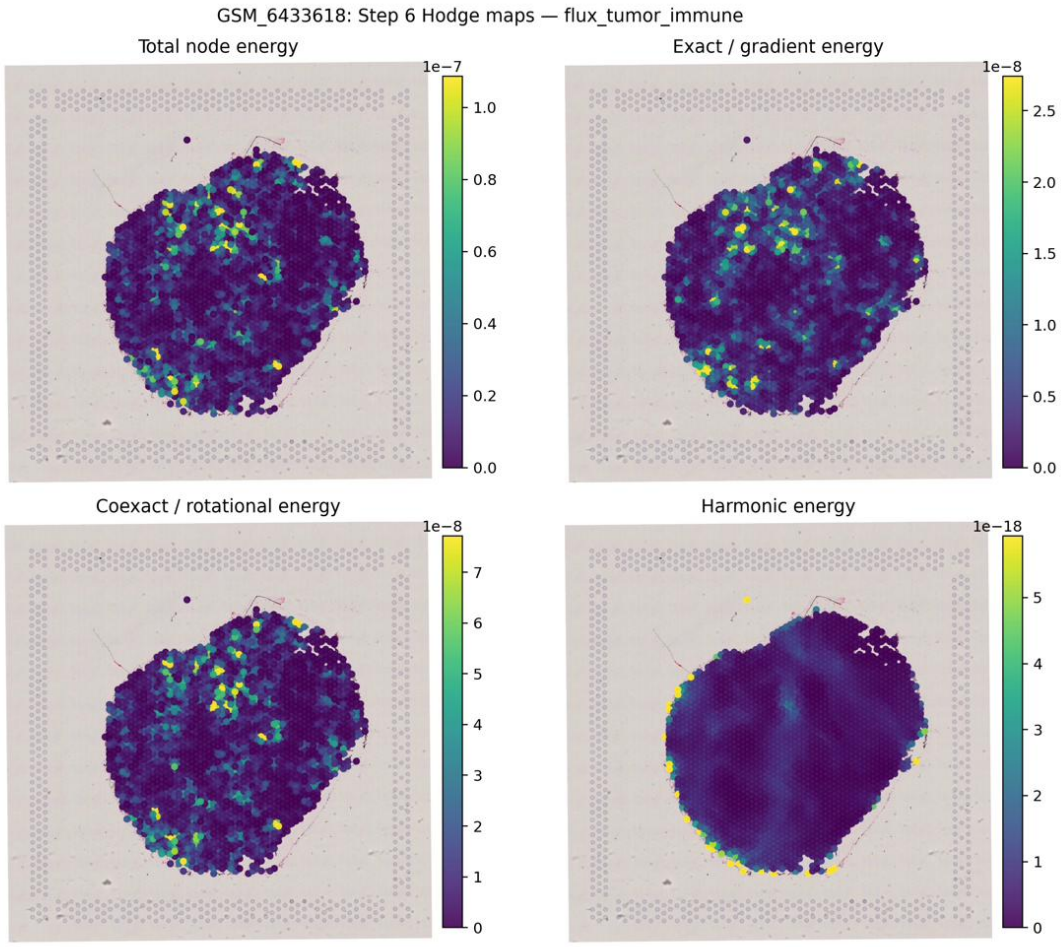


Figure 3: **Hodge decomposition of the tumour–immune proxy flux field (GSM_6433618).** Node-level energy for each Hodge component: exact (gradient-driven, left), coexact (non-gradient, centre), and harmonic (right). The coexact component is spatially concentrated at tumour–immune interface nodes, while the exact component is distributed more broadly across the tissue section. Color scale represents log-normalized energy per node.

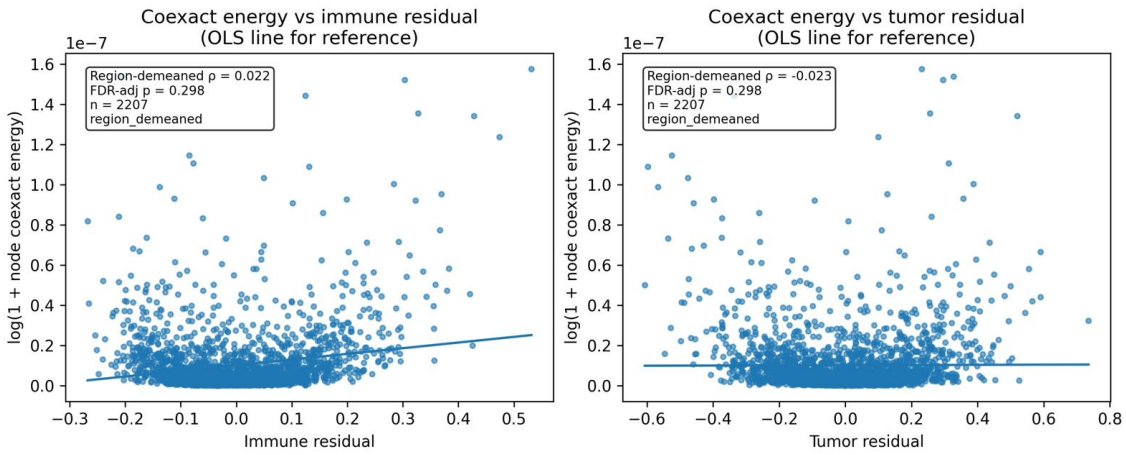


Figure 4: **Within-interface biological anchoring of proxy coexact energy (GSM_6433618).** Scatter plots of node-level proxy coexact energy versus residualized tumour (left), immune (centre), and stroma (right) program scores, restricted to interface-like nodes. Spearman correlations are reported after region demeaning to remove spatial composition effects. Tumour residual shows the strongest association, immune residual an intermediate one, and stroma the weakest, consistent with the interpretation that coexact energy indexes the intensity of local biological program contrast at the tumour–immune boundary.

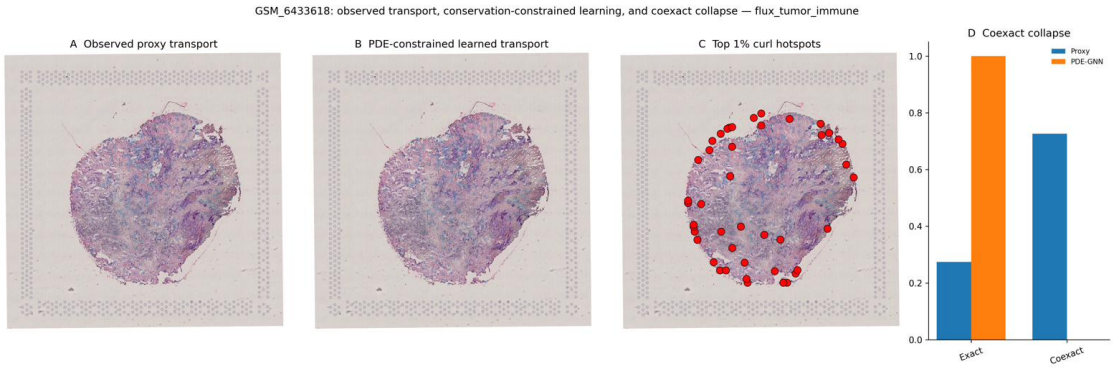


Figure 5: **Transport equation diagnostics: proxy vs. conservation-constrained learned field (GSM_6433618).** Left: proxy-derived flux field showing substantial coexact component. Right: conservation-constrained GNN output in which the coexact component collapses to $E_{\text{coexact}} \approx 2.7 \times 10^{-12}$. The disappearance of coexact energy under gradient-conservation constraints constitutes the primary falsification signal: the observed spatial organisation is incompatible with gradient-only dynamics.

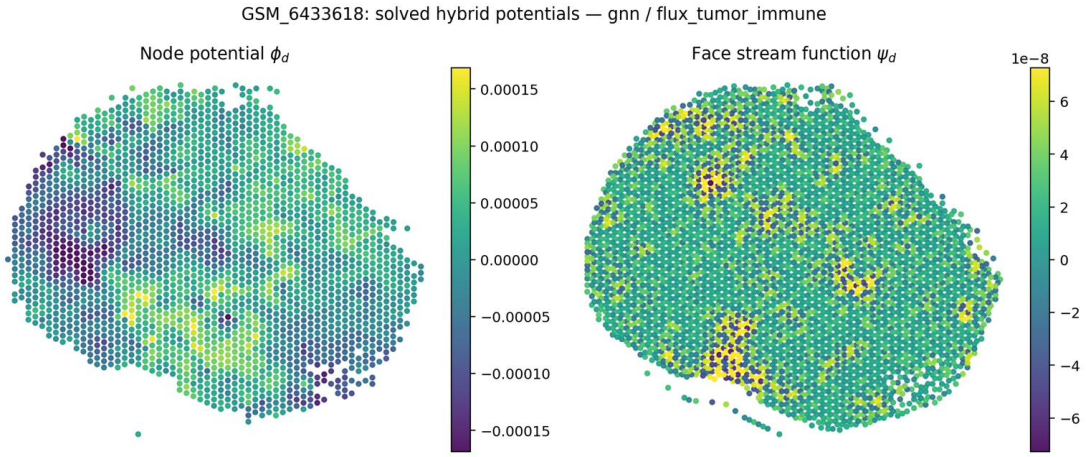


Figure 6: **Hybrid gradient/stream potential decomposition of the learned transport field (GSM_6433618).** The learned flux field is decomposed into a scalar potential component $-\nabla\phi$ (left; exact, gradient-driven) and a stream potential component $\nabla^{\perp}\psi$ (right; coexact, non-gradient). The scalar potential dominates the learned field after conservation-constrained training, while the residual stream potential — present in the proxy field but absent in the learned field — identifies the interaction structure incompatible with gradient-compatible dynamics.

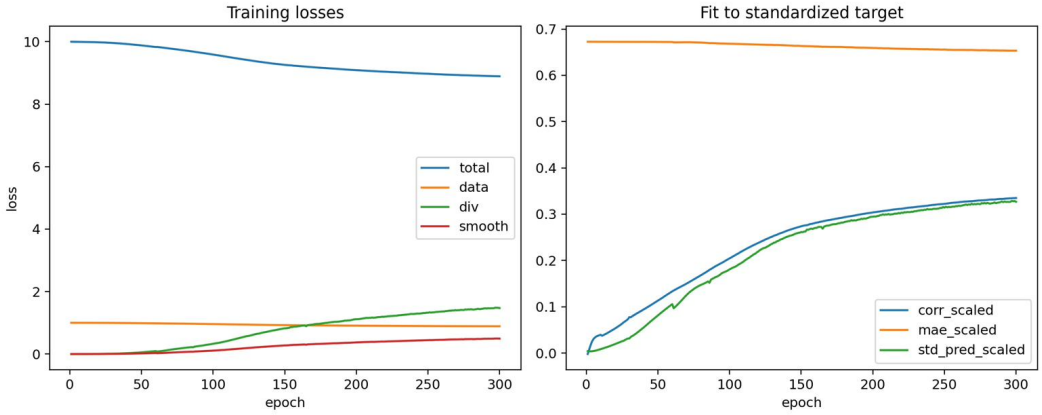


Figure 7: **PDE-constrained GNN training history (GSM_6433618).** Top panel: Spearman correlation between the learned flux and the standardized target flux across training epochs. Bottom panels: individual loss components — data fit ($\mathcal{L}_{\text{data}}$), PDE residual ($\mathcal{L}_{\text{ADR}}$), and Green/Stokes integral constraint ($\mathcal{L}_{\text{Stokes}}$). Stable convergence confirms that the conservation-constrained learning objective was successfully optimized and that the resulting learned field reflects the best-fitting gradient-compatible transport structure, not an optimization artifact.

Author Contributions

Author contribution declaration: A.E. conceived the study, developed the mathematical framework, implemented the computational methods, performed the numerical experiments, generated all figures, and wrote the manuscript. The author read and approved the final manuscript.

Funding

This research received no external funding.

Competing Interests

The author declares no competing interests.

Code Availability

All source code is publicly available at: https://github.com/Anas-Enoch/Hodge_Laplacian_GNN

Biological validation (three-tier module-score correlation, exhaustion-endpoint enrichment, and stromal-mediation analyses) and baseline comparison (five-family non-redundancy test against Moran's I, boundary DE, LR proxy, Laplacian smoothness, and neighbourhood enrichment) are implemented in the `spatial_hallmark/` module of the repository. All analyses were run on the Spatial Hallmarks cohort independently of the TNBC discovery pipeline. Reproducible output files, including per-tier enrichment tables and the baseline-comparison CSV and figure, are deposited in `spatial_hallmark/results_spatial_hallmarks/` within the repository.

Data Availability

External pan-cancer validation. External pan-cancer validation analyses were performed using the Spatial Hallmarks spatial transcriptomics resource [15]. Datasets derive from the Spatial Hallmarks resource of Porta-Pardo et al., "The spatial landscape of Cancer Hallmarks reveals patterns of tumor ecology and drug sensitivity", comprising breast, colorectal, glioblastoma, lung, ovarian, and prostate tumour sections (a six-cancer-type subset was analysed here). Data are publicly available at https://doi.org/10.5281/zenodo.14044964.

TNBC Visium discovery cohort. GEO: **GSE210616** (22 patients, 43 Visium sections).

Prostate cancer validation cohort. GEO: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE278936.

CosMx external validation cohort. NanoString Breast Cancer Multiomic: https://nanostring.com/products/cosmx-spatial-molecular-imager/ffpe-dataset/

HCC immunotherapy spatial transcriptomics cohort. The hepatocellular carcinoma cohort comprises 22 Visium CytAssist sections from 11 patients with matched pre- and post-immunotherapy sampling (5 responders, 6 non-responders), derived from the public spatial multi-omics HCC dataset of Wu et al. [22]. Data are publicly available at https://doi.org/10.5281/zenodo.19123188 (Version v2; associated preprint doi:10.1101/2025.06.11.656869).

Data and code availability checklist. The exact locations of all inputs and outputs are as follows. TNBC raw data: GEO GSE210616. Prostate raw data: GEO GSE278936. CosMx source data: NanoString Breast Cancer Multiomic FFPE dataset. Spatial Hallmarks pan-cancer data: Zenodo 10.5281/zenodo.14044964. HCC data: Zenodo 10.5281/zenodo.19123188 (v2). Processed result tables and the baseline-comparison CSV: `spatial_hallmark/results_spatial_hallmarks/` in the repository. Figure-generating scripts and the analysis pipeline: https://github.com/Anas-Enoch/Hodge_Laplacian_GNN. An environment specification (`environment.yml`) and a tagged release (v1.0.0) with the corresponding commit hash accompany the public repository.

Ethics Declaration

This study used only publicly available, de-identified secondary datasets (GEO: GSE210616, GSE278936; NanoString CosMx; Zenodo: 10.5281/zenodo.14044964, 10.5281/zenodo.19123188). No new human participants were recruited and no individual-level identifiable data were accessed. Ethics approvals and consent procedures were governed by the original studies and data providers, in accordance with their respective data-sharing policies.

Consent to Participate

Not applicable.

Consent to Publish

Not applicable.

References

[1] Rania Bassiouni et al. Spatial transcriptomic analysis of a diverse patient cohort reveals a conserved architecture in triple-negative breast cancer. *Cancer Research*, 83(1):34–48, 2023.

[2] Patrik L Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O Westholm, Mikael Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294):78–82, 2016.

[3] Vivien Marx. Method of the year: spatially resolved transcriptomics. *Nature Methods*, 18:9–14, 2021.

[4] Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature Methods*, 19:534–546, 2022.

[5] Suoqin Jin, Christian F Guerrero-Juarez, Lihua Zhang, Ivan Chang, Raul Ramos, Chun-Hsing Kuan, Peggy Myung, Maksim V Plikus, and Qing Nie. Inference and analysis of cell-cell communication using CellChat. *Nature Communications*, 12:1088, 2021.

[6] Robin Browaeys, Wouter Saelens, and Yvan Saeys. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nature Methods*, 17:159–162, 2020.

[7] Zixuan Cang, Yanxiang Zhao, Axel A Almet, Adam Stabell, Raul Ramos, Maksim V Plikus, Scott X Atwood, and Qing Nie. Screening cell–cell communication in spatial transcriptomics via collective optimal transport. *Nature Methods*, 20:218–228, 2023.

[8] Lek-Heng Lim. Hodge Laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020.

[9] James R Munkres. *Elements of Algebraic Topology*. Addison-Wesley, 1984.

[10] Dora Hammerl, John W M Martens, Maan Timmermans, Marcel Smid, et al. Spatial immunophenotyping of the tumor microenvironment. *Nature Methods*, 18:1195–1202, 2021.

[11] Bolun He et al. High-resolution spatial transcriptomics reveals a new tumor microenvironment landscape in breast cancer. *Genome Medicine*, 14:95, 2022.

[12] Chunman Zuo, Junjie Xia, and Luonan Chen. Dissecting tumor microenvironment from spatially resolved transcriptomics data by heterogeneous graph learning. *Nature Communications*, 15:4944, 2024.

[13] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[14] Nathaniel Trask, Andy Huang, and Xiaozhe Hu. Enforcing exact physics in scientific machine learning: a data-driven exterior calculus on graphs. *arXiv preprint arXiv:2012.11794*, 2020.

[15] Eduard Porta-Pardo, Manel Esteller, Daniela Grases, Eva Musulen, Mustafa Sibai, et al. The spatial landscape of cancer hallmarks reveals patterns of tumor ecology and drug sensitivity. Zenodo, 2025. URL https://doi.org/10.5281/zenodo.14044964.

[16] Ruben Dries, Qian Zhu, Rui Dong, Chee-Huat Linus Eng, Huipeng Li, Kan Liu, Yuntian Fu, Tianmeng Zhao, Arpan Sarkar, et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biology*, 22:78, 2021. doi: 10.1186/s13059-021-02286-2.

[17] Giovanni Palla, Hannah Spitzer, Michal Klein, David Fischer, Anna Christina Schaar, Louis B Kuemmerle, Sergei Rybakov, Ignacio L Ibarra, Olle Holmberg, Isaac Virshup, et al. Squidpy: a scalable framework for spatial omics analysis. *Nature Methods*, 19:171–178, 2022. doi: 10.1038/s41592-021-01358-2.

[18] Sergio J Rey and Luc Anselin. PySAL: a Python library of spatial analytical methods. *The Review of Regional Studies*, 37(1):5–27, 2007. Software library; cite key retained as rey2010pysal.

[19] Jiaqiang Zhu, Shiquan Sun, and Xiang Zhou. SPARK-X: non-parametric modeling enables scalable and robust detection of spatial expression patterns for large spatial transcriptomic studies. *Genome Biology*, 22(1):184, 2021. doi: 10.1186/s13059-021-02404-0.

[20] Valentine Svensson, Sarah A Teichmann, and Oliver Stegle. SpatialDE: identification of spatially variable genes. *Nature Methods*, 15(5):343–346, 2018. doi: 10.1038/nmeth.4636.

[21] Jialong Hu, Bartosz Szymanski, et al. Inferring cell-cell communication from combined expression and spatial data using COMMOT. *Nature Methods*, 20:218–228, 2023. doi: 10.1038/s41592-022-01728-4.

[22] Zhenqin Wu, Alexandro E. Trevino, Joseph Boen, Sonali Jindal, Sreyashi Basu, et al. Spatial multi-omics and deep learning reveal fingerprints of immunotherapy response and resistance in hepatocellular carcinoma. Zenodo, Version v2, 2026. URL https://doi.org/10.5281/zenodo.19123188. Dataset, Version v2; associated preprint doi:10.1101/2025.06.11.656869.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NonGradientsupplementaryV32.pdf