PHISA: POLARIZED HYPERGRAPHS INDUCE SPECTRAL ADAPTATION OF BIOMEDICAL LLMS FOR SURVIVAL PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Tissue organization and molecular information provide complementary perspectives that jointly inform cancer survival prediction. However, existing multimodal survival methods struggle to characterize the *diverse relational patterns of tissue interactions*, weakening the *connection between tissue organization and molecular representation learning*. To address this limitation, we propose PHISA, which encodes diverse interaction tendencies into a structural representation of patient tissue to guide molecular representation learning with pretrained biomedical knowledge for survival prediction. Specifically, PHISA first forms a polarized hypergraph by learning phase orientations for regions within each hyperedge, such that relative phases encode cooperative and antagonistic tendencies. The resulting polarized hypergraph governs the propagation of representations across tissue regions, with global structural coordinates and local responses capturing tissue structure. PHISA then aggregates the global and local structural components into a patient tissue descriptor. This structural descriptor modulates spectral coefficients along fixed principal singular directions of pretrained biomedical LLM weights. Finally, PHISA combines tissue representations with molecular representations refined by spectral adaptation conditioned on the patient tissue descriptor for survival prediction. Experiments on four TCGA cohorts show that PHISA achieves a mean C-index of 0.750, representing a 6.1% relative improvement over the strongest evaluated baseline. Ablations show predictive gains from polarized tissue modeling, with additional benefit from conditioning molecular refinement on patient tissue structure. Further analyses reveal spatial variation in tissue region importance and significant survival differences between groups defined by predicted risk scores.

1 INTRODUCTION

Cancer survival analysis supports prognostic assessment and patient risk stratification by modeling survival outcomes over time Clark et al. (2003). Patient prognosis is associated with clinical characteristics and tumor biology, motivating the consideration of complementary information sources Jee et al. (2024). Clinical records, medical images, and molecular profiles provide observations of patient health and disease at different levels Liu et al. (2026). Among these sources, whole slide images (WSIs) reveal tissue morphology and spatial organization, while molecular profiles characterize genomic alterations and gene expression patterns Chen et al. (2022); Brussee et al. (2025). Integrating these complementary sources yields patient representations that capture distinct aspects of tumor biology relevant to survival prediction Long et al. (2026); Hu et al. (2025).

Tumor biology is closely related to cellular organization in the tumor microenvironment, where interactions among local cell populations are associated with tissue function and patient outcomes Danenberg et al. (2022). Cellular interactions among local cell populations can involve distinct and potentially opposing functional roles Rahal et al. (2025). Tertiary lymphoid structures in a mouse model of lung adenocarcinoma illustrate opposing functional roles among local cell populations. Dendritic cells presenting tumor antigens interact with T cells to support antitumor immunity, whereas regulatory T cells suppress the local response (Figure 1, upper left). Depleting regulatory T cells increases T cell proliferation within the lymphoid structures, supporting the local suppressive interac-

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

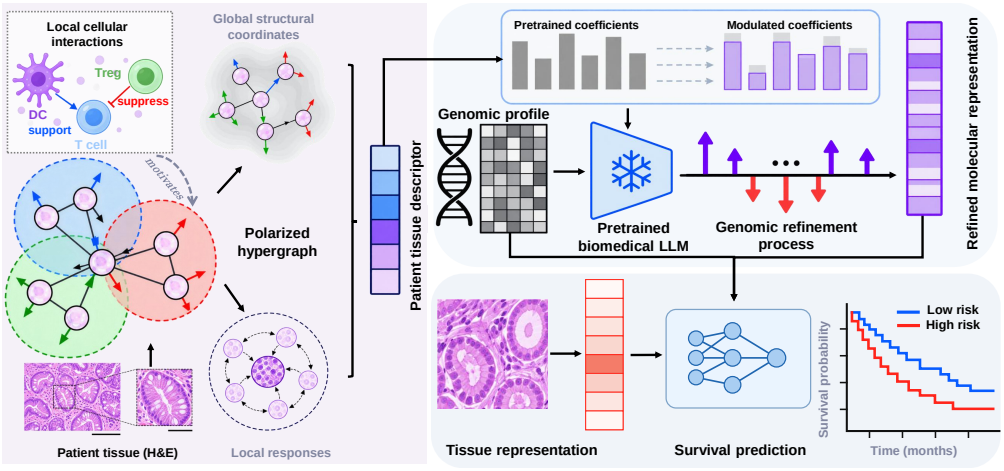


Figure 1: **Overview of PHISA**. Local cellular interactions motivate polarized hypergraph modeling, where relative phases encode cooperative and antagonistic tendencies. The polarized hypergraph governs regional representation propagation, while global structural coordinates and local responses are integrated into the patient tissue descriptor that conditions spectral adaptation of a pretrained biomedical LLM. Tissue and refined molecular representations are combined for survival prediction.

tion Joshi et al. (2015). For survival prediction, tissue representations should capture the functional diversity of cellular interactions that coexist within the same microenvironment.

Accordingly, cellular interaction diversity motivates modeling relationships among tissue regions. Regional relationships can be represented by pairwise edges in graphs, while hypergraphs connect groups of regions through shared hyperedges. Learnable mappings between nodes and hyperedges allow information transfer to vary across associations Duta et al. (2023). Directional hypergraph models further encode head and tail memberships to represent interaction direction Mule et al. (2026). Beyond interaction direction, cooperative and antagonistic relationships are difficult to distinguish when the same tissue region participates in local interactions with different tendencies.

Molecular profiles provide complementary information about biological processes associated with local tissue function Jaume et al. (2024). Biomedical knowledge links gene activity to biological processes and helps interpret molecular observations in tissue. Large language models (LLMs) pretrained on biomedical literature offer a potential foundation for molecular representation learning Labrak et al. (2024). Regional interaction patterns observed in patient tissue provide structural context that complements biomedical knowledge during molecular representation learning. However, general associations from pretraining corpora offer limited insight into each patient's tissue organization. Collectively, *tissue representations struggle to capture diverse regional interaction patterns sufficiently to inform molecular representation learning*.

To address this problem, we propose **PHISA** (**P**olarized **H**ypergraphs **I**nduce **S**pectral **A**daptation) for multimodal survival prediction (Figure 1). PHISA encodes diverse regional interaction patterns in tissue representations that guide molecular representation learning with pretrained biomedical knowledge. Specifically, PHISA first uses polarized hypergraph modeling to represent cooperative and antagonistic tendencies within each hyperedge. Relative phases between learned regional orientations encode these tendencies and modulate regional representation propagation across tissue regions. Global structural coordinates and local responses characterize complementary aspects of tissue structure and are integrated into the patient tissue descriptor. Spectral adaptation of pretrained biomedical LLM weights is then conditioned on the patient tissue descriptor to refine molecular representations. Finally, PHISA combines tissue representations with refined molecular representations for survival prediction. Experiments on four TCGA cohorts demonstrate PHISA's predictive advantage over the evaluated baselines. Ablations show gains from polarized tissue modeling and additional improvements when patient tissue descriptors guide molecular refinement. The proposed spectral adaptation outperforms alternative adaptation methods using the same biomedical LLM. Complementary analyses reveal spatial variation in tissue region importance and significantly different survival distributions between risk groups. Our main contributions are summarized as follows:

- A Polarized Hypergraphs Induce Spectral Adaptation (PHISA) framework is proposed to incorporate patient tissue organization into molecular representation learning for multimodal survival prediction. To our knowledge, PHISA establishes the first paradigm that conditions biomedical LLM adaptation on learned patient tissue structure for this task.
- A polarized hypergraph is designed to explicitly encode cooperative and antagonistic tendencies via relative phases of learned regional orientations. The resulting hypergraph captures complementary global and local tissue structure for expressive tissue representation learning.
- To enable adaptive molecular refinement, PHISA conditions biomedical LLM adaptation on patient tissue structure. Spectral coefficient modulation integrates tissue context with pretrained biomedical knowledge, yielding biologically informed molecular representations.
- Experiments on TCGA cohorts show that PHISA outperforms the evaluated baselines in survival prediction. Ablations support the predictive contribution of polarized modeling, while adaptation comparisons show the effectiveness of the proposed spectral modulation.

2 Related Work

Multimodal survival prediction. Existing models integrate histopathology and molecular measurements via fusion or interactions between representations. MCAT relates histology patches to genomic features with co-attention Chen et al. (2021). To organize molecular inputs around biological processes, SurvPath models dense interactions between pathways and tissue regions Jaume et al. (2024). Alongside interaction modeling, PIBD addresses redundancy via disentanglement of shared and modality-specific information Zhang et al. (2024). M^2Surv Qu et al. (2025a) further uses hypergraphs spanning multiple slides and guided by genomic information to model tissue relationships.

Hypergraph learning. Hypergraph methods aggregate group information for node representation updates. AllSet uses learnable multiset functions across nodes and hyperedges Chien et al. (2022). To distinguish node responses within hyperedges, ED-HNN conditions messages on receiving nodes Wang et al. (2023). SheafHNN allows node information to undergo different transformations across associations Duta et al. (2023). For directed interactions, DSHN extends sheaf hypergraph learning with mappings that encode head and tail memberships of hyperedges Mule et al. (2026).

Adaptation of pretrained models. Parameter-efficient adaptation varies in update parameterization and the sources that determine adaptation parameters. For update parameterization, LoRA represents weight updates with low-rank matrices Hu et al. (2022). PiSSA initializes trainable factors from principal singular components of pretrained weights, whereas SVFT learns update coefficients in a basis formed by fixed singular vectors Meng et al. (2024); Lingam et al. (2024). Beyond update parameterization, conditioning information can determine the adaptation parameters. For example, Hyperformer uses shared hypernetworks conditioned on task information, while HyperTuning generates parameters from task examples Mahabadi et al. (2021); Phang et al. (2023). Meanwhile, MAC and Generative Adapter derive model modulations conditioned on textual contexts Tack et al. (2024); Chen et al. (2025). Combining parameterization with conditioning, Transformer^2 selects or combines learned parameters for singular components according to contextual information Sun et al. (2025).

3 The Proposed Method

PHISA integrates tissue structure and genomic information for multimodal survival prediction. Section 3.1 first models regional relationships with a polarized hypergraph and derives structural encodings for the pathology representation. Section 3.2 then aggregates these encodings into the patient tissue descriptor to condition spectral adaptation of the pretrained biomedical LLM. Finally, the refined genomic representation is combined with the pathology representation for survival prediction.

3.1 Polarized Hypergraphs and Interference-Aware Structural Encodings

Tissue regions are represented as nodes in a hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{C}, \mathbf{W})$, with hyperedges grouping related regions. The non-negative incidence matrix $\mathbf{C} \in \mathbb{R}_{\geq 0}^{|\mathcal{V}|\times|\mathcal{E}|}$ describes structural associations, and $\mathbf{W}$ is a diagonal matrix of hyperedge weights. With diagonal node and hyperedge degree matrices

$\mathbf{D}_v$ and $\mathbf{D}_e$, the standard normalized hypergraph Laplacian is

$$\mathbf{L} = \mathbf{I} - \mathbf{D}_v^{-1/2}\mathbf{C}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{C}^\top\mathbf{D}_v^{-1/2}. \tag{1}$$

To further characterize regional relationships within each hyperedge, we consider a **polarized** hypergraph. For region v in hyperedge e, the orientation is represented by a phase $\phi_{ve} \in (-\pi, \pi)$. Specifically, we define the polarized incidence matrix $\tilde{\mathbf{C}} \in \mathbb{C}^{|\mathcal{V}|\times|\mathcal{E}|}$ as

$$\tilde{C}_{ve} = C_{ve}e^{i\phi_{ve}}, \quad \phi_{ve} = \pi\tanh(f_\phi([\mathbf{x}_v\|\mathbf{x}_e])), \tag{2}$$

where $\mathbf{x}_v$ is the feature representation of region v, $\mathbf{x}_e$ is the aggregated feature of regions in hyperedge e, and f_ϕ is a learnable projection. The resulting $\tilde{\mathbf{C}}$ defines the polarized hypergraph Laplacian $\tilde{\mathbf{L}}$ as

$$\tilde{\mathbf{L}} = \mathbf{I} - \mathbf{D}_v^{-1/2}\tilde{\mathbf{C}}\mathbf{W}\mathbf{D}_e^{-1}\tilde{\mathbf{C}}^*\mathbf{D}_v^{-1/2}, \tag{3}$$

where * denotes the conjugate transpose. Setting all incidence phases to zero recovers the standard Laplacian $\mathbf{L}$. With non-negative hyperedge weights, $\mathbf{I} - \tilde{\mathbf{L}}$ is Hermitian positive semidefinite. Consequently, $\tilde{\mathbf{L}}$ is Hermitian and admits an orthonormal eigenbasis with real eigenvalues.

For regions u and v in hyperedge e, the incidence entries satisfy $\tilde{C}_{ue}\tilde{C}^*_{ve} = C_{ue}C_{ve}e^{i(\phi_{ue}-\phi_{ve})}$. Thus, the relative phase $\phi_{ue} - \phi_{ve}$ modulates the contribution of e to the (u, v) entry of $\mathbf{I} - \tilde{\mathbf{L}}$.

Assumption 1 (Opposing Regional Features) *Let u and v be regions in hyperedge e. Denote the contribution of e to the transition $\mathbf{A} = \mathbf{I} - \mathbf{L}$ by $\mathbf{A}^{(e)}$. Assume equal local coefficients $A^{(e)}_{uu} = A^{(e)}_{uv} = a > 0$ and opposing features $\mathbf{x}_u = -\mathbf{x}_v \neq \mathbf{0}$.*

Under Assumption 1, the contributions of u and v to the response at u through hyperedge e are

$$\Delta\mathbf{x}_u = A^{(e)}_{uu}\mathbf{x}_u + A^{(e)}_{uv}\mathbf{x}_v = a\mathbf{x}_u - a\mathbf{x}_u = \mathbf{0}. \tag{4}$$

By contrast, polarization preserves the diagonal coefficient a and changes the off-diagonal coefficient to $ae^{i\Delta\phi_{uv}}$. The resulting response is $\Delta\tilde{\mathbf{x}}_u = a\mathbf{x}_u + ae^{i\Delta\phi_{uv}}\mathbf{x}_v$, where $\Delta\phi_{uv} = \phi_{ue} - \phi_{ve}$.

Theorem 2 (Response Energy under Phase Interference) *For the configuration in Assumption 1, the squared norm of the polarized local response is*

$$\|\Delta\tilde{\mathbf{x}}_u\|^2 = 2a^2\|\mathbf{x}_u\|^2\big(1 - \cos(\Delta\phi_{uv})\big) > 0, \tag{5}$$

provided that $\Delta\phi_{uv} \neq 0$.

A nonzero relative phase thus yields a nonzero polarized response by altering the cancellation between opposing feature contributions. The proof is provided in Appendix A.

Regional features are propagated using the polarized transition:

$$\tilde{\mathbf{A}} = \mathbf{I} - \tilde{\mathbf{L}} = \mathbf{D}_v^{-1/2}\tilde{\mathbf{C}}\mathbf{W}\mathbf{D}_e^{-1}\tilde{\mathbf{C}}^*\mathbf{D}_v^{-1/2}, \quad \mathbf{X}^{(l+1)} = \mathrm{modReLU}\Big(\tilde{\mathbf{A}}\mathbf{X}^{(l)}\mathbf{\Theta}^{(l)}\Big), \tag{6}$$

where $\mathbf{X}^{(l)}$ contains the representations of tissue regions at layer l, with $\mathbf{X}^{(0)}$ formed by stacking $\{\mathbf{x}_v\}_{v\in\mathcal{V}}$. Relative phases in $\tilde{\mathbf{A}}$ modulate feature aggregation, and $\mathbf{\Theta}^{(l)}$ transforms the aggregated features. The transformed features are then activated by $\mathrm{modReLU}(\cdot)$ Trabelsi et al. (2018).

We further introduce structural encodings to characterize the relationships among tissue regions. The global construction begins with a structural seed $r_v = \sqrt{\tilde{A}_{vv}}$, normalized as $\bar{\mathbf{r}} = \mathbf{r}/\|\mathbf{r}\|_2$. Stacking the responses at successive propagation orders forms a complex Krylov matrix:

$$\mathbf{K}_d = \left[\tilde{\mathbf{A}}\bar{\mathbf{r}}, \left(\tilde{\mathbf{A}}\right)^2\bar{\mathbf{r}}, \ldots, \left(\tilde{\mathbf{A}}\right)^d\bar{\mathbf{r}}\right], \qquad \mathbf{S} = \mathrm{qf}(\mathbf{K}_d), \tag{7}$$

where $\mathrm{qf}(\cdot)$ returns an orthonormal basis for the column space of $\mathbf{K}_d$. Each column of $\mathbf{S}$ defines a structural coordinate across tissue regions. Global phase ambiguity of each structural coordinate is resolved by deterministic phase normalization.

The global construction admits a spectral interpretation. For eigenpairs $(\lambda_j, \mathbf{u}_j)$ of $\tilde{\mathbf{L}}$ and $c_j = \mathbf{u}_j^* \bar{\mathbf{r}}$, the structural seed at propagation order m is $\tilde{\mathbf{A}}^m \bar{\mathbf{r}} = \sum_j (1-\lambda_j)^m \mathbf{u}_j c_j$. For $0 \le \lambda_j < 1$, smaller eigenvalues yield more slowly decaying spectral weights as m increases. The following proposition bounds normalized eigenvector differences after accounting for relative phase.

Proposition 3 (Spectral Bound on Regional Discrepancy) *Let $\mathbf{u}_j$ be a unit-norm eigenvector of $\tilde{\mathbf{L}}$ with eigenvalue λ_j. For regions u and v in hyperedge e with $w_e C_{ue} C_{ve} > 0$,*

$$\left| \frac{u_{j,u}}{\sqrt{d_u}} - \frac{u_{j,v}}{\sqrt{d_v}} e^{i\Delta\phi_{uv}} \right|^2 \le \frac{2 d_e \lambda_j}{w_e C_{ue} C_{ve}}, \tag{8}$$

where $w_e = W_{ee}$ and $\Delta\phi_{uv} = \phi_{ue} - \phi_{ve}$.

The proof is provided in Appendix B. For fixed regions u, v, and hyperedge e, Eq. (8) shows that smaller eigenvalues yield tighter discrepancy bounds. This links slower spectral decay to tighter regional discrepancy bounds.

Complementing the global coordinates, local structure is encoded by the polarized diffusion operator:

$$\mathbf{P} = \mathbf{D}_v^{-1} \tilde{\mathbf{C}} \mathbf{W} \mathbf{D}_e^{-1} \tilde{\mathbf{C}}^* = \mathbf{D}_v^{-1/2} \tilde{\mathbf{A}} \mathbf{D}_v^{1/2}. \tag{9}$$

At propagation order k, the diagonal of $\mathbf{P}^k$ gives the local structural responses:

$$\mathbf{Z}_v^{(k)} = \left(\mathbf{P}^k\right)_{vv} = \left(\tilde{\mathbf{A}}^k\right)_{vv} \in \mathbb{R}_{\ge 0}. \tag{10}$$

Each regional response aggregates transition coefficient products over anchored closed walks.

Accordingly, combining the local responses for propagation orders $1, \ldots, K$ with d global coordinates yields the Dual-Scale Interference-Aware Structural Encoding (DS-IASE):

$$\mathbf{E}_{DS\text{-}IASE}(v) = \Big[\underbrace{|\mathbf{S}_v^{(1)}|, \angle\mathbf{S}_v^{(1)}, \ldots, |\mathbf{S}_v^{(d)}|, \angle\mathbf{S}_v^{(d)}}_{\text{Macro}}, \underbrace{\mathbf{Z}_v^{(1)}, \ldots, \mathbf{Z}_v^{(K)}}_{\text{Micro}} \Big]. \tag{11}$$

Macro denotes the magnitudes and phases of the global structural coordinates, whereas Micro denotes the local structural responses. The resulting structural encoding is combined with the propagated regional representations from Eq. (6) to form the pathology representation.

3.2 Spectral Adaptation of Biomedical LLMs Guided by Tissue Structure

The regional structural encodings from Section 3.1 provide tissue context for refining genomic representations with a pretrained biomedical LLM. These encodings define the patient tissue descriptor:

$$\mathbf{\Psi}_{DS} = \rho(\{\mathbf{E}_{DS\text{-}IASE}(v) \mid v \in \mathcal{V}\}) \in \mathbb{R}^{d_\Psi}, \tag{12}$$

where $\rho(\cdot)$ denotes aggregation over tissue regions.

Genomic information is represented by $\mathbf{H}_{gene}$, with gene expression organized into functional groups. Projection $\mathcal{P}_{in}$ maps $\mathbf{H}_{gene}$ to the LLM embedding space, yielding $\mathbf{H}^{(0)} = \mathcal{P}_{in}(\mathbf{H}_{gene})$.

We introduce Topology-Coherent Spectral Adaptation (TCSA), in which $\mathbf{H}^{(0)}$ serves as genomic input and $\mathbf{\Psi}_{DS}$ conditions the adaptation. TCSA performs adaptation within the principal singular subspace of selected pretrained projection weights Meng et al. (2024). For such a projection matrix $\mathbf{W}_0 \in \mathbb{R}^{d_{\text{out}} \times d_{\text{in}}}$, retaining its r principal singular components yields the decomposition:

$$\mathbf{W}_0 = \mathbf{W}_{res} + \mathbf{U}_0 \operatorname{diag}(\boldsymbol{\sigma}_0) \mathbf{V}_0^\top. \tag{13}$$

Here, $\mathbf{U}_0 \in \mathbb{R}^{d_{\text{out}} \times r}$ and $\mathbf{V}_0 \in \mathbb{R}^{d_{\text{in}} \times r}$ contain the principal left and right singular directions. The vector $\boldsymbol{\sigma}_0 \in \mathbb{R}^r$ contains the corresponding singular values, and $\mathbf{W}_{res}$ denotes the residual pretrained component. These pretrained quantities remain fixed during adaptation, while the patient tissue descriptor modulates the spectral coefficients:

$$\boldsymbol{\sigma}_{mod} = \boldsymbol{\sigma}_0 \odot \mathbf{s}_{tune} \odot \cos(\mathbf{W}_\phi \mathbf{\Psi}_{DS}), \tag{14}$$

where $\mathbf{W}_\phi \in \mathbb{R}^{r\times d_\Psi}$ maps $\mathbf{\Psi}_{DS}$ to r modulation arguments, and $\mathbf{s}_{tune} \in \mathbb{R}^r$ scales the spectral coefficients. Both $\mathbf{W}_\phi$ and $\mathbf{s}_{tune}$ are learned and shared across patients.

The resulting spectral coefficient change induces the weight update:

$$\Delta\boldsymbol{\sigma} = \boldsymbol{\sigma}_{mod} - \boldsymbol{\sigma}_0, \qquad \Delta\mathbf{W} = \mathbf{U}_0 \,\mathrm{diag}(\Delta\boldsymbol{\sigma})\mathbf{V}_0^\top. \tag{15}$$

Applying the adapted projection to an input row vector $\mathbf{x} \in \mathbb{R}^{1\times d_{\text{in}}}$, the adapted output is

$$\mathbf{y} = \mathbf{x}\mathbf{W}_0^\top + \mathbf{x}\Delta\mathbf{W}^\top = \mathbf{x}\mathbf{W}_0^\top + \mathbf{x}\mathbf{V}_0 \,\mathrm{diag}(\Delta\boldsymbol{\sigma})\mathbf{U}_0^\top. \tag{16}$$

The orthonormal singular directions preserve perturbation magnitude.

Proposition 4 (Spectral Perturbation within a Fixed Subspace) *Given the column orthonormality $\mathbf{U}_0^\top\mathbf{U}_0 = \mathbf{I}$ and $\mathbf{V}_0^\top\mathbf{V}_0 = \mathbf{I}$, the induced weight perturbation satisfies*

$$\|\Delta\mathbf{W}\|_F^2 = \|\Delta\boldsymbol{\sigma}\|_2^2 = \left\|\boldsymbol{\sigma}_{mod} - \boldsymbol{\sigma}_0\right\|_2^2. \tag{17}$$

The proof is provided in Appendix C. The Frobenius norm of the update is determined by $\Delta\boldsymbol{\sigma}$.

With the adapted projections, the LLM maps $\mathbf{H}^{(0)}$ to $\mathbf{H}_{LLM}$. Projection $\mathcal{P}_{out}$ maps $\mathbf{H}_{LLM}$ back to the genomic feature space, yielding the residual update:

$$\Delta\mathbf{H}_{gene} = \mathcal{P}_{out}(\mathbf{H}_{LLM}), \quad \widetilde{\mathbf{H}}_{gene} = \mathbf{H}_{gene} + \Delta\mathbf{H}_{gene}. \tag{18}$$

Throughout TCSA adaptation, the pretrained LLM parameters and the previously learned survival model remain fixed. The refined genomic representation $\widetilde{\mathbf{H}}_{gene}$ is combined with the pathology representation for survival prediction. For patient i, the model predicts interval hazards h_{ik}, with survival probability $S_{ik} = \prod_{m=1}^{k}(1-h_{im})$. Training minimizes the weighted negative log-likelihood:

$$\mathcal{L}_{\text{surv}} = -\sum_i \left[(1-c_i)\left(\log S_{i,y_i-1} + \log h_{i,y_i}\right) + (1-\alpha)c_i \log S_{i,y_i}\right], \tag{19}$$

where y_i denotes the interval containing the observed time, $c_i = 1$ indicates censoring, $S_{i0} = 1$, and α controls the weighting of censored patients. Further training details are provided in Appendix D.

4 Experiment

Experiments evaluate PHISA for survival prediction using tissue structure to guide molecular representation learning. Experimental setup and overall evaluation are presented in Section 4.1. Building on this evaluation, ablation studies examine the proposed designs (Section 4.2), followed by analysis of tissue conditioning during model adaptation (Section 4.3). The evaluation then extends to spatial visualization of tissue region importance and patient risk stratification (Section 4.4).

4.1 Experimental Setup and Overall Evaluation

Datasets and Protocols. We evaluate PHISA on four TCGA cohorts: COAD (n=461, 78.74% censored), CESC (n=307, 76.22%), READ (n=172, 83.72%), and LUAD (n=585, 63.93%). Patients have paired WSIs and transcriptomic profiles. WSIs are represented as polarized hypergraphs, and transcriptomic profiles are organized into functional groups. We use stratified five-fold cross-validation with the same folds across three random seeds. All slides from each patient remain in the same fold, with event status used for stratification. Performance is evaluated using the Concordance Index (C-index) and Integrated Brier Score (IBS). Implementation details are provided in Appendix D.

Baselines. We benchmark PHISA against representative survival methods across three categories. **(1) Classical Survival Modeling (CSM):** CoxPH Fox & Weisberg (2002) and DeepSurv Katzman et al. (2018) provide proportional hazards baselines with linear and neural risk functions, respectively. **(2) Unimodal Pathology Representation (UPR):** CLAM Lu et al. (2021) represents WSIs through attention-based MIL, while TMEGL Shao et al. (2024), SGCT Shi et al. (2025), and I2HGC Han et al.

(2025) use graph or hypergraph relations to capture structural organization in pathology. **(3) Multimodal Survival Learning (MSL):** MCAT Chen et al. (2021) and MOTCat Xu & Chen (2023) model cross-modal interactions between histology and genomic representations, while SurvPath Jaume et al. (2024) organizes transcriptomic information by biological pathway to model interactions with pathology. HGSNet Ding et al. (2025) and MRePath Qu et al. (2025b) use hypergraphs to model pathology structure and combine it with molecular information for survival prediction.

Table 1: Survival prediction performance on TCGA cohorts. Results are reported as mean ± standard deviation across three random seeds, each averaged over the same five folds.

	Method	C-index ↑					IBS ↓				
		COAD	CESC	READ	LUAD	**Mean**	COAD	CESC	READ	LUAD	**Mean**
CSM	CoxPH	0.702±0.068	0.549±0.092	0.668±0.073	0.559±0.069	0.620	0.195±0.056	0.307±0.066	0.220±0.061	0.292±0.079	0.254
	DeepSurv	0.715±0.059	0.562±0.081	0.687±0.068	0.561±0.054	0.631	0.180±0.047	0.290±0.061	0.209±0.059	0.289±0.062	0.242
UPR	CLAM	0.722±0.056	0.573±0.077	0.707±0.065	0.578±0.066	0.645	0.174±0.046	0.282±0.057	0.193±0.060	0.283±0.075	0.233
	TMEGL	0.745±0.057	0.603±0.076	0.731±0.064	0.587±0.063	0.667	0.157±0.037	0.266±0.068	0.162±0.048	0.276±0.065	0.215
	SGCT	0.727±0.065	0.571±0.074	0.709±0.067	0.575±0.063	0.646	0.172±0.051	0.281±0.055	0.192±0.063	0.280±0.071	0.231
	I2HGC	0.754±0.062	0.645±0.053	0.760±0.063	0.618±0.048	0.694	0.153±0.058	0.232±0.045	0.151±0.051	0.255±0.061	0.198
MSL	MCAT	0.743±0.069	0.629±0.079	0.737±0.059	0.586±0.065	0.673	0.159±0.042	0.236±0.052	0.160±0.056	0.276±0.068	0.208
	MOTCat	0.748±0.059	0.638±0.084	0.742±0.076	0.594±0.065	0.681	0.156±0.053	0.242±0.060	0.158±0.070	0.271±0.064	0.206
	SurvPath	0.759±0.053	0.651±0.063	0.743±0.058	0.599±0.052	0.688	0.152±0.050	0.230±0.054	0.158±0.052	0.269±0.064	0.202
	HGSNet	0.754±0.051	0.636±0.068	0.767±0.062	0.604±0.041	0.690	0.153±0.056	0.240±0.062	0.149±0.065	0.266±0.072	0.202
	MRePath	0.776±0.049	0.654±0.072	0.773±0.052	0.624±0.059	0.707	0.143±0.044	0.228±0.048	0.144±0.058	0.249±0.069	0.191
	PHISA	**0.801**±0.056	**0.673**±0.069	**0.875**±0.062	**0.650**±0.047	**0.750**	**0.128**±0.038	**0.219**±0.052	**0.112**±0.047	**0.240**±0.058	**0.175**

Overall Performance. Table 1 shows that PHISA achieves the highest mean C-index and lowest mean IBS on all TCGA cohorts, with averages across cohorts of 0.750 and 0.175. The gains cover both patient risk ranking and survival probability estimation. I2HGC achieves the strongest unimodal mean C-index (0.694), exceeding several multimodal baselines. Among the evaluated baselines, MRePath achieves the best average performance on both metrics; PHISA improves its mean C-index by 0.043 while reducing mean IBS by 0.016. Paired comparisons with MRePath across matched folds are provided in Appendix E.1. The largest gains over MRePath occur on READ, where the C-index increases from 0.773 to 0.875 and IBS decreases from 0.144 to 0.112. On COAD, CESC, and LUAD, the C-index improves by 0.019–0.026 and IBS decreases in each cohort.

4.2 Ablation Study

To examine the contributions of PHISA's designs, we first analyze propagation and structural encoding in the tissue representation, followed by tissue-guided genomic refinement.

Polarized Propagation and Structural Encoding. Real and polarized propagation are compared across combinations of Macro and Micro in DS-IASE (Eq. (11)). Macro represents global structural coordinates, whereas Micro captures local closed-walk responses. As shown in Table 2, combining both encodings improves the mean C-index by 0.010 under real propagation and by 0.032 under polarized propagation. Within the polarized setting, Macro+Micro achieves a mean C-index of 0.720, compared with 0.704 for Macro alone and 0.710 for Micro alone. This full polarized Macro+Micro configuration is used as the DS-IASE baseline in Table 3.

Model Adaptation Guided by Tissue Structure. Starting from the full DS-IASE configuration, we examine tissue-guided adaptation. The "Uncond." variant retains LLM-based genomic refinement but removes patient-specific structural conditioning. As shown in Table 3, this increases the mean C-index from 0.720 to 0.728, indicating a benefit from genomic representation refinement. Conditioning spectral adaptation on the patient-specific descriptor Ψ_{DS} further increases the mean C-index to 0.750, indicating additional benefit from tissue-guided refinement. To test patient–descriptor correspondence, "Shuffled" retains structural conditioning but permutes descriptors across patients. Shuffling reduces the mean C-index to 0.717, below both "Uncond." and matched PHISA. Together, these results support matching each patient with its own tissue descriptor during genomic refinement.

4.3 Analysis of Biomedical LLM Adaptation

The system-level ablation in Table 3 supports the predictive value of conditioning adaptation on patient tissue structure. Building on this result, we further compare adaptation strategies and assess the sensitivity of TCSA to patient-specific structural conditioning.

Table 2: Ablation of polarized propagation and structural encoding. Macro and Micro denote the global structural coordinates in Eq. (7) and local structural responses in Eq. (10). Results are reported as mean ± standard deviation across three random seeds, each averaged over the same five folds.

Macro	Micro	Real					Polarized				
		COAD	CESC	READ	LUAD	**Mean**	COAD	CESC	READ	LUAD	**Mean**
×	×	0.748±0.061	0.631±0.072	0.755±0.066	0.608±0.052	0.685	0.750±0.059	0.633±0.068	0.759±0.067	0.610±0.050	0.688
✓	×	0.752±0.058	0.633±0.068	0.758±0.069	0.611±0.048	0.688	0.768±0.052	0.645±0.065	0.781±0.058	0.622±0.045	0.704
×	✓	0.755±0.055	0.638±0.070	0.761±0.063	0.613±0.049	0.691	0.771±0.054	0.649±0.062	0.795±0.060	0.625±0.048	0.710
✓	✓	0.759±0.057	0.640±0.071	0.765±0.065	0.618±0.051	0.695	**0.780**±0.049	**0.658**±0.061	**0.810**±0.055	**0.635**±0.042	**0.720**

Table 3: **Ablation and adaptation analysis of PHISA.** "Uncond." removes structural conditioning, while "Shuffled" uses permuted patient-level structural descriptors. The adaptation comparison reports PEFT parameter counts and the mean C-index across cohorts.

System-level Ablation						Adaptation Strategy		
Variant	**COAD**	**CESC**	**READ**	**LUAD**	**Mean**	**Method**	**PEFT Params**	**Mean C-index ↑**
DS-IASE	0.780±0.049	0.658±0.061	0.810±0.055	0.635±0.042	0.720	LoRA	27.26M	0.719
Uncond.	0.785±0.052	0.663±0.065	0.822±0.058	0.641±0.046	0.728	PiSSA	27.26M	0.728
Shuffled	0.776±0.058	0.655±0.068	0.805±0.066	0.631±0.051	0.717	SVFT	262.14K	0.740
PHISA	**0.801±0.056**	**0.673±0.069**	**0.875±0.062**	**0.650±0.047**	**0.750**	**TCSA**	**36.86K**	**0.750**

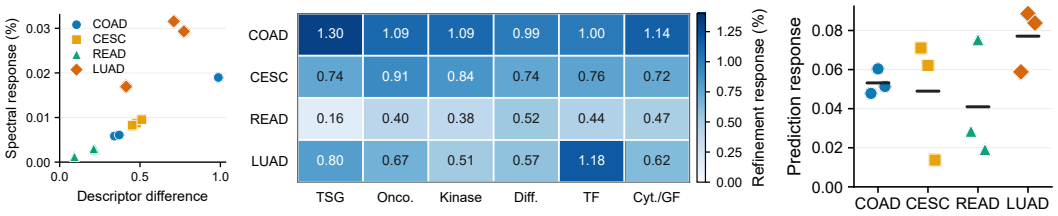
Comparison with Parameter-Efficient Adaptation. LoRA, PiSSA, SVFT, and TCSA are compared using the same BioMistral checkpoint and frozen upstream multimodal model. All strategies adapt the query and value projections with rank $r = 64$. The genomic projections and residual readout follow the same training setup, data partitions, and survival objectives. LoRA, PiSSA, and SVFT learn task-level adaptation parameters, whereas TCSA additionally conditions its spectral coefficients on patient tissue structure. Table 3 summarizes PEFT parameter counts and mean C-index; complete cohort-wise results and total trainable parameter counts are provided in Appendix E.2. TCSA achieves the highest mean C-index among the evaluated strategies. Compared with SVFT, the strongest alternative, TCSA increases the mean from 0.740 to 0.750 using 36.86K PEFT parameters versus 262.14K for SVFT. Complete cohort-wise results show higher means on all four cohorts.

Sensitivity to Structural Conditioning. TCSA uses the patient tissue descriptor to condition spectral modulation. To assess sensitivity, we perform descriptor substitutions in cases from all four TCGA cohorts. For each case, the model parameters and pathology and transcriptomic inputs are held fixed. Only the structural descriptor supplied to TCSA is replaced with that of another patient from the same cohort. Responses are measured relative to the outputs obtained using the patient's own descriptor. Figure 2a shows that descriptor substitution changes the spectral adaptation state. The corresponding residual refinement exhibits different relative responses across the six transcriptomic functional groups (Figure 2b). In the illustrated cases, the tumor suppressor group shows a comparatively large response in COAD, whereas the transcription factor group is largest in LUAD. Prediction logits also change under descriptor substitution (Figure 2c). Together, these observations trace the response to tissue conditioning from spectral modulation through genomic refinement to prediction.

4.4 Patient Stratification and Spatial Interpretability

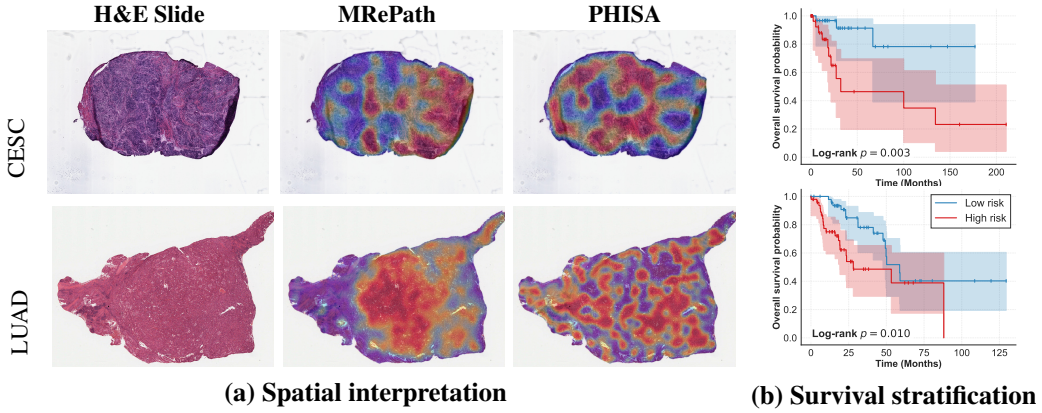
Building on the analysis of tissue conditioning, the complete PHISA model is further examined through spatial importance patterns and prognostic stratification.

Spatial Distribution of Tissue Region Importance. Figure 3(a) presents representative H&E slides and spatial importance maps from MRePath and PHISA for CESC and LUAD. In LUAD, MRePath emphasizes a largely continuous central area, whereas PHISA highlights spatially separated regions of higher importance. The spatially heterogeneous weighting is consistent with PHISA's modeling of differentiated regional structure. In CESC, the methods differ mainly in the location and extent of regions with higher importance. Complete spatial comparisons across all four cohorts are provided in Appendix E.3. Appendix F quantifies polarized propagation through a matched zero-phase comparison and the dispersion of closed-walk responses across propagation steps.



(a) (b) (c)

Figure 2: Responses to structural descriptor substitution in a fixed model. (a) Spectral response versus descriptor difference. (b) Relative residual refinement responses across six transcriptomic functional groups. (c) Changes in prediction logits.



(a) Spatial interpretation **(b) Survival stratification**

Figure 3: **Spatial interpretation and prognostic stratification.** (a) Representative H&E slides and relative importance maps for CESC and LUAD; warmer colors indicate higher relative importance, with PHISA showing more spatially separated regions with higher relative importance in LUAD. (b) Kaplan–Meier curves for groups split by median predicted risk; shaded regions denote confidence intervals and p-values are from log-rank tests.

Prognostic Stratification. PHISA uses patient tissue context to refine molecular representations, with the resulting risk scores evaluated for prognostic stratification within each cohort. Patients are divided into high- and low-risk groups using the median predicted risk score within each cohort. Figure 3 (b) presents Kaplan–Meier curves for CESC and LUAD, showing lower survival probabilities in the high-risk groups and significant log-rank separation. In CESC, survival declines more rapidly in the high-risk group during early follow-up, while the low-risk group maintains higher survival probabilities over extended intervals, resulting in sustained separation. In LUAD, survival declines earlier in the high-risk group, with clearer separation during early and intermediate follow-up and closer curves over part of the later interval. Complete analyses across all four cohorts are provided in Appendix E.4. By relating predicted risk groups to observed survival distributions, the Kaplan–Meier analysis complements the individual risk ranking evaluated by the C-index in Table 1.

5 Conclusion

PHISA integrates polarized hypergraph modeling of tissue relationships with spectral adaptation of a pretrained biomedical LLM guided by patient tissue structure. The framework supports multimodal survival prediction across four TCGA cohorts and improves performance over the evaluated baselines. Ablation and descriptor substitution analyses support the contributions of polarized tissue modeling and molecular refinement guided by patient tissue structure. Spatial importance and prognostic stratification provide complementary views of tissue patterns and patient survival differences. The current evaluation is limited to TCGA cohorts and one biomedical language model, while broader evaluation would require additional data and computational resources. Future work will broaden the evaluation to independent cohorts and additional biomedical language models.

AI USE STATEMENT

In this work, we used generative AI tools to assist with revising and polishing the manuscript, including improvements to language, clarity, and presentation. We reviewed all AI-assisted content and take responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

Implementation and training details are provided in Appendix D, including data preprocessing, model configuration, and optimization settings. Code for the proposed method is available in an anonymous repository at `https://anonymous.4open.science/r/PHISA/`.

REFERENCES

Siemen Brussee, Giorgio Buzzanca, Anne MR Schrader, and Jesper Kers. Graph neural networks in histopathology: Emerging trends and future directions. *Medical Image Analysis*, 101:103444, 2025.

Richard J. Chen, Ming Y. Lu, Wei-Hung Weng, Tiffany Y. Chen, Drew FK. Williamson, Trevor Manz, Maha Shady, and Faisal Mahmood. Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3995–4005, 2021.

Richard J Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Jana Lipkova, Zahra Noor, Muhammad Shaban, Maha Shady, Mane Williams, Bumjin Joo, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer cell*, 40(8):865–878, 2022.

Tong Chen, Hao Fang, Patrick Xia, Xiaodong Liu, Ben Van Durme, Luke Zettlemoyer, Jianfeng Gao, and Hao Cheng. Generative adapter: Contextualizing language models in parameters with a single forward pass. In *International Conference on Learning Representations*, volume 2025, pp. 27450–27470, 2025.

Eli Chien, Chao Pan, Jianhao Peng, and Olgica Milenkovic. You are allset: A multiset function framework for hypergraph neural networks. In *International Conference on Learning Representations (ICLR)*, 2022.

Taane G Clark, Michael J Bradburn, Sharon B Love, and Douglas G Altman. Survival analysis part i: basic concepts and first analyses. *British journal of cancer*, 89(2):232–238, 2003.

Esther Danenberg, Helen Bardwell, Vito RT Zanotelli, Elena Provenzano, Suet-Feung Chin, Oscar M Rueda, Andrew Green, Emad Rakha, Samuel Aparicio, Ian O Ellis, et al. Breast tumor microenvironment structures are associated with genomic features and clinical outcome. *Nature genetics*, 54 (5):660–669, 2022.

Saisai Ding, Linjin Li, Ge Jin, Jun Wang, Shihui Ying, and Jun Shi. Hgmsurvnet: A two-stage hypergraph learning network for multimodal cancer survival prediction. *Medical Image Analysis*, 104:103661, 2025. ISSN 1361-8415.

Iulia Duta, Giulia Cassarà, Fabrizio Silvestri, and Pietro Liò. Sheaf hypergraph networks. *Advances in Neural Information Processing Systems*, 36:12087–12099, 2023.

John Fox and Sanford Weisberg. Cox proportional-hazards regression for survival data. *An R and S-PLUS companion to applied regression*, 2002:7, 2002.

Xiangmin Han, Huijian Zhou, Zhiqiang Tian, Shaoyi Du, and Yue Gao. Inter-intra hypergraph computation for survival prediction on whole slide images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(7):6006–6021, 2025.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*, 2022.

Yongfei Hu, Xinyu Li, Ying Yi, Yan Huang, Guangyu Wang, and Dong Wang. Deep learning-driven survival prediction in pan-cancer studies by integrating multimodal histology-genomic data. *Briefings in Bioinformatics*, 26(2):bbaf121, 2025.

Guillaume Jaume, Anurag Vaidya, Richard J. Chen, Drew F.K. Williamson, Paul Pu Liang, and Faisal Mahmood. Modeling dense multimodal interactions between biological pathways and histology for survival prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11579–11590, June 2024.

Justin Jee, Christopher Fong, Karl Pichotta, Thinh Ngoc Tran, Anisha Luthra, Michele Waters, Chenlian Fu, Mirella Altoe, Si-Yang Liu, Steven B Maron, et al. Automated real-world data integration improves cancer outcome prediction. *Nature*, 636(8043):728–736, 2024.

Nikhil S Joshi, Elliot H Akama-Garren, Yisi Lu, Da-Yae Lee, Gregory P Chang, Amy Li, Michel DuPage, Tuomas Tammela, Natanya R Kerper, Anna F Farago, et al. Regulatory t cells in tumor-associated tertiary lymphoid structures suppress anti-tumor t cell responses. *Immunity*, 43(3): 579–590, 2015.

Jared L Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. *BMC medical research methodology*, 18(1):24, 2018.

Yanis Labrak, Adrien Bazoge, Emmanuel Morin, Pierre-Antoine Gourraud, Mickael Rouvier, and Richard Dufour. Biomistral: A collection of open-source pretrained large language models for medical domains. In *Findings of the association for computational linguistics: acl 2024*, pp. 5848–5864, 2024.

Vijay Lingam, Atula Tejaswi, Aditya Vavre, Aneesh Shetty, Gautham K Gudur, Joydeep Ghosh, Alex Dimakis, Eunsol Choi, Aleksandar Bojchevski, and Sujay Sanghavi. Svft: Parameter-efficient fine-tuning with singular vectors. *Advances in neural information processing systems*, 37:41425–41446, 2024.

Fei Liu, Stephan Beck, Lei Yang, Huiyan Luo, and Kang Zhang. Advancing ai for multi-omics and clinical data integration in basic and translational cancer research. *Nature Reviews Cancer*, pp. 1–16, 2026.

Lifan Long, Yilun Li, Daoqiang Zhang, and Yan Wang. Mohd: Multi-modal survival prediction through hierarchical decoupling of whole-slide image pyramids and genomics. *Medical Image Analysis*, pp. 104098, 2026.

Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Richard J Chen, Matteo Barbieri, and Faisal Mahmood. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature biomedical engineering*, 5(6):555–570, 2021.

Rabeeh Karimi Mahabadi, Sebastian Ruder, Mostafa Dehghani, and James Henderson. Parameter-efficient multi-task fine-tuning for transformers via shared hypernetworks. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 565–576, 2021.

Fanxu Meng, Zhaohui Wang, and Muhan Zhang. Pissa: Principal singular values and singular vectors adaptation of large language models. In *Advances in Neural Information Processing Systems*, volume 37, pp. 121038–121072, 2024.

Emanuele Mule, Stefano Fiorini, Antonio Purificato, Federico Siciliano, Stefano Coniglio, and Fabrizio Silvestri. Directional sheaf hypergraph networks: Unifying learning on directed and undirected hypergraphs. In *International Conference on Learning Representations*, volume 2026, pp. 81483–81512, 2026.

Jason Phang, Yi Mao, Pengcheng He, and Weizhu Chen. Hypertuning: Toward adapting large language models without back-propagation. In *International Conference on Machine Learning*, pp. 27854–27875. pmlr, 2023.

Mingcheng Qu, Guang Yang, Donglin Di, Yue Gao, Tonghua Su, Yang Song, and Lei Fan. Memory-augmented incomplete multimodal survival prediction via cross-slide and gene-attentive hypergraph learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 318–327. Springer, 2025a.

Mingcheng Qu, Guang Yang, Donglin Di, Tonghua Su, Yue Gao, Yang Song, and Lei Fan. Multimodal cancer survival analysis via hypergraph learning with cross-modality rebalance. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI-25*, pp. 1802–1810. International Joint Conferences on Artificial Intelligence Organization, 8 2025b.

Zahraa Rahal, Roy El Darzi, Seyed Javad Moghaddam, Tina Cascone, and Humam Kadara. Tumour and microenvironment crosstalk in nsclc progression and response to therapy. *Nature reviews Clinical oncology*, 22(7):463–482, 2025.

Wei Shao, YangYang Shi, Daoqiang Zhang, JunJie Zhou, and Peng Wan. Tumor micro-environment interactions guided graph learning for survival analysis of human cancers from whole-slide pathological images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11694–11703, June 2024.

Jun Shi, Tong Shu, Zhiguo Jiang, Wei Wang, Haibo Wu, and Yushan Zheng. Slide-based graph collaborative training for histopathology whole slide image analysis. *IEEE Transactions on Medical Imaging*, pp. 1–1, 2025.

Qi Sun, Edoardo Cetin, and Yujin Tang. Transformer-squared: Self-adaptive llms. In *International Conference on Learning Representations*, volume 2025, pp. 13878–13895, 2025.

Jihoon Tack, Jaehyung Kim, Eric Mitchell, Jinwoo Shin, Yee Whye Teh, and Jonathan Richard Schwarz. Online adaptation of language models with a memory of amortized contexts. *Advances in Neural Information Processing Systems*, 37:130109–130135, 2024.

Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, João Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J. Pal. Deep complex networks. In *International Conference on Learning Representations*, 2018.

Peihao Wang, Shenghao Yang, Yunyu Liu, Zhangyang Wang, and Pan Li. Equivariant hypergraph diffusion neural operators. In *International Conference on Learning Representations (ICLR)*, 2023.

Yingxue Xu and Hao Chen. Multimodal optimal transport-based co-attention transformer with global structure consistency for survival prediction. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 21184–21194, 2023.

Yilan Zhang, Yingxue Xu, Jianqi Chen, Fengying Xie, and Hao Chen. Prototypical information bottlenecking and disentangling for multimodal cancer survival prediction. In *International Conference on Learning Representations*, volume 2024, pp. 36854–36876, 2024.

A Proof of Response-Energy Retention via Phase Interference

This section provides the detailed derivation of Theorem 2 under the idealized regional configuration in Assumption 1. We compare the isolated local response induced by the real-valued transition with its polarized counterpart.

A.1 Setup and Real-Valued Baseline

Recall the real-valued transition operator

$$\mathbf{A} = \mathbf{D}_v^{-1/2}\mathbf{C}\mathbf{W}\mathbf{D}_e^{-1}\mathbf{C}^\top\mathbf{D}_v^{-1/2}. \tag{20}$$

Under Assumption 1, we isolate the contributions of regions u and v through hyperedge e. Since $A_{uu}^{(e)} = A_{uv}^{(e)} = a > 0$ and $\mathbf{x}_v = -\mathbf{x}_u$, the real-valued response is

$$\Delta\mathbf{x}_u = A_{uu}^{(e)}\mathbf{x}_u + A_{uv}^{(e)}\mathbf{x}_v = a\mathbf{x}_u - a\mathbf{x}_u = \mathbf{0}. \tag{21}$$

Thus, the squared response norm satisfies $\|\Delta\mathbf{x}_u\|^2 = 0$, corresponding to complete local cancellation under the idealized real-valued setting.

A.2 Derivation in the Complex Domain

We now analyze the corresponding local response $\Delta\tilde{\mathbf{x}}_u$ under the complex-valued polarized transition operator:

$$\tilde{\mathbf{A}} = \mathbf{D}_v^{-1/2}\tilde{\mathbf{C}}\mathbf{W}\mathbf{D}_e^{-1}\tilde{\mathbf{C}}^*\mathbf{D}_v^{-1/2}. \tag{22}$$

For the shared hyperedge e connecting nodes u and v, its contribution to the polarized transition is

$$\tilde{A}_{uv}^{(e)} = \frac{C_{ue}C_{ve}W_{ee}}{\sqrt{d_u d_v}\, d_e} e^{i(\phi_{ue}-\phi_{ve})} = A_{uv}^{(e)} e^{i\Delta\phi_{uv}}, \tag{23}$$

where $\Delta\phi_{uv} = \phi_{ue} - \phi_{ve}$ denotes the relative phase within the shared hyperedge. For the corresponding self-contribution, the phase difference vanishes, and therefore

$$\tilde{A}_{uu}^{(e)} = A_{uu}^{(e)}. \tag{24}$$

Using $A_{uu}^{(e)} = A_{uv}^{(e)} = a > 0$, the local response for node u is

$$\begin{aligned} \Delta\tilde{\mathbf{x}}_u &= \tilde{A}_{uu}^{(e)}\mathbf{x}_u + \tilde{A}_{uv}^{(e)}\mathbf{x}_v \\ &= a\mathbf{x}_u + ae^{i\Delta\phi_{uv}}\mathbf{x}_v. \end{aligned} \tag{25}$$

Substituting $\mathbf{x}_v = -\mathbf{x}_u$ gives

$$\Delta\tilde{\mathbf{x}}_u = a\mathbf{x}_u\left(1 - e^{i\Delta\phi_{uv}}\right). \tag{26}$$

A.3 Response Energy via Euler's Formula

The squared response norm is

$$\begin{aligned} \|\Delta\tilde{\mathbf{x}}_u\|^2 &= \left\|a\mathbf{x}_u\left(1 - e^{i\Delta\phi_{uv}}\right)\right\|^2 \\ &= a^2\|\mathbf{x}_u\|^2\left|1 - e^{i\Delta\phi_{uv}}\right|^2. \end{aligned} \tag{27}$$

Using Euler's formula, $e^{i\Delta\phi} = \cos(\Delta\phi) + i\sin(\Delta\phi)$, we obtain

$$\begin{aligned} \left|1 - e^{i\Delta\phi_{uv}}\right|^2 &= |(1 - \cos\Delta\phi_{uv}) - i\sin\Delta\phi_{uv}|^2 \\ &= (1 - \cos\Delta\phi_{uv})^2 + \sin^2\Delta\phi_{uv} \\ &= 2\big(1 - \cos(\Delta\phi_{uv})\big). \end{aligned} \tag{28}$$

Therefore,

$$\|\Delta\tilde{\mathbf{x}}_u\|^2 = 2a^2\|\mathbf{x}_u\|^2\big(1 - \cos(\Delta\phi_{uv})\big). \tag{29}$$

A.4 Conclusion of the Proof

Under the non-zero relative phase condition in Theorem 2, $\Delta\phi_{uv} \neq 0 \pmod{2\pi}$. Hence,

$$1 - \cos(\Delta\phi_{uv}) > 0.$$

Since $a > 0$ and $\|\mathbf{x}_u\| > 0$, it follows that

$$\|\Delta\tilde{\mathbf{x}}_u\|^2 > 0. \tag{30}$$

Thus, the polarized transition converts the complete local cancellation of the real-valued response into a non-zero phase-dependent response. This establishes Theorem 2.

B Proof of Spectral Bound on Regional Discrepancy

Proof. Let $\mathbf{u}_j$ be a unit-norm eigenvector of the normalized polarized hypergraph Laplacian $\tilde{\mathbf{L}}$ with eigenvalue λ_j. By the Rayleigh quotient,

$$\mathbf{u}_j^* \tilde{\mathbf{L}} \mathbf{u}_j = \lambda_j. \tag{31}$$

Expanding the normalized phase-aware Dirichlet energy gives

$$\lambda_j = \frac{1}{2} \sum_{e \in \mathcal{E}} \frac{w_e}{d_e} \sum_{u,v \in e} C_{ue} C_{ve} \left| \frac{u_{j,u}}{\sqrt{d_u}} - \frac{u_{j,v}}{\sqrt{d_v}} e^{i\Delta\phi_{uv}} \right|^2. \tag{32}$$

Since every term in Eq. (32) is non-negative, the contribution of any interacting pair (u, v) within hyperedge e satisfies

$$\frac{w_e C_{ue} C_{ve}}{2d_e} \left| \frac{u_{j,u}}{\sqrt{d_u}} - \frac{u_{j,v}}{\sqrt{d_v}} e^{i\Delta\phi_{uv}} \right|^2 \leq \lambda_j. \tag{33}$$

Rearranging yields

$$\left| \frac{u_{j,u}}{\sqrt{d_u}} - \frac{u_{j,v}}{\sqrt{d_v}} e^{i\Delta\phi_{uv}} \right|^2 \leq \frac{2d_e \lambda_j}{w_e C_{ue} C_{ve}}, \tag{34}$$

which establishes Proposition 3. Hence, low-frequency eigenvectors induce small normalized phase-aware discrepancies on strongly coupled interactions. □

C Proof of Spectral Perturbation within a Fixed Subspace

This section provides the proof of Proposition 4, which characterizes the weight perturbation induced by Topology-Coherent Spectral Adaptation (TCSA) within the selected principal singular subspace.

Proof. Recall that the pretrained weight matrix is decomposed as

$$\mathbf{W}_0 = \mathbf{W}_{res} + \mathbf{U}_0 \operatorname{diag}(\boldsymbol{\sigma}_0) \mathbf{V}_0^\top. \tag{35}$$

TCSA modifies only the spectral coefficients within the retained singular subspace. Defining

$$\Delta\boldsymbol{\sigma} = \boldsymbol{\sigma}_{mod} - \boldsymbol{\sigma}_0, \tag{36}$$

the induced weight perturbation is

$$\Delta\mathbf{W} = \mathbf{U}_0 \operatorname{diag}(\Delta\boldsymbol{\sigma}) \mathbf{V}_0^\top. \tag{37}$$

Hence, $\Delta\mathbf{W}$ is confined to the singular subspace spanned by $\mathbf{U}_0$ and $\mathbf{V}_0$.

Since the columns of $\mathbf{U}_0$ and $\mathbf{V}_0$ are orthonormal,

$$\mathbf{U}_0^\top \mathbf{U}_0 = \mathbf{I}, \qquad \mathbf{V}_0^\top \mathbf{V}_0 = \mathbf{I}. \tag{38}$$

The squared Frobenius norm of the induced weight perturbation therefore satisfies

$$\begin{aligned} \|\Delta\mathbf{W}\|_F^2 &= \operatorname{Tr}\big(\Delta\mathbf{W}^\top \Delta\mathbf{W}\big) \\ &= \operatorname{Tr}\big(\mathbf{V}_0 \operatorname{diag}(\Delta\boldsymbol{\sigma})^\top \mathbf{U}_0^\top \mathbf{U}_0 \operatorname{diag}(\Delta\boldsymbol{\sigma}) \mathbf{V}_0^\top\big) \\ &= \operatorname{Tr}\big(\operatorname{diag}(\Delta\boldsymbol{\sigma})^\top \operatorname{diag}(\Delta\boldsymbol{\sigma}) \mathbf{V}_0^\top \mathbf{V}_0\big) \\ &= \operatorname{Tr}\big(\operatorname{diag}(\Delta\boldsymbol{\sigma})^\top \operatorname{diag}(\Delta\boldsymbol{\sigma})\big) \\ &= \|\Delta\boldsymbol{\sigma}\|_2^2 \\ &= \big\|\boldsymbol{\sigma}_{mod} - \boldsymbol{\sigma}_0\big\|_2^2. \end{aligned} \tag{39}$$

This establishes Proposition 4.

Substituting the TCSA modulation

$$\boldsymbol{\sigma}_{mod} = \boldsymbol{\sigma}_0 \odot \mathbf{s}_{tune} \odot \cos(\mathbf{W}_\phi \boldsymbol{\Psi}_{DS}) \tag{40}$$

gives, for the j-th spectral coefficient,

$$\Delta\sigma_j = \sigma_{0,j} \left[s_{tune,j} \cos\Big((\mathbf{W}_\phi \boldsymbol{\Psi}_{DS})_j \Big) - 1 \right]. \tag{41}$$

Consequently,

$$\|\Delta \mathbf{W}\|_F^2 = \sum_{j=1}^{r} \sigma_{0,j}^2 \left[s_{tune,j} \cos\Big((\mathbf{W}_\phi \boldsymbol{\Psi}_{DS})_j \Big) - 1 \right]^2. \tag{42}$$

Thus, the magnitude of the induced weight perturbation is fully determined by the changes in the spectral coefficients within the fixed singular subspace. □

Table 4: **Baseline characteristics and survival statistics across four TCGA cohorts.** *N*: Total number of patients. *Age*: Median age in years (Interquartile Range). *Gender*: Number of female patients (percentage). *Censoring*: Percentage of right-censored subjects. *Follow-up*: Median follow-up time in months (Interquartile Range).

Dataset	Size	Patient Characteristics		Survival & Follow-up		
	N	Age (Years)	Female, n (%)	Events (n)	Censoring (%)	Median Time (IQR)
TCGA-COAD	461	67 (56–76)	216 (46.9%)	98	78.74	22.67 (13.52–37.53)
TCGA-CESC	307	47 (38–58)	307 (100.0%)	73	76.22	21.02 (12.51–44.58)
TCGA-READ	172	60 (53–71)	74 (43.0%)	28	83.72	21.25 (15.32–38.12)
TCGA-LUAD	585	66 (59–72)	328 (56.1%)	211	63.93	23.78 (14.06–38.21)

D Implementation Details

D.1 Datasets and Preprocessing

Cohorts and WSI Processing. Four TCGA datasets serve as the evaluation benchmark: COAD (n=461), CESC (n=307), READ (n=172), and LUAD (n=585). Table 4 summarizes the patient demographics and survival outcomes for each cohort. The CLAM framework facilitates the automated processing of raw whole-slide images. Tissue regions are segmented via Otsu thresholding to eliminate background components. The resulting tissue areas are partitioned into discrete 256×256 patches at $20\times$ magnification. We extract up to 8,192 patches per patient to ensure a balance between computational demand and pathological detail. A ResNet-50 model pretrained on ImageNet serves as the feature extractor. This backbone is chosen for its proven capability in identifying histological structures within pathology images. Feature vectors of 1024 dimensions are produced by the C4 stage of the model followed by adaptive average pooling. A trainable linear layer projects these high-dimensional features into a 512-dimensional space.

Omics and Clinical Data. Transcriptomic data are retrieved from the TCGA Pan-Cancer Atlas. Expression profiles undergo RSEM normalization and log2 transformation. Gene features are mapped to biological pathway signatures instead of using traditional variance-based filters. This procedure yields a biologically informed feature set containing 4,000 to 5,000 genes per cohort. Patient demographics and tumor stages are encoded via normalization for continuous variables and one-hot encoding for categorical features.

D.2 Polarized Hypergraph Construction

WSI Polarized Hypergraph. Each whole-slide image is represented as a hypergraph $\mathcal{G}$ to encapsulate morphological arrangements and semantic correlations among tissue patches. Nodes $v \in \mathcal{V}$

within the hypergraph correspond to the sampled histological regions from the segmented areas. Approximated k-nearest neighbor connections are established via the Hierarchical Navigable Small World (HNSW) algorithm to identify structural proximities in the feature space. Spatial relations are determined by the physical coordinates to maintain the integrity of the tissue architecture. Latent dependencies are derived from feature similarities to link distant yet phenotypically related regions. These associations are grouped into hyperedges $e \in \mathcal{E}$ representing localized multicellular niches to model high-order interactions. A mean pooling operation aggregates the features of patches $\mathbf{x}_v$ within each hyperedge to produce the niche representation $\mathbf{x}_e$. A polarization network maps the concatenation of node features $\mathbf{x}_v$ and niche features $\mathbf{x}_e$ to a phase value ϕ_{ve} through a tanh activation. This procedure generates the complex-valued incidence matrix $\tilde{\mathbf{C}}$ where magnitudes denote structural connections and phases encode regional orientations, with relative phases characterizing interaction tendencies.

Biologically Informed Genomic Representation. Transcriptomic information is organized into functional groups. Gene expression within each group is encoded into a group representation, and the resulting representations form the genomic representation $\mathbf{H}_{\text{gene}}$ used for multimodal survival modeling. During TCSA adaptation, $\mathbf{H}_{\text{gene}}$ is further refined through the pretrained biomedical LLM as described in Section 3.2.

D.3 Experimental Configuration

Training Protocol. All computational experiments are developed utilizing the PyTorch 2.2 and PyTorch Geometric libraries and executed on a local machine configured with one NVIDIA RTX 3090 GPU offering 24 GB of VRAM. A stratified 5-fold cross-validation scheme is adopted to maintain consistent event-status proportions across data splits. Optimization proceeds for a maximum limit of 30 epochs via the Rectified Adam algorithm. The baseline learning rate is initialized at 10^{-4} alongside a weight decay coefficient of 10^{-4}. To address the hardware memory bottleneck caused by extensive tissue patch sequences, parameter updates are performed by computing gradients for one complete patient bag per iteration. For each random seed, performance is averaged over the same five folds. Final results are reported as the mean and standard deviation across the three seed-specific estimates.

Large Language Model and Adaptation Configuration. BioMistral 7B is used as the pretrained biomedical LLM. Genomic features are projected into its embedding space before being processed by the LLM, and the LLM output is projected back to the genomic feature space for residual refinement.

Training follows a fold-specific two-stage procedure. For each cross-validation fold, the multimodal survival model is first trained using the corresponding training partition. TCSA adaptation is then initialized from the multimodal survival checkpoint obtained from the same fold. During this adaptation stage, the multimodal survival model and the pretrained BioMistral backbone remain fixed.

TCSA is applied to the query and value projection matrices in the self-attention modules. For each adapted pretrained weight matrix $\mathbf{W}_0$, singular value decomposition is used to obtain the principal singular directions $\mathbf{U}_0$ and $\mathbf{V}_0$ and the corresponding singular values $\boldsymbol{\sigma}_0$, with spectral rank $r = 64$. The modulated spectral coefficients are computed as

$$\boldsymbol{\sigma}_{mod} = \boldsymbol{\sigma}_0 \odot \mathbf{s}_{tune} \odot \cos(\mathbf{W}_\phi \boldsymbol{\Psi}_{DS}). \tag{43}$$

Optimization during TCSA adaptation is restricted to the genomic projection layers, $\mathbf{W}_\phi$, and $\mathbf{s}_{tune}$.

Across Uncond., Shuffled, and PHISA, the genomic input and output projections and spectral scaling parameters remain trainable under the same freezing protocol. In Uncond., modulation reduces to $\boldsymbol{\sigma}_0 \odot \mathbf{s}_{tune}$, with the structural conditioner inactive. The conditioning parameters $\mathbf{W}_\phi$ are trained in Shuffled and PHISA.

Visualization Details. Region importance in Figure 3 is obtained by summing attention weights from genomic queries to tissue regions over functional groups and averaging across heads. For visualization, scores are clipped to the 5th–95th percentile range, rescaled, and spatially smoothed before conversion to percentile ranks. Colors therefore indicate relative importance within each displayed slide.

Table 5: **Paired comparison between PHISA and MRePath across matched held-out folds.** Each fold-level result is averaged over three random seeds. Positive Δ_{C} indicates higher C-index for PHISA, while positive Δ_{IBS} indicates lower IBS for PHISA.

	C-index				IBS			
Cohort	Δ_{C}	**95% CI**	**Positive folds**	p	Δ_{IBS}	**95% CI**	**Positive folds**	p
COAD	+0.025	[0.019, 0.031]	5/5	< 0.001	+0.015	[0.011, 0.019]	5/5	< 0.001
CESC	+0.019	[0.015, 0.023]	5/5	< 0.001	+0.009	[0.003, 0.015]	5/5	0.011
READ	+0.102	[0.097, 0.107]	5/5	< 0.001	+0.032	[0.029, 0.035]	5/5	< 0.001
LUAD	+0.026	[0.020, 0.032]	5/5	< 0.001	+0.009	[0.005, 0.013]	5/5	0.002

For Figure 2, responses are measured relative to each patient's own descriptor. Descriptor difference is measured by Euclidean distance. Spectral response is the norm of the change in spectral coefficients divided by the norm of the pretrained coefficients, with both vectors concatenated across adapted projections. Functional group response is the norm of the change in residual refinement divided by its magnitude under the matched condition. Both responses are expressed as percentages. Prediction response is the Euclidean norm of the logit difference. Each illustrated case uses three donor descriptors from the same cohort; scatter points denote individual substitutions, while heatmap entries and horizontal bars report their means.

E Additional Experimental Results

This section provides additional experimental results supporting the main evaluation, including paired comparisons with MRePath, cohort-wise adaptation comparisons, spatial importance visualizations, and prognostic stratification across all four TCGA cohorts.

E.1 Paired Statistical Comparison with MRePath

We further assess the performance differences between PHISA and MRePath, the strongest multimodal baseline in Table 1, using matched held-out folds. The same five patient-level folds are used across methods, with each fold evaluated under three random seeds. Performance is first averaged across seeds within each fold, yielding five paired fold-level observations for each cohort. For fold k, the C-index difference is defined as

$$\Delta_{\mathrm{C},k} = C_{\mathrm{PHISA},k} - C_{\mathrm{MRePath},k},$$

and the IBS difference as

$$\Delta_{\mathrm{IBS},k} = IBS_{\mathrm{MRePath},k} - IBS_{\mathrm{PHISA},k}.$$

Positive values therefore indicate better PHISA performance for both metrics. We report the mean paired difference, its 95% confidence interval, the number of folds with positive differences, and the two-sided paired t-test p-value.

PHISA improves both C-index and IBS on all five matched folds in each of the four cohorts. For C-index, the mean paired gains are 0.025, 0.019, 0.102, and 0.026 on COAD, CESC, READ, and LUAD, respectively, with all 95% confidence intervals above zero. The corresponding IBS improvements are 0.015, 0.009, 0.032, and 0.009, and are likewise positive across all five folds for every cohort. The paired analysis complements the mean $\pm$ standard deviation results in Table 1 by showing that the observed performance gains are consistent across matched held-out folds.

E.2 Complete Comparison of Adaptation Strategies

Table 6 reports the complete cohort-wise adaptation comparison corresponding to the compact summary in Table 3. All methods use the same BioMistral checkpoint, data partitions, and adaptation setup.

TCSA achieves the highest mean C-index among the evaluated adaptation strategies, with the highest mean C-index on each of the four cohorts. It uses 36.86K trainable PEFT parameters, compared with

Table 6: **Complete comparison of adaptation strategies for BioMistral.** C-index is reported as the mean ± standard deviation across three random seeds.

Method	Trainable params		C-index ↑				
	PEFT	Total	COAD	CESC	READ	LUAD	Mean
LoRA	27.26M	29.36M	0.768 ± 0.061	0.651 ± 0.079	0.829 ± 0.071	0.628 ± 0.051	0.719
PiSSA	27.26M	29.36M	0.779 ± 0.060	0.656 ± 0.072	0.851 ± 0.064	0.626 ± 0.046	0.728
SVFT	262.14K	2.36M	0.786 ± 0.058	0.660 ± 0.066	0.867 ± 0.065	0.645 ± 0.044	0.740
TCSA	**36.86K**	**2.14M**	$\mathbf{0.801 \pm 0.056}$	$\mathbf{0.673 \pm 0.069}$	$\mathbf{0.875 \pm 0.062}$	$\mathbf{0.650 \pm 0.047}$	**0.750**

262.14K for SVFT and 27.26M for LoRA and PiSSA. The corresponding total trainable parameter count is 2.14M.

E.3 Complete Spatial Interpretation Across Four Cohorts

Figure 4 provides the complete four-cohort spatial comparison corresponding to the compact visualization in the main paper.

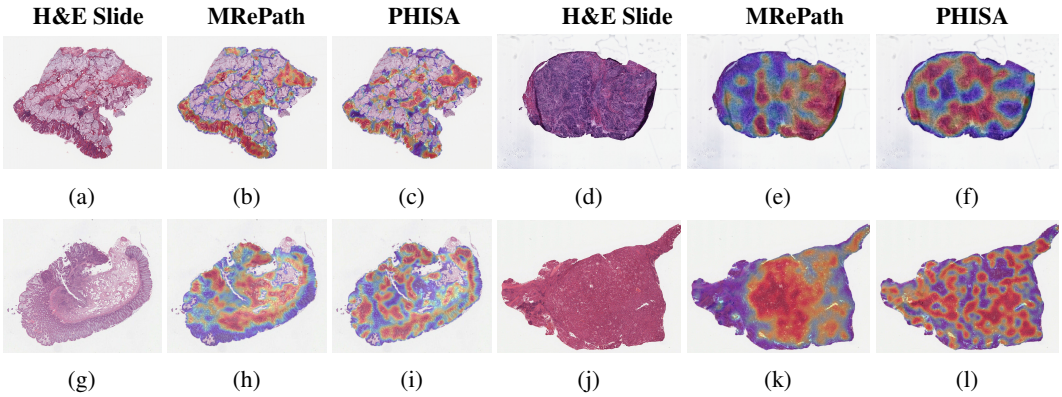


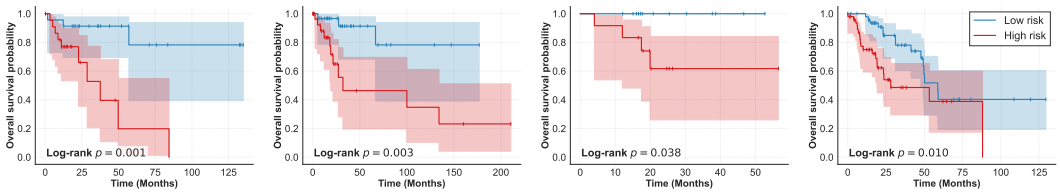
Figure 4: **Complete spatial importance visualization across four TCGA cohorts.** Representative H&E slides and corresponding importance maps from MRePath and PHISA are shown for (a–c) COAD, (d–f) CESC, (g–i) READ, and (j–l) LUAD. Warmer colors indicate regions with higher relative importance for survival prediction.

The illustrated cases from the four cohorts exhibit different spatial weighting patterns. In LUAD, MRePath emphasizes a largely continuous central area, whereas PHISA assigns higher importance to multiple areas separated by regions of lower importance. READ similarly exhibits alternating areas of higher and lower importance within the tissue. These cases illustrate how regional weighting can vary within tissue emphasized by both models. In CESC, the two methods differ mainly in the location and extent of highly weighted regions, while their overall distributions are more similar in COAD. These visualizations provide a qualitative view of the local weighting patterns produced by the complete model.

E.4 Complete Prognostic Stratification Across Four Cohorts

Figure 5 reports the complete four-cohort Kaplan–Meier analysis corresponding to the compact visualization in the main paper.

The predicted high-risk groups show lower estimated survival probabilities than the low-risk groups across all four cohorts, with log-rank $p < 0.05$ in every comparison. COAD and CESC show sustained separation between the two risk groups over follow-up. In READ, survival in the high-risk group declines early and subsequently plateaus, whereas the low-risk curve remains relatively stable. LUAD shows clearer separation during early and intermediate follow-up, with partial convergence



(a) COAD (b) CESC (c) READ (d) LUAD

Figure 5: **Complete Kaplan–Meier survival analysis across four TCGA cohorts.** Patients are stratified into high- and low-risk groups using the median predicted risk score within each cohort. Shaded regions denote confidence intervals, and p-values are obtained from log-rank tests.

later in the observed interval. These analyses complement the C-index evaluation by relating predicted risk groups to observed survival distributions.

F Empirical Analysis of Anti-Dilution Mechanisms

This section empirically examines the feature retention and structural response mechanisms motivated in Section 3.1. The analyses complement the macroscopic WSI visualizations in Figure 4 by focusing on localized responses in high-variation tissue regions.

F.1 Selection of High-Variation ROIs

For the local mechanism analysis, we rank hyperedges according to their within-region feature variation. Using the feature representation captured at the final polarized layer, we compute for each hyperedge e

$$\mathrm{Var}(e) = \frac{1}{|e|} \sum_{v \in e} \|\mathbf{x}_v - \bar{\mathbf{x}}_e\|^2, \tag{44}$$

where $\bar{\mathbf{x}}_e$ denotes the mean feature representation within hyperedge e. We rank the hyperedges by $\mathrm{Var}(e)$ and retain the K hyperedges with the largest values as ROIs for subsequent analysis. ROI ranking is based on feature variation, while the learned polarization phases are used in the subsequent response analysis. These high-variation regions provide local contexts for examining the phase-dependent propagation responses motivated by Assumption 1 and Theorem 2.

F.2 Feature Retention and Structural Dispersion across Propagation

To examine the phase-dependent feature response motivated by Theorem 2 together with the local structural responses defined in Eq. (10), we analyze feature retention and closed-walk dispersion across propagation orders within the high-variation ROIs selected in Appendix F.1.

Phase-Dependent Feature Retention. For each ROI e, we center its node features around $\bar{\mathbf{x}}_e$ and use the centered features as a localized initial signal. We compare propagation with the learned polarization phases against a matched zero-phase counterpart that retains the same incidence magnitudes, hypergraph topology, and degree normalization. Let $m \in \{\mathrm{pol}, \mathrm{real}\}$ denote the polarized and matched zero-phase operators, respectively. We define

$$E_{m,e}^{(k)} = \frac{1}{|\mathcal{V}_e|} \sum_{v \in \mathcal{V}_e} \left\| \mathbf{x}_{m,e,v}^{(k)} \right\|_2^2, \quad R_{m,e}^{(k)} = \frac{E_{m,e}^{(k)}}{E_e^{(0)}}, \quad G_e^{(k)} = \log_{10} \frac{R_{\mathrm{pol},e}^{(k)} + \varepsilon}{R_{\mathrm{real},e}^{(k)} + \varepsilon}, \tag{45}$$

where $E_e^{(0)}$ is the energy of the common initial signal and ε is a small numerical constant. A positive $G_e^{(k)}$ indicates greater local feature retention under polarized propagation relative to the matched zero-phase counterpart.

Closed-Walk Structural Dispersion. We further examine the micro-spatial responses $\mathbf{Z}_v^{(k)} = (\mathbf{P}^k)_{vv}$. Let $\mathcal{V}_{\mathrm{sel}}$ denote the union of nodes contained in the selected ROIs. Their node-wise dispersion

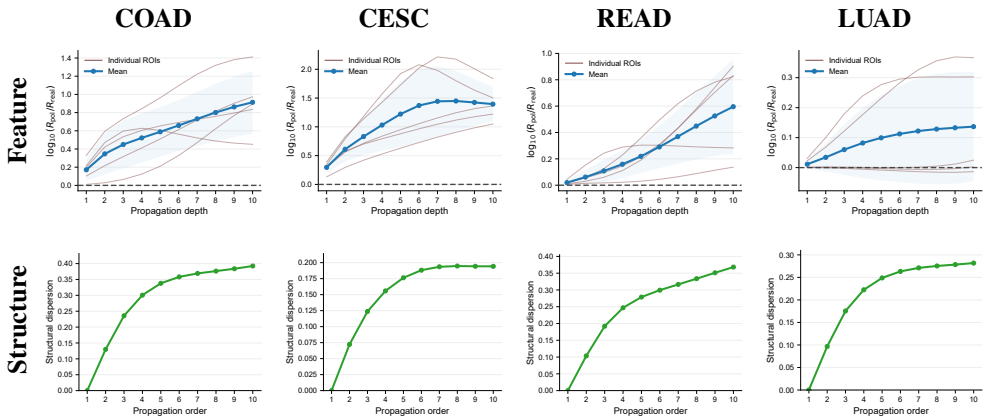


Figure 6: Quantitative propagation analysis for cases from four TCGA cohorts. The first row reports phase-dependent feature retention in the high-variation ROIs, measured by $\log_{10}(R_{\mathrm{pol}}/R_{\mathrm{real}})$ between the learned polarized operator and its matched zero-phase counterpart. Thin curves denote individual ROIs; the thick curve and shaded region show the mean and one standard deviation across ROIs. The second row reports the structural dispersion $D_Z^{(k)}$ of the closed-walk responses across the selected ROI nodes.

at propagation order k is measured by

$$D_Z^{(k)} = \frac{\mathrm{Std}_{v\in\mathcal{V}_{\mathrm{sel}}}\left(|\mathbf{Z}_v^{(k)}|\right)}{\mathrm{Mean}_{v\in\mathcal{V}_{\mathrm{sel}}}\left(|\mathbf{Z}_v^{(k)}|\right)+\varepsilon}. \tag{46}$$

This normalized dispersion characterizes how differentiated the local closed-walk responses become across the selected tissue regions as the propagation order increases.

Figure 6 summarizes the two responses across cases from the four TCGA cohorts. The mean log retention ratio remains positive across the displayed propagation depths, indicating higher feature retention under the learned polarization than under the matched zero-phase counterpart. Structural dispersion increases over the early propagation orders and then exhibits cohort-specific saturation or continued growth, showing how higher-order closed walks progressively differentiate local structural responses.

F.3 Spatial Visualization of Closed-Walk Responses

To provide a spatial view of the micro-spatial component in DS-IASE, we visualize the first- and second-order closed-walk responses around the highest-variation ROI in each cohort. For propagation order $k \in \{1, 2\}$, the node-wise response is $\mathbf{Z}_v^{(k)} = (\mathbf{P}^k)_{vv}$. Within each local crop, we display its relative deviation from the mean response as

$$D_v^{(k)} = \frac{|\mathbf{Z}_v^{(k)}|}{\mathrm{Mean}_{u\in\mathcal{V}_{\mathrm{crop}}}|\mathbf{Z}_u^{(k)}|+\varepsilon} - 1, \tag{47}$$

where $\mathcal{V}_{\mathrm{crop}}$ contains the nodes in the displayed local tissue region. This normalization emphasizes the spatial variation of the closed-walk response within each local context.

Figure 7 compares the first- and second-order response patterns across the four cohorts. The first-order responses are comparatively uniform within the displayed neighborhoods, whereas the second-order responses exhibit clearer local variation. This change illustrates how an additional propagation step introduces greater spatial differentiation into the micro-spatial structural response.

The local response maps complement the whole-slide importance visualizations in Figure 3 by showing the spatial organization of the structural quantities used by the micro-spatial path.

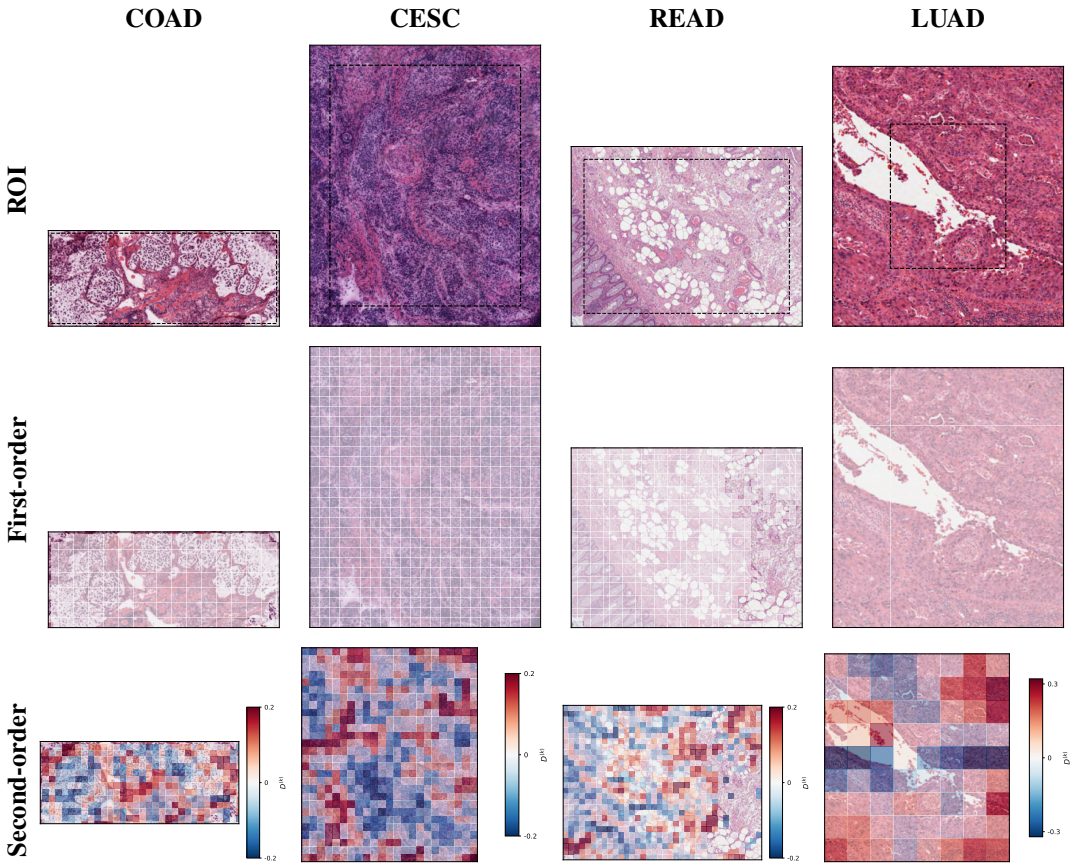


Figure 7: Spatial visualization of micro-spatial closed-walk responses for cases from four TCGA cohorts. The first row shows the local H&E regions centered on the highest-variation ROIs. The second and third rows show the relative first- and second-order closed-walk responses, respectively, normalized by the mean response within each displayed region. The same symmetric display range is used for the first- and second-order responses within each cohort, while display ranges are determined separately across cohorts.

G Empirical Analysis of Global Phase Synchronization

This section examines how the phase alignment of learned polarized spectral modes varies with spectral frequency. Proposition 3 bounds the normalized phase-aware discrepancy of an eigenmode by a quantity that scales with its Laplacian eigenvalue λ_j for a fixed interaction. We therefore analyze angular phase alignment on the same high-variation hyperedges selected in Appendix F.1, providing an empirical view of the low-frequency bias underlying the macro-spectral coordinates in Eq. (7).

G.1 Synchronization Metric and Experimental Setup

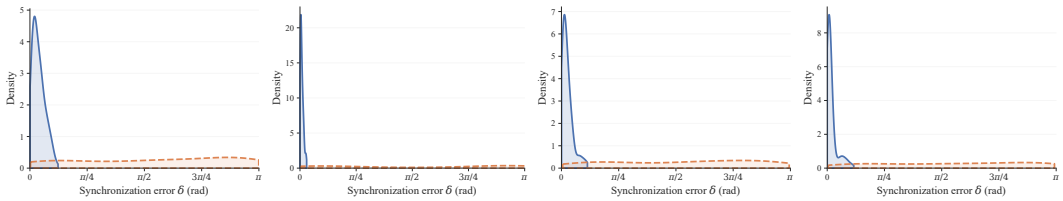
For an interacting node pair (u, v) within hyperedge e, let

$$\Delta\phi_{uv}^{(e)} = \phi_{ue} - \phi_{ve} \tag{48}$$

denote the learned relative incidence phase. For eigenmode $\mathbf{u}_j$, we define the angular synchronization error as

$$\delta_j(u, v; e) = \left| \mathrm{Arg}\left(e^{\,i[(\angle u_{j,u} - \angle u_{j,v}) - \Delta\phi_{uv}^{(e)}]} \right) \right|. \tag{49}$$

A smaller $\delta_j(u, v; e)$ corresponds to closer angular agreement between the relative phase of the eigenmode and the learned local polarization.



(a) COAD (b) CESC (c) READ (d) LUAD

Figure 8: Kernel density estimates of synchronization errors at the spectral extremes for cases from four TCGA cohorts. The lowest non-zero Laplacian mode shows a greater concentration of errors near zero, whereas the largest-eigenvalue mode exhibits a broader angular distribution.

Phase estimation uses interactions whose two eigenvector entries satisfy the magnitude stability criterion. For each eigenmode, we summarize the retained interactions by the Mean Absolute Synchronization Error (MASE),

$$\mathrm{MASE}_j = \frac{1}{|\mathcal{P}_j|} \sum_{(u,v,e)\in\mathcal{P}_j} \delta_j(u, v; e), \tag{50}$$

where $\mathcal{P}_j$ contains the evaluated interactions from the selected high-variation hyperedges.

The angular statistics provide an interpretable diagnostic of the spectral behavior associated with Proposition 3. The proposition characterizes the normalized complex phase-aware discrepancy, while δ_j and MASE_j summarize the corresponding angular alignment across spectral modes.

G.2 Phase Alignment at Spectral Extremes

We first compare the synchronization-error distributions at the two spectral extremes. Specifically, we examine the lowest non-zero Laplacian mode and the mode associated with the largest Laplacian eigenvalue. This endpoint comparison examines whether the low-frequency phase alignment follows the frequency-dependent behavior motivated by Proposition 3.

Across the four cohort cases, the synchronization errors of the lowest-frequency mode are more concentrated near zero, whereas the highest-frequency mode exhibits a broader angular distribution. Although the detailed distribution shapes vary across cohorts, the same endpoint contrast is observed in COAD, CESC, READ, and LUAD. The lower errors at the low-frequency endpoint correspond to closer angular agreement with the learned local polarization.

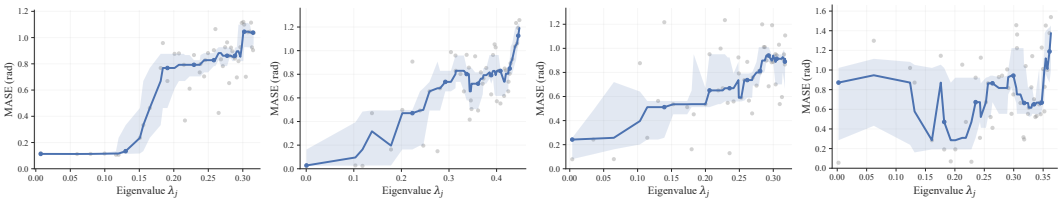
G.3 Synchronization Error across Spectral Frequencies

We further examine phase alignment across the first 50 non-zero Laplacian modes by relating MASE_j to the corresponding eigenvalue λ_j.

Figure 9 shows the mode-wise MASE values together with a rolling local median and interquartile range. The synchronization errors exhibit frequency-dependent variation across the analyzed spectrum, with distinct local trajectories across the four cohorts. COAD, CESC, and READ show an overall shift toward larger errors at higher eigenvalues, while LUAD displays a more variable spectral pattern. These observations complement the endpoint comparison in Figure 8 by showing how angular alignment evolves across intermediate spectral modes. The observed frequency-dependent structure is qualitatively consistent with Proposition 3, which associates smaller eigenvalues with a tighter phase-aware discrepancy bound for a fixed interaction.

G.4 Discussion

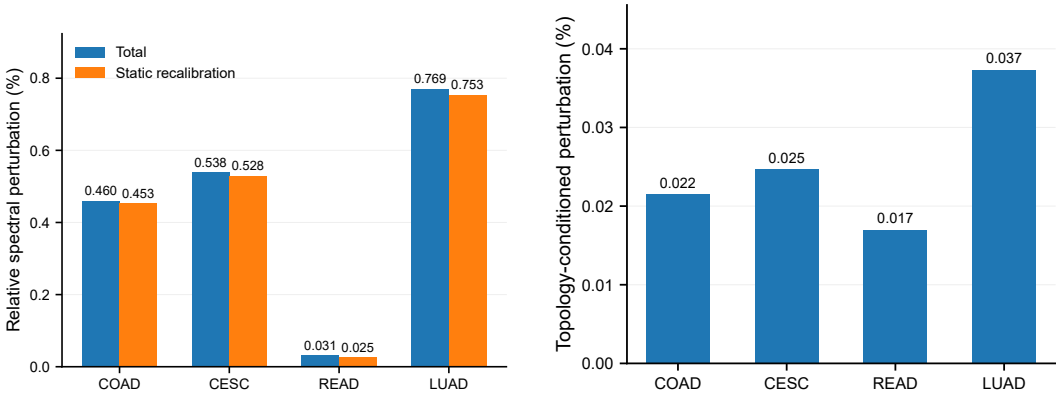
The endpoint and spectrum-wide analyses provide complementary views of the learned spectral phase structure. The endpoint comparison shows tighter angular alignment for the lowest non-zero Laplacian mode than for the largest-eigenvalue mode across cases from the four TCGA cohorts. The spectrum-wide analysis extends this comparison to the first 50 non-zero modes and reveals frequency-dependent variation with cohort-specific local trajectories.



(a) COAD (b) CESC (c) READ (d) LUAD

Figure 9: Synchronization error across the first 50 non-zero Laplacian modes for cases from four TCGA cohorts. Gray points denote the mode-wise MASE values. The solid curve shows the rolling local median, and the shaded region shows the corresponding interquartile range. The four cohorts exhibit distinct frequency-dependent synchronization patterns across the analyzed spectral range.

Together, these observations are qualitatively consistent with Proposition 3, which associates smaller eigenvalues with a tighter phase-aware discrepancy bound for a fixed interaction. The empirical phase-alignment patterns therefore provide additional context for the low-frequency bias of the macro-spectral construction in Eq. (7).



(a) Total spectral perturbation and the patient-independent spectral recalibration.

(b) Magnitude of the topology-conditioned spectral modulation.

Figure 10: Spectral characterization of TCSA adaptation for cases from four TCGA cohorts. (a) Relative magnitudes of the total spectral perturbation and the patient-independent spectral recalibration induced by $\mathbf{s}_{\text{tune}}$. (b) Relative magnitude of the topology-conditioned component $\Delta\boldsymbol{\sigma}_{\text{topo}}$, shown on a separate scale for visibility. All quantities are normalized by the ℓ_2 norm of the retained pretrained singular values.

H Empirical Characterization of TCSA Spectral Adaptation

Proposition 4 shows that TCSA modifies each adapted pretrained projection within the fixed principal singular subspace,

$$\Delta W = U_0 \,\mathrm{diag}(\Delta\boldsymbol{\sigma}) V_0^\top, \tag{51}$$

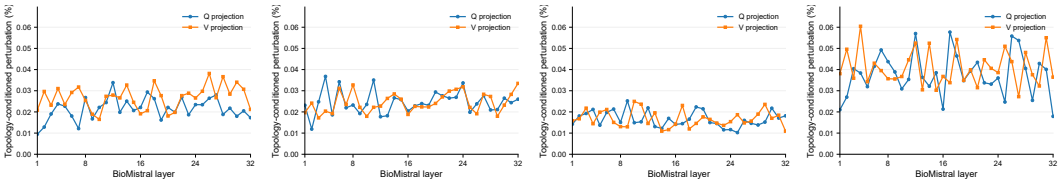
where the singular directions U_0 and V_0 remain fixed. We further characterize the spectral coefficient changes realized in the trained model.

For patient i, the modulated spectral coefficients are

$$\boldsymbol{\sigma}_{\text{mod}}^{(i)} = \boldsymbol{\sigma}_0 \odot \mathbf{s}_{\text{tune}} \odot \cos(W_\phi \Psi_{\text{DS}}^{(i)}), \tag{52}$$

which gives

$$\Delta\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_{\text{mod}}^{(i)} - \boldsymbol{\sigma}_0 = \underbrace{\boldsymbol{\sigma}_0 \odot (\mathbf{s}_{\text{tune}} - \mathbf{1})}_{\Delta\boldsymbol{\sigma}_{\text{static}}} + \underbrace{\boldsymbol{\sigma}_0 \odot \mathbf{s}_{\text{tune}} \odot \left[\cos(W_\phi \Psi_{\text{DS}}^{(i)}) - \mathbf{1}\right]}_{\Delta\boldsymbol{\sigma}_{\text{topo}}^{(i)}}. \tag{53}$$



(a) COAD (b) CESC (c) READ (d) LUAD

Figure 11: Layer-wise topology-conditioned spectral modulation across the 32 BioMistral attention layers for cases from four TCGA cohorts. Query and value projections exhibit distinct layer-wise profiles, with the modulation distributed non-uniformly across model depth.

The first term represents the patient-independent spectral recalibration learned for each adapted projection, while the second captures the topology-conditioned modulation induced by the patient-specific descriptor $\Psi_{\mathrm{DS}}^{(i)}$. We characterize their magnitudes relative to the retained pretrained spectrum using

$$\eta_{\mathrm{total}}^{(i)} = \frac{\|\Delta\boldsymbol{\sigma}_i\|_2}{\|\boldsymbol{\sigma}_0\|_2}, \quad \eta_{\mathrm{static}} = \frac{\|\Delta\boldsymbol{\sigma}_{\mathrm{static}}\|_2}{\|\boldsymbol{\sigma}_0\|_2}, \quad \eta_{\mathrm{topo}}^{(i)} = \frac{\left\|\Delta\boldsymbol{\sigma}_{\mathrm{topo}}^{(i)}\right\|_2}{\|\boldsymbol{\sigma}_0\|_2}. \tag{54}$$

For the global characterization in Figure 10, the coefficient vectors are aggregated across all 64 adapted query and value projections before computing these norms.

Figure 10 shows that the learned spectral perturbations remain small relative to the retained pretrained spectrum, staying below 1% in the four cohort cases considered here. The patient-independent recalibration closely tracks the magnitude of the total perturbation, while the topology-conditioned component remains visible at a smaller absolute scale across all four cohorts. Figure 10(b) therefore displays this component separately to resolve its cross-cohort variation. These relative norms characterize the realized parameter-space modulation, while the predictive role of structural conditioning is examined separately in the main experiments.

To examine how the topology-conditioned component is distributed across BioMistral, we compute for layer ℓ and projection $p \in \{Q, V\}$,

$$\eta_{\mathrm{topo},\ell,p}^{(i)} = \frac{\left\|\Delta\boldsymbol{\sigma}_{\mathrm{topo},\ell,p}^{(i)}\right\|_2}{\|\boldsymbol{\sigma}_{0,\ell,p}\|_2}. \tag{55}$$

As shown in Figure 11, the topology-conditioned modulation varies across the 32 BioMistral attention layers. The query and value projections exhibit distinct layer-wise profiles, and the magnitude of the modulation changes locally across model depth. Across cohorts, the profiles also differ in their local patterns and overall scales.

Together with the global characterization in Figure 10, these results show that TCSA realizes a small but structured spectral adaptation within the fixed pretrained singular subspace.