Quiver Laplacians and Feature Selection

Otto Sumray[1,2], Heather A. Harrington[2,3,4,5], Vidit Nanda[2]

ABSTRACT. The challenge of selecting the most relevant features of a given dataset arises ubiquitously in data analysis and dimensionality reduction. However, features found to be of high importance for the entire dataset may not be relevant to subsets of interest, and vice versa. Given a feature selector and a fixed decomposition of the data into subsets, we describe a method for identifying selected features which are compatible with the decomposition into subsets. We achieve this by re-framing the problem of finding compatible features to one of finding sections of a suitable quiver representation. In order to approximate such sections, we then introduce a Laplacian operator for quiver representations valued in Hilbert spaces. We provide explicit bounds on how the spectrum of a quiver Laplacian changes when the representation and the underlying quiver are modified in certain natural ways. Finally, we apply this machinery to the study of peak-calling algorithms which measure chromatin accessibility in single-cell data. We demonstrate that eigenvectors of the associated quiver Laplacian yield locally and globally compatible features.

1. Introduction

When working with large quantities of unstructured data, it is often convenient – and on occasion, indispensable – to employ a family of real-valued *features* defined on the dataset. It has, for instance, become standard practice within natural language processing to embed words into Euclidean space based on their relative proximity to each other in a substantial corpus of text [23]. In this case, the coordinates serve as features. Regardless of how such an embedding has been obtained, one immediately seeks to simplify matters by identifying those features which are most relevant to the task at hand. There are several ways of quantifying the relevance of a given feature — one may use domain-specific context, optimise a statistical quantity such as variance, or exploit the presence of labelled training data, etc. This process of *feature selection* forms a cornerstone of data analysis. Examples of well-known feature selectors include principal components [26], LASSO [32] and recursive feature elimination [15], to name but a few.

Email of corresponding author: otto.sumray@maths.ox.ac.uk

[1]LUDWIG CANCER RESEARCH, UNIVERSITY OF OXFORD
[2]MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD
[3]MAX PLANCK INSTITUTE OF MOLECULAR CELL BIOLOGY AND GENETICS
[4]CENTRE FOR SYSTEMS BIOLOGY DRESDEN
[5]FACULTY OF MATHEMATICS, TECHNISCHE UNIVERSITÄT DRESDEN

We focus here on feature selection in the presence of an auxiliary decomposition of the given dataset. As a toy example of the situation which we have in mind, let us consider the plots in Figure 1 of two features defined on datasets which have been decomposed into two overlapping subsets A and B.
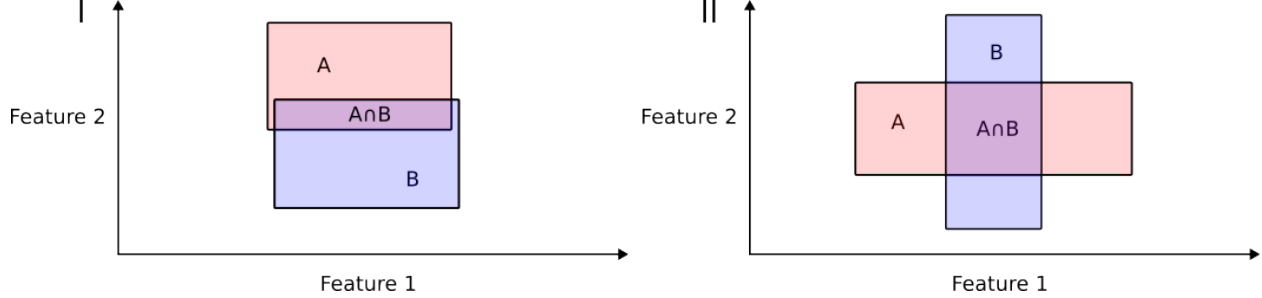

FIGURE 1. Two features defined on a dataset equipped with a decomposition into two overlapping subsets.

Suppose we rank the relevance of features by their variance. In Figure 1.I, feature 1 is most relevant for subsets A, B, and $A \cap B$, but would not be more relevant than feature 2 on their union $A \cup B$. On the other hand in Figure 1.II, feature 1 is most relevant for A, and feature 2 for B, both are equally relevant for $A \cap B$ and $A \cup B$. Such unfortunate inconsistencies arise quite frequently, especially in biological contexts — we are often interested in features defined on datasets that are naturally organised by types, sub-types and so forth, such as disease or cell types. For such data, any feature selection process (which remains agnostic to the decomposition) is liable to select features that are irrelevant to certain subsets while ignoring features that are relevant to certain subsets.

This Paper. Here we present a principled framework for selecting features on a finite set X that are consistent with respect to a decomposition of X into a *cover* $\mathscr{U}$. Crucially, the subsets of X which comprise $\mathscr{U}$ are allowed to admit arbitrary overlaps. Given the immense diversity of available feature selectors and their inner workings, we adopt an abstract approach. Let $\mathbb{K}$ denote either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. For our purposes, a feature selector $\mathbf{S}$ is any deterministic process that

(1) accepts as input some finite collection $\mathscr{F}$ of functions $X \to \mathbb{K}$, and
(2) outputs a subspace $\mathbf{S}_X^{\mathscr{F}}$ of the $\mathbb{K}$-linear span $\mathbf{V}(\mathscr{F})$ of $\mathscr{F}$.

We assume an inner product structure on $\mathbf{V}(\mathscr{F})$, but impose no further constraints on $\mathbf{S}$; as such, the results and constructions of this paper may be applied verbatim to any of the standard feature selectors. Similarly, there are no constraints on the cover $\mathscr{U}$ besides the requirement that $X = \bigcup_{U \in \mathscr{U}} U$.

For the purposes of these introductory remarks, let us call a nonempty subset $\sigma \subset \mathscr{U}$ *admissible* whenever the intersection $|\sigma| := \bigcap_{U \in \sigma} U$ is nonempty. Our quest here is to isolate the largest subspace of $\mathbf{S}_X^{\mathscr{F}}$ that is compatible with respect to the decomposition of X into admissible subsets. Two forms of compatibility are studied here: *local* and *global*.

Given an admissible subset $\sigma \subset \mathscr{U}$, a feature $v : |\sigma| \to \mathbb{R}$ in $\mathbf{S}^{\mathscr{F}}_{|\sigma|}$ is locally compatible with $\mathscr{U}$ if for every admissible $\tau \subset \mathscr{U}$ with $|\tau| \supset |\sigma|$, there is some feature $v^+ : |\tau| \to \mathbb{R}$ in $\mathbf{S}^{\mathscr{F}}_{|\tau|}$ whose restriction to $|\sigma|$ coincides with v. In other words, a locally compatible feature is one which continues to be selected as we pass to larger admissible subsets. Conversely, consider a collection of features $\{v_\sigma\}$, one for each admissible σ. Such a collection is globally compatible with $\mathscr{U}$ if for each pair $|\sigma| \subset |\tau|$ of admissible subsets, the orthogonal projection $\pi^{\mathscr{F}}_\tau : \mathbf{V}(\mathscr{F}) \twoheadrightarrow \mathbf{S}^{\mathscr{F}}_{|\tau|}$ sends v_σ to v_τ. Thus, every globally compatible feature is generated by functions $\{v_U : U \to \mathbb{R} \mid U \in \mathscr{U}\}$ such that the orthogonality relation

$$(v_U - v_{U'}) \perp \mathbf{S}^{\mathscr{F}}_{U \cap U'}$$

holds in $\mathbf{V}(\mathscr{F})$ for every $U, U' \in \mathscr{U}$.

The first question addressed in our work here is:

> *Fix a feature selector $\mathbf{S}$ and a cover $\mathscr{U}$ of X. Given an input $\mathscr{F}$ family of functions $X \to \mathbb{R}$, to what extent can we efficiently discover all of the locally and globally compatible features output by $\mathbf{S}$?*

We approach this problem by recasting it in the language of quiver representations. The underlying quiver $Q^{\mathscr{U}}$ is a directed graph whose vertices are admissible sets, with a single edge $\sigma \to \tau$ whenever $|\sigma| \subset |\tau|$. The output of $\mathbf{S}$ with input $\mathscr{F}$ induces a representation $\mathbf{A}_\bullet = \mathbf{A}^{\mathbf{S},\mathscr{F}}_\bullet$ of Q valued in the category of finite dimensional Hilbert spaces. Explicitly, each vertex σ is assigned the Hilbert space $\mathbf{A}_\sigma := \mathbf{S}^{\mathscr{F}}_{|\sigma|}$, while each edge $\sigma \to \tau$ is assigned the linear map $\mathbf{A}_{\sigma \to \tau} : \mathbf{A}_\sigma \to \mathbf{A}_\tau$ obtained by restricting $\pi^{\mathscr{F}}_\tau$ to $\mathbf{S}^{\mathscr{F}}_{|\sigma|}$. We show that the **sections** of this representation (as introduced in [30] and generalised to the multilinear setting in [24]), recover all the globally compatible features. Similarly, sections of this representation over certain augmented sub-quivers yield all the locally compatible features of $\mathbf{S}$.

The presence of noise severely impairs the compatibility framework described above. The main difficulty is that one never obtains an exact equality between (projections of) features selected over an admissible set $|\tau|$ and features selected over an admissible subset $|\sigma|$. The best-case scenario is that features in $\mathbf{A}_\sigma$ lie *near*[1] the restriction of some feature in $\mathbf{A}_\tau$. In fact, exact equality may not even be desirable — often, features are highly correlated, and collapsing together such features will further reduce the dimension of the space of selected features. Thus, there are many compelling reasons to seek a less rigid compatibility framework built around approximate sections of quiver representations. With this consideration in mind, the second question addressed in this paper is:

> *To what extent can we define and efficiently compute the approximate sections of a quiver representation?*

We address this challenge by introducing a new Laplacian operator.

The **quiver Laplacian** L introduced here is a (Hermitian, positive semidefinite) endomorphism of the *total space* $\mathrm{Tot}(\mathbf{A}_\bullet) := \prod_\sigma \mathbf{A}_\sigma$. In the special case where every $\mathbf{A}_\sigma$ is

---

[1]Here proximity is determined by the inner product metric on $\mathbf{V}(\mathscr{F})$.

one-dimensional and all orthogonal projection maps are identities, L coincides with the ordinary graph Laplacian of $Q^{\mathcal{U}}$. We note that Laplacians have been defined for several discrete algebraic/combinatorial objects: from the original case of graphs [3], to simplicial complexes [12, 17], hypergraphs [2] and principal bundles on graphs [10]. Most closely related to our quiver Laplacian are the analogous operators defined for cellular sheaves in [16] (but see also [11, 25]). The space of sections of a quiver representation coincides precisely with the kernel of the affiliated quiver Laplacian. Armed with the quiver Laplacian, we are able to define, for every $\epsilon > 0$, the *ϵ-approximate sections* of a quiver representation as those normalised vectors $x \in \text{Tot}(\mathbf{A}_\bullet^{\mathbf{S}, \mathscr{F}})$ which satisfy $\langle Lx, x \rangle \leq \epsilon$. Thus, the spans of eigenvectors of L associated to small nonzero eigenvalues of L delineate special families of approximate sections. Motivated by these considerations, we establish a host of fundamental properties for quiver Laplacians and their spectra.

Finally, we use the spectrum of an appropriate quiver Laplacian to study feature selection (as given by a *peak-calling* algorithm) for chromatin accessibility[2] data from [28]. Our dataset comes from single-cell sequencing. This type of data has several properties which make the Laplacian-driven approach particularly appropriate:

(1) The dimensionality is massive — a naïve estimate assumes that every genomic position might be accessible, giving around 3 billion features. Therefore, feature selection (via peak-calling) becomes a crucial intermediate step towards making the analysis tractable.

(2) Different cell types have different relevant features, and it is important to make the feature selection process compatible with the types of cells under consideration. This makes it necessary to consider peaks not only for the entire dataset, but also for certain distinguished subsets.

(3) In many relevant cases where the data arise from diseases such as cancer or from the immune system (e.g. T cells), the cells might exhibit plasticity and their type is therefore not well-defined. In such cases, our distinguished subsets might overlap and should be modelled by a cover rather than a partition.

(4) There is some stochasticity in the output of peak-calling algorithms — in particular, the accessible genomic positions are only known approximately. Thus, the accessibility status of nearby positions is not independent; we capture this phenomenon via the inner product structure on feature space.

Summary of Results. Let Q be a finite quiver with vertex set Q_0, edge set Q_1 and maps $s, t : Q_1 \rightarrow Q_0$ which specify the source and target vertices of each edge. We consider representations of Q — each such representation $\mathbf{A}_\bullet$ assigns a finite-dimensional (real or complex) Hilbert space $\mathbf{A}_v$ to every vertex $v \in Q_0$ and a linear map $\mathbf{A}_e : \mathbf{A}_{s(e)} \rightarrow \mathbf{A}_{t(e)}$ to every edge $e \in Q_1$. We write $L_{\mathbf{A}}$ for the Laplacian of $\mathbf{A}_\bullet$ (see Definition 4.2 below

---

[2]Chromatin accessibility measures the extent to which regions within the genome (among a specific family of cells) are reachable by cellular machinery, such as transcription factors.

for details), and arrange its eigenvalues as

$$0 \leq \lambda_1(\mathbf{A}) \leq \lambda_2(\mathbf{A}) \leq \cdots \leq \lambda_n(\mathbf{A}),$$

where $n := \dim \mathrm{Tot}(\mathbf{A}_\bullet)$ is the total dimension of $\mathbf{A}_\bullet$.

Here is a simplified version of our first main result, which provides upper bounds on Laplacian eigenvalues of a representation $\mathbf{A}'_\bullet$ in terms of eigenvalues of $\mathbf{A}_\bullet$ whenever there exists a family of vertex-indexed linear maps $\mathbf{A}_v \to \mathbf{A}'_v$.

THEOREM (I). *Let $\mathbf{A}_\bullet$ and $\mathbf{A}'_\bullet$ be two representations of Q, and consider any vertex-indexed collection $\tau = \{\tau_v : \mathbf{A}_v \to \mathbf{A}'_v \mid v \in Q_0\}$ of linear maps. Write q for the dimension of $\ker \tau$ viewed as a (block diagonal) map $\mathrm{Tot}(\mathbf{A}_\bullet) \to \mathrm{Tot}(\mathbf{A}'_\bullet)$. There exist positive constants α, β, γ – each dependent on τ – such that the inequality*

$$\lambda_k(\mathbf{A}') \leq \alpha \cdot \lambda_{k+q}(\mathbf{A}) + \beta \cdot \sqrt{\lambda_{k+q}(\mathbf{A})} + \gamma$$

holds for every k in $\{1, 2, \ldots, n - q\}$.

The full statement of this result, along with explicit formulas for α, β and γ, is recorded in Theorem 5.2. If τ is a morphism of quiver representations $\mathbf{A}_\bullet \to \mathbf{A}'_\bullet$, then both β and γ vanish and we are left with a much simpler bound $\lambda_k(\mathbf{A}') \leq \alpha \cdot \lambda_{k+q}(\mathbf{A})$.

Our next main result takes the form of a *spectral stability* result for feature selectors. The main challenge here is that the total spaces of two representations $\mathbf{A}_\bullet^{\mathbf{S}, \mathscr{F}}$ and $\mathbf{A}_\bullet^{\mathbf{T}, \mathscr{F}}$ may have differnt dimensions, so even the number of eigenvalues of the corresponding Laplacians could be different. Nevertheless, let $\mu_\mathbf{S}$ and $\mu_\mathbf{T}$ be the probability measures on $\mathbb{R}$ given by taking the average of Dirac measures concentrated at the eigenvalues $\{\lambda_i(\mathbf{A}^{\mathbf{S}, \mathscr{F}})\}$ and $\{\lambda_j(\mathbf{A}^{\mathbf{T}, \mathscr{F}})\}$, respectively. We obtain, in Corollary 5.5, the following bound on the 1-Wasserstein distance $W_1(\mu_\mathbf{S}, \mu_\mathbf{T})$ between these two measures.

THEOREM (II). *Let $\mathbf{S}$ and $\mathbf{T}$ be two feature selectors defined on a finite set X which is equipped with a cover $\mathscr{U}$. Fix a finite collection $\mathscr{F}$ of functions $X \to \mathbb{K}$. For each admissible set σ, define*

$$c_\sigma := \left| \dim \mathbf{S}_{|\sigma|}^{\mathscr{F}} - \dim \mathbf{T}_{|\sigma|}^{\mathscr{F}} \right| \quad \text{and} \quad \epsilon_\sigma := \mathrm{dist}_{\mathbf{Gr}} \left(\mathbf{S}_{|\sigma|}^{\mathscr{F}}, \mathbf{T}_{|\sigma|}^{\mathscr{F}} \right),$$

where the latter expression invokes the Grassmannian distance between subspaces of $\mathbf{V}(\mathscr{F})$. Define $c := \sum_\sigma c_\sigma$ and $\epsilon := \max_\sigma \epsilon_\sigma$. There exist trigonometric polynomials f and g such that

$$W_1(\mu_\mathbf{S}, \mu_\mathbf{T}) \leq \frac{|Q_1^{\mathscr{U}}|}{m} \left[f(\epsilon) \cdot c + g(\epsilon) \cdot (m - c) \right],$$

where $|Q_1^{\mathscr{U}}|$ is the number of edges in $Q^{\mathscr{U}}$ and $m = \max\{\dim \mathrm{Tot}(\mathbf{A}_\bullet^{\mathbf{S}, \mathscr{F}}), \dim \mathrm{Tot}(\mathbf{A}_\bullet^{\mathbf{T}, \mathscr{F}})\}$.

We note that $c = 0$ if and only if the two feature selectors assign equidimensional spaces to each admissible set. Explicit formulas for f and g are provided in Section 5. Crucially, we have $g(0) = 0$; in the equidimensional case, the first term in our bound vanishes and the second term approaches 0 as $\epsilon \to 0$.

Our next main result provides several flavours of eigenvalue interlacing inequalities for quiver Laplacians which hold when quivers (and the overlaid representations) are modified in certain natural ways.

THEOREM (III). *Let* $\mathbf{A}_\bullet$ *be a representation of* Q. *The following inequalities hold for appropriate integers* i. *The constants* k, r, w_1, w_2 *which appear below depend on* $\mathbf{A}_\bullet$ *as well as the modification performed to create a new representation* $\mathbf{A}'_\bullet$ *of a new quiver* Q'.

(1) *If* Q' *is obtained by removing edges from* Q *and* $\mathbf{A}'_\bullet$ *is the restriction of* $\mathbf{A}_\bullet$ *to* Q', *then*

$$\lambda_{k+i}(\mathbf{A}_\bullet) \leq \lambda_{k+i}(\mathbf{A}'_\bullet) \leq \lambda_{k+r+i}(\mathbf{A}_\bullet).$$

(2) *If* Q' *is obtained by removing vertices (and incident edges) from* Q *and* $\mathbf{A}'_\bullet$ *is the restriction of* $\mathbf{A}_\bullet$ *to* Q' *then*

$$\lambda_i(\mathbf{A}_\bullet) - w_1 \leq \lambda_i(\mathbf{A}'_\bullet) \leq \lambda_{k+i}(\mathbf{A}_\bullet) - w_2.$$

(3) *If* Q' *and* $\mathbf{A}'_\bullet$ *are the result of performing an admissible homotopy of* Q, *then*

$$\varphi^{-2}\lambda_i(\mathbf{A}_\bullet) \leq \lambda_i(\mathbf{A}'_\bullet) \leq \varphi^2\lambda_i(\mathbf{A}_\bullet)$$

where φ *is the golden ratio.*

(4) *If* Q' *and* $\mathbf{A}'_\bullet$ *are the result of performing an admissible Kron reduction of* $\mathbf{A}_\bullet$, *then*

$$\lambda_i(\mathbf{A}_\bullet) \leq \lambda_i(\mathbf{A}'_\bullet) \leq \lambda_{i+r}(\mathbf{A}_\bullet).$$

Detailed versions of these results, including explicit formulas for various constants as well as precise definitions of admissible homotopy and Kron reduction, are located in Section 6. In order to state and prove these results in an appropriately general setting, we define *sheaves* on quivers — these are almost identical to cellular sheaves [6, 5, 16] on graphs, except that our graphs are allowed to have self-loops.

As mentioned above, much of our work here has been motivated by the desire to better understand and improve peak-calling algorithms on datasets X which come naturally equipped with a decomposition $\mathscr{U}$ into subsets. For this purpose, it is convenient to combine the quiver representations whose sections give locally and globally compatible features of a given feature selector $\mathbf{S}$ into a single representation of a larger quiver. This *combined representation*, denoted $\mathbf{N}_\bullet^{\mathbf{S}, \mathscr{F}}$, is described in Section 3.1. By design, its sections correspond to features which are both locally and globally compatible with $\mathscr{U}$ along $\mathscr{F}$, and we naturally seek its approximate sections. Unfortunately, extracting spectral data for its Laplacian presents a serious computational challenge — both the underlying quiver and the total dimension are typically much larger than $Q^{\mathscr{U}}$ and $\mathbf{A}_\bullet^{\mathbf{S}, \mathscr{F}}$. The third main result of this paper is a remedy which takes the form of a *reduction theorem*.

THEOREM (IV). *Let* $\mathbf{S}$ *be a feature selector on a set* X *equipped with a cover* $\mathscr{U}$ *such that* $Q^{\mathscr{U}}$ *is weakly connected. Let* $\mathbf{N}_\bullet^{\mathbf{S}, \mathscr{F}}$ *be the combined representation associated to a finite collection* $\mathscr{F}$ *of functions* $X \to \mathbb{K}$. *Fix a vertex* $\sigma \in Q_0^{\mathscr{U}}$. *There exists a new quiver* Q' *with a single vertex and* $|Q_0^{\mathscr{U}}|$ *edges along with a representation* $\mathbf{A}'_\bullet$ *of* Q' *satisfying two properties:*

(1) *For each vertex τ of $Q^{\mathscr{U}}$, let id_τ be the identity map on $\mathbf{S}^{\mathscr{F}}_{|\tau|}$ and let $\iota^{\mathscr{F}}_\tau : \mathbf{S}^{\mathscr{F}}_{|\tau|} \hookrightarrow \mathbf{V}(\mathscr{F})$ denote the inclusion map (which is adjoint to $\pi^{\mathscr{F}}_\tau$). Then, the Laplacian of $\mathbf{A}'_\bullet$ is*

$$L_{\mathbf{A}'} = \sum_{\tau \in Q^{\mathscr{U}}_0} \left[\mathrm{id}_\sigma - \pi^{\mathscr{F}}_\sigma \iota^{\mathscr{F}}_\tau \pi^{\mathscr{F}}_\tau \iota^{\mathscr{F}}_\sigma \right]$$

(2) *Let n be the total dimension of $\mathbf{N}^{\mathbf{S},\mathscr{F}}_\bullet$. There exist positive constants C_1 and C_2, which depend only on $\mathbf{A}^{\mathbf{S},\mathscr{F}}_\bullet$, such that the inequalities*

$$C_1 \cdot \lambda_i(\mathbf{N}^{\mathbf{S},\mathscr{F}}_\bullet) \le \lambda_i(\mathbf{A}'_\bullet) \le C_2 \cdot \lambda_{i+k}(\mathbf{N}^{\mathbf{S},\mathscr{F}}_\bullet)$$

hold for every i in $\{1, 2, \ldots, n\}$, with $k := n - \dim \mathbf{S}^{\mathscr{F}}_{|\sigma|}$.

The full statement of this result is recorded in Theorem 7.1.

Finally, we examined a single-cell chromatin accessibility (ATAC-seq) dataset from [28]; here X is a dataset consisting of around 3×10^4 tumour-infiltrating T cells, while the input $\mathscr{F}$ consists of 20,000 binary-valued features $X \to \mathbb{R}$. To build a cover $\mathscr{U}$ of X, we used a standard notion of genomic proximity between cells, constructed a 15-nearest neighbour graph, and expanded the communities of that graph by one hop. This produces a quiver $Q^{\mathscr{U}}$ with 87 vertices. We then ran the MACS2 peak-calling algorithm [39] to find the 1,000 most accessible regions for each admissible set and built a quiver Laplacian for a variant of the combined representation of $\mathbf{S}$ (as described in Theorem 7.2).

OBSERVATION (V). *The first 3,000 eigenvectors of the quiver Laplacian described above had the following properties:*

(1) *The majority of eigenvectors displayed a tight grouping of genomic positions (around 150 consecutive base-pairs out of approximately 3 billion) across the vertices.*

(2) *Most eigenvectors were supported in most of the vector spaces assigned to vertices of the quiver. In particular, one such eigenvector that was supported on every vertex, lay near the gene FAM72D, which is involved in the cell cycle and hence active for every type of T cell.*

(3) *Conversely, the support of a handful of eigenvectors was localised to a very small subset of vertices. One such eigenvector lay near the gene STAT3, which regulates Th1, Th17 and Treg cells; this eigenvector was supported on just 3 vertices, each of which corresponds to admissible subpopulations particularly rich in these three types of T cells.*

Thus, eigenvectors of the quiver Laplacian were able to isolate relevant and consistent meta-features across different scales within the given dataset.

Outline. The paper is structured as follows: §2 describes how a feature selector together with a cover of the data yields a quiver representation. In §3 we describe both local and global compatibility of features with respect to the cover. We also describe how these features can be computed simultaneously as sections of a larger quiver representation. §4 introduces the quiver Laplacian and how its eigenvectors may be used to approximate the space of sections. §5 studies how a transformation of quiver representations relates

the spectra of the respective representations and bounds their spectral distance. This then gives a stability of the spectrum of the representation associated to a feature selector. §6 bounds the changes in the spectrum of a quiver Laplacian when the underlying quiver undergoes various operations. §7 then uses these operations to simplify the computation of approximate sections of a feature selector. Finally, in §8 we use quiver Laplacians to extract (approximately) compatible peaks in single-cell chromatin accessibility data.

Acknowledgements

2. From Feature Selection to Quiver Representations

Given a finite set X and a field $\mathbb{K}$, let X^* be the $\mathbb{K}$-vector space which consists of all functions $X \to \mathbb{K}$. Denote by $\mathrm{Sub}(X^*)$ the collection of all subsets of X^* (not necessarily subspaces), and for each such subset $\mathscr{F}$ write $\mathbf{V}(\mathscr{F})$ to indicate the subspace of X^* given by the $\mathbb{K}$-linear span of $\mathscr{F}$. Let $\mathbf{Gr}_k(X^*)$ be the Grassmannian of k-dimensional subspaces of X^* for each $k \geq 0$, and consider the disjoint union

$$\mathbf{Gr}_\infty(X^*) := \coprod_{k \geq 0} \mathbf{Gr}_k(X^*).$$

Here is the main object of study in this work.

DEFINITION 2.1. A **feature selector** on X is any map $\mathbf{S} : \mathrm{Sub}(X^*) \to \mathbf{Gr}_\infty(X^*)$, such that for each input $\mathscr{F} \subset X^*$, the corresponding output $\mathbf{S}_X^{\mathscr{F}}$ is a subspace of $\mathbf{V}(\mathscr{F})$.

We now fix a feature selector $\mathbf{S}$ on X; readers may wish to keep the following concrete example in mind throughout the remainder of §2.

EXAMPLE 2.2. Among the most well-known and ubiquitous feature selectors is *principal components analysis* [**19**] — here,

- X is a finite origin-centered subset of some real Euclidean space $\mathbb{R}^n$,
- $\mathbb{K}$ is the field $\mathbb{R}$ of real numbers, and
- $\mathscr{F}$ is the set $\{x \mapsto \langle x, p \rangle\}$ of maps given by taking inner products with the unit vectors p.

For a fixed $k \ll n$, the output $\mathbf{S}_X^{\mathscr{F}}$ is the vector space spanned by the k inner product functions (or equivalently, the k unit vectors) along which the variance in X is maximised.

Our goal here is to find a principled framework for relating the vector subspaces $\mathbf{S}_Y^{\mathscr{F}}$ to each other for a fixed $\mathscr{F}$ as Y ranges over subsets of X. Already for principal components analysis, it is clear that there need not be any coherent relationship between $\mathbf{S}_X^{\mathscr{F}}$ and $\mathbf{S}_Y^{\mathscr{F}}$ for arbitrary $Y \subset X$ — for instance, neither one is guaranteed to be a subspace of the other within $\mathbf{V}(\mathscr{F})$. It therefore becomes necessary to constrain the family of subsets under consideration. Here we will describe how to relate the subspaces $\mathbf{S}_Y^{\mathscr{F}}$ for Y ranging over the lattice of subsets generated from a chosen *cover* $\mathscr{U}$ of X. We recall that any such $\mathscr{U}$ is a finite collection of nonempty subsets satisfying $X = \bigcup_{U \in \mathscr{U}} U$.

DEFINITION 2.3. The **quiver associated to a cover** $\mathscr{U}$ of X is the directed graph $Q^{\mathscr{U}}$ whose

(1) vertices are subsets $\sigma \subset \mathscr{U}$ whose *support* $|\sigma| := \bigcap_{U \in \sigma} U$ is nonempty; and,
(2) there is a unique directed edge $\sigma \to \tau$ whenever σ properly contains τ (and hence $|\sigma| \subset |\tau|$).

We now re-assemble the algebraic data of $\mathbf{S}$ to produce a *representation* of $Q^{\mathscr{U}}$, i.e., an assignment of vector spaces to vertices and linear maps to edges [29]. Implicit in the construction below is the choice of an inner product structure on X^*, which allows us to define adjoints of linear maps, and hence, orthogonal projections onto subspaces. For each vertex σ of $Q^{\mathscr{U}}$, let $\iota_\sigma^{\mathscr{F}} : \mathbf{S}_{|\sigma|}^{\mathscr{F}} \hookrightarrow \mathbf{V}(\mathscr{F})$ denote the inclusion map and let $\pi_\sigma^{\mathscr{F}} : \mathbf{V}(\mathscr{F}) \twoheadrightarrow \mathbf{S}_{|\sigma|}^{\mathscr{F}}$ denote its adjoint, i.e., the orthogonal projection in X^* with respect to our chosen inner product.

DEFINITION 2.4. For each subset $\mathscr{F} \subset X^*$, the **S-representation** of $Q^{\mathscr{U}}$ along $\mathscr{F}$ consists of the following assignments $\mathbf{A}_\bullet = \mathbf{A}_\bullet^{\mathbf{S},\mathscr{F}}$:

(1) every vertex σ of $Q^{\mathscr{U}}$ is assigned the vector space $\mathbf{A}_\sigma := \mathbf{S}_{|\sigma|}^{\mathscr{F}}$, which is a subspace of $\mathbf{V}(\mathscr{F}|_{|\sigma|})$ and hence[3] of $\mathbf{V}(\mathscr{F})$; moreover,
(2) every edge $\sigma \to \tau$ of $Q^{\mathscr{U}}$ is assigned the $\mathbb{K}$-linear map $\mathbf{A}_{\sigma \to \tau} : \mathbf{A}_\sigma \to \mathbf{A}_\tau$ given by $\pi_\tau^{\mathscr{F}} \circ \iota_\sigma^{\mathscr{F}}$.

It may be worth noting that the vertices of $Q^{\mathscr{U}}$ are simplices in the *nerve* of the cover $\mathscr{U}$ [8], and every edge $\sigma \to \tau$ in $Q^{\mathscr{U}}$ corresponds to a face relation (where the simplex τ lies in the boundary of the simplex σ). In particular, it follows from the definition of $Q^{\mathscr{U}}$ above that the existence of adjacent edges $\sigma \to \tau \to \gamma$ forces the existence of the edge $\sigma \to \gamma$. We call $\mathbf{A}_\bullet$ a **sheaf** on the nerve of $\mathscr{U}$ if it satisfies the following *associativity criterion*: for every such pair of adjacent edges, the map $\mathbf{A}_{\sigma \to \gamma}$ must equal the composite $\mathbf{A}_{\tau \to \gamma} \circ \mathbf{A}_{\sigma \to \tau}$. In general, we do not expect this associative property to hold when $\mathbf{A}_\bullet$ arises from a feature selector. The next section contains our attempt to rectify this shortcoming.

---

[3]Here we adopt the convention that any function $f : |\sigma| \to \mathbb{K}$ extends to all of X by setting $f = 0$ on $X \setminus |\sigma|$. Conversely, any function $X \to \mathbb{K}$ automatically restricts to a function on $|\sigma|$.

3. Compatibility and Sections

Here we use the **S**-representation to define and study two notions of compatibility for a feature selector **S** on X against a cover $\mathscr{U}$ when invoked with input $\mathscr{F} \subset X^*$:

(1) Fix a vertex $\sigma \in Q_0^{\mathscr{U}}$. An element $v_\sigma \in \mathbf{A}_\sigma^{\mathbf{S},\mathscr{F}}$ is *locally compatible* with $\mathscr{U}$ if

$$\iota_\sigma^{\mathscr{F}}(v_\sigma) = \iota_\tau^{\mathscr{F}} \circ \mathbf{A}_{\sigma \to \tau}(v_\sigma)$$

holds for every edge $\sigma \to \tau$ in Q_1.

(2) A collection $\{v_\sigma \in \mathbf{A}_\sigma \mid \sigma \in Q_0^{\mathscr{U}}\}$ is *globally compatible* with $\mathscr{U}$ if the equality

$$\mathbf{A}_{\sigma \to \tau}(v_\sigma) = v_\tau$$

holds for every edge $\sigma \to \tau$ in $Q^{\mathscr{U}}$.

This notion of local compatibility checks whether or not (the orthogonal projections of) a given feature v_σ which has been selected by **S** over $|\sigma|$ continue to be selected over all greater subsets $|\tau| \supset |\sigma|$ generated by the cover $\mathscr{U}$. In this sense, local compatibility tests the robustness of selected features as we decrease the depth of the cover. Already for principal component analysis, it is clear that not every feature will be locally compatible, i.e., in general $\mathbf{A}_{\sigma \to \tau}$ has a non-trivial kernel. Global compatibility, on the other hand, tests features horizontally across the quiver — to be globally compatible, the vectors of a family $\{v_\sigma\}$ must map coherently onto each other under the linear maps of $\mathbf{A}_\bullet$ as we traverse zigzag paths in $Q^{\mathscr{U}}$ of the form

$$\sigma_1 \to \tau_1 \leftarrow \sigma_2 \to \tau_2 \leftarrow \cdots \leftarrow \sigma_k \to \tau_k.$$

Locally compatible feature   Globally compatible feature   Locally and globally compatible feature
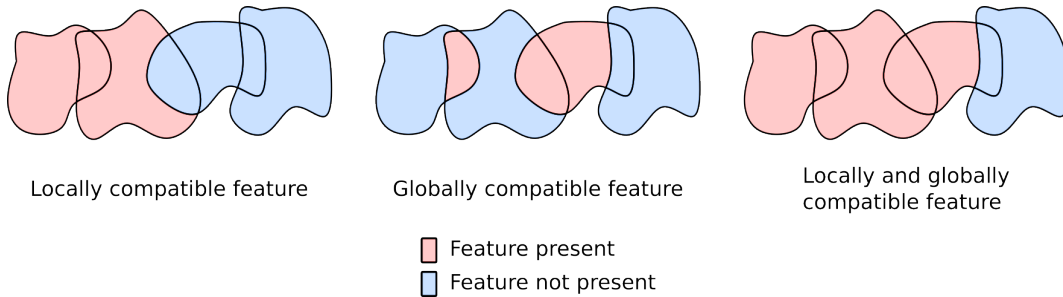
☐ Feature present
☐ Feature not present

FIGURE 2. Example of different compatibility conditions.

We defer the study of local compatibility for now, and will focus on global compatibility.

3.1. Global Compatibility and Sections. Let $Q = (s, t : Q_1 \to Q_0)$ be a finite quiver and let $\mathbf{A}_\bullet$ be a representation of Q valued in the category $\mathbf{FdHilb}(\mathbb{K})$ of finite-dimensional Hilbert spaces over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The *total space* of $\mathbf{A}_\bullet$ is defined as the product

$$\mathrm{Tot}(\mathbf{A}_\bullet) := \prod_{i \in Q_0} \mathbf{A}_i. \tag{1}$$

Since Q is finite, $\text{Tot}(\mathbf{A}_\bullet)$ inherits a finite-dimensional Hilbert space structure from its $\mathbf{A}_i$ factors. Motivated by the global compatibility criterion from Definition 3.2, we highlight a relevant subspace of the total space below.

DEFINITION 3.1. A tuple $\gamma := (\gamma_i \in \mathbf{A}_i \mid i \in Q_0)$ in $\text{Tot}(\mathbf{A}_\bullet)$ is called a **section** of $\mathbf{A}_\bullet$ if for every edge e in Q_1 we have $\mathbf{A}_e(\gamma_{s(e)}) = \gamma_{t(e)}$.

It is straightforward to check that sums and scalar multiples of sections are also sections; we write $\Gamma(\mathbf{A}_\bullet)$ to denote the vector space of sections of $\mathbf{A}_\bullet$. This space was introduced in [**30**] along with an effective algorithm for its computation — see [**30**, Sec 5.2]. Sections are closely related to the compatibility criteria from Definition 3.2. In particular, if $\mathbf{S}$ is a feature selector on a set X, then the sections of $\mathbf{A}_\bullet^{\mathbf{S},\mathscr{F}}$ are exactly globally compatible features of $\mathbf{S}$ against the cover $\mathscr{U}$. Sections can also be used to compute locally compatible features, as we will now describe.

3.2. Local Compatibility.

DEFINITION 3.2. A feature selector $\mathbf{S}$ on X is **locally compatible** with $\mathscr{U}$ on $\mathscr{F} \subset X^*$ if the following triangle of vector spaces commutes for each edge $\sigma \to \tau$ of $Q^{\mathscr{U}}$:

$$
\begin{array}{ccc}
\mathbf{A}_\sigma^{\mathbf{S},\mathscr{F}} & \xrightarrow{\ \mathbf{A}_{\sigma\to\tau}^{\mathbf{S},\mathscr{F}}\ } & \mathbf{A}_\tau^{\mathbf{S},\mathscr{F}} \\
& \searrow{\scriptstyle \iota_\sigma^{\mathscr{F}}} \quad \swarrow{\scriptstyle \iota_\tau^{\mathscr{F}}} & \\
& \mathbf{V}(\mathscr{F}) &
\end{array}
$$

Equivalently, the identity $\iota_\sigma^{\mathscr{F}} = \iota_\tau^{\mathscr{F}} \circ \pi_\tau^{\mathscr{F}} \circ \iota_\sigma^{\mathscr{F}}$ holds for every edge $\sigma \to \tau$ of $Q^{\mathscr{U}}$.

It will be useful to rephrase this compatibility condition entirely within the realm of quiver representations by blowing up the above triangle to a square. To this end, we introduce another representation $\mathbf{C}_\bullet = \mathbf{C}_\bullet^{\mathscr{F}}$ of $Q^{\mathscr{U}}$ which may be associated to every $\mathscr{F} \subset X^*$ (and which does not depend on $\mathbf{S}$). This *constant representation* assigns $\mathbf{C}_\sigma := \mathbf{V}(\mathscr{F})$ to every vertex σ and the identity map to each edge $\sigma \to \tau$. For each vertex σ, there is an evident inclusion map

$$\iota_\sigma^{\mathscr{F}} : \mathbf{A}_\sigma^{\mathbf{S},\mathscr{F}} \hookrightarrow \mathbf{C}_\sigma^{\mathscr{F}},$$

since the codomain is $\mathbf{V}(\mathscr{F})$ and the domain is its subspace $\mathbf{S}_{|\sigma|}^{\mathscr{F}}$. Now, for every edge $\sigma \to \tau$ we have a diagram of four linear maps:

$$
\begin{array}{ccc}
\mathbf{A}_\sigma^{\mathbf{S},\mathscr{F}} & \xrightarrow{\ \mathbf{A}_{\sigma\to\tau}^{\mathbf{S},\mathscr{F}}\ } & \mathbf{A}_\tau^{\mathbf{S},\mathscr{F}} \\
{\scriptstyle \iota_\sigma^{\mathscr{F}}}\downarrow & & \downarrow{\scriptstyle \iota_\tau^{\mathscr{F}}} \\
\mathbf{C}_\tau^{\mathscr{F}} & \xrightarrow[\ \mathbf{C}_{\sigma\to\tau}^{\mathscr{F}}\ ]{} & \mathbf{C}_\sigma^{\mathscr{F}}
\end{array}
\tag{2}
$$

The vertical maps $\iota^{\mathscr{F}}$ prescribe a *morphism of quiver representations* from $\mathbf{A}_\bullet^{\mathbf{S},\mathscr{F}}$ to $\mathbf{C}_\bullet^{\mathscr{F}}$ if and only if the above diagram commutes for every edge $\sigma \to \tau$ in $Q^{\mathscr{U}}$. By definition of

$\mathbf{C}_\bullet^{\mathcal{F}}$, the bottom edge of this diagram is precisely $\mathbf{V}(\mathcal{F}) \xrightarrow{=} \mathbf{V}(\mathcal{F})$, so this square is an elementary modification of the triangle from Definition 3.2. Thus, we have arrived at the following straightforward result.

PROPOSITION 3.3. *The feature selector* $\mathbf{S}$ *is locally compatible with* $\mathcal{U}$ *on* $\mathcal{F} \subset X^*$ *if and only if the collection of maps*

$$\left\{ \iota_\sigma^{\mathcal{F}} : \mathbf{A}_\sigma^{\mathbf{S},\mathcal{F}} \hookrightarrow \mathbf{C}_\sigma^{\mathcal{F}} \right\}$$

indexed by vertices σ *of* $Q^{\mathcal{U}}$ *prescribe a morphism of quiver representations* $\mathbf{A}_\bullet^{\mathbf{S},\mathcal{F}} \to \mathbf{C}_\bullet^{\mathcal{F}}$.

As described earlier, a morphism τ between quiver representations $\mathbf{A}_\bullet$ and $\mathbf{A}_\bullet'$ is a collection

$$\tau = \left\{ \tau_v : \mathbf{A}_v \to \mathbf{A}_v' \mid v \in Q_0 \right\}$$

of linear maps τ_v such that

$$\mathbf{A}_e' \circ \tau_{s(e)} = \tau_{t(e)} \circ \mathbf{A}_{t(e)} \tag{3}$$

for each edge $e \in Q_1$. The commutivity condition in (3) can be too strong for quiver representations obtained from data, so we will define a *transformation* τ between quiver representations as a collection of linear maps

$$\tau = \left\{ \tau_v : \mathbf{A}_v \to \mathbf{A}_v' \mid v \in Q_0 \right\}$$

without the commutivity condition.

DEFINITION 3.4. Let Q be a quiver. For each vertex $u \in Q_0$ define the **floret** of based at u to be the following quiver $F^{Q,u}$:

(1) the set of vertices is the disjoint union

$$F_0^{Q,u} = \{u\} \sqcup \{v_e \mid (e : u \to v) \in Q_1\}.$$

In other words, there is a distinguished copy of u, which we label u_0 henceforth, and an additional object denoted v_e for every edge in Q_1 from u to v.

(2) Define the set of edges of $F^{Q,u}$ as

$$F_1^{Q,u} = \{e_{LD}, e_{DL} : u_0 \to v_e \mid (e : u \to v) \in Q_1\}.$$

In other words, for each edge $e : u \to v$ in Q there are two distinct edges e_{LD} and e_{DL} from u_0 to v_e in $F^{Q,u}$.

Now suppose $\mathbf{A}_\bullet$ and $\mathbf{A}_\bullet'$ are representations of the quiver Q with a transformation $\tau : \mathbf{A}_\bullet \to \mathbf{A}_\bullet'$. Define the following representation $\mathbf{F}_\bullet^{\tau,u}$ of $F^{Q,u}$:

$$\mathbf{F}_{u_0}^{\tau,u} = \mathbf{A}_u \qquad\qquad \mathbf{F}_{v_e}^{\tau,u} = \mathbf{A}_v'$$

$$\mathbf{F}_{e_{LD}}^{\tau,u} = \mathbf{A}_e' \circ \tau_u \qquad\qquad \mathbf{F}_{e_{DL}}^{\tau,u} = \tau_{t(e)} \circ \mathbf{A}_e.$$

See Figure 3.4 for an illustration.

Constructing the floret now allows us to compute locally compatible features: in the case where the transformation is $\iota^{\mathcal{F}} : \mathbf{A}_\bullet^{\mathbf{S},\mathcal{F}} \to \mathbf{C}_\bullet^{\mathcal{F}}$ as in Proposition 3.3, the sections $\Gamma(\mathbf{F}_\bullet^{\iota^{\mathcal{F}},\sigma})$ are exactly the locally compatible features at σ.
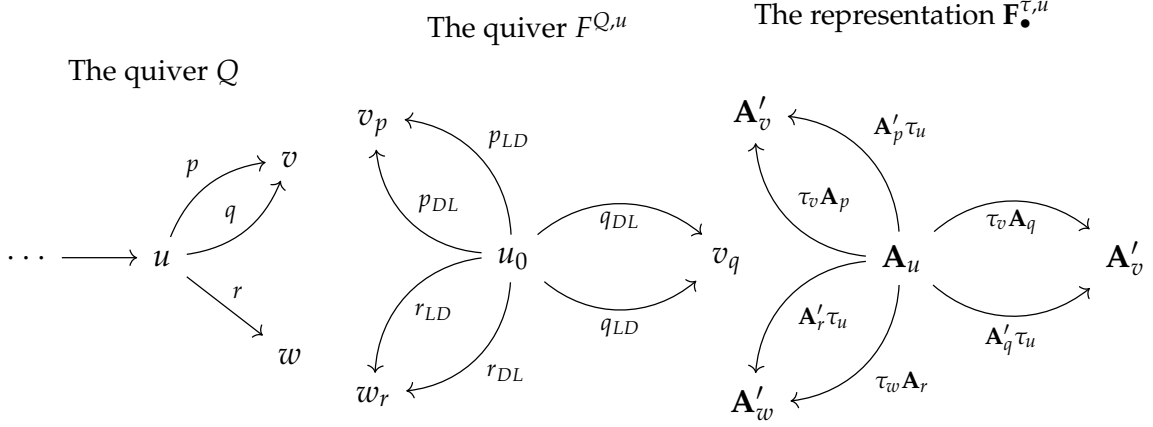
FIGURE 3. An illustration of the floret based at u and its representation.

3.3. Bicompatible features. In Section 4 below, we will consider the problem of discovering features which satisfy approximate versions of the local and global compatibility criteria. For this purpose, it will be helpful to construct a single quiver whose sections correspond to features which are both locally and globally compatible. The first step in this construction is a method for gluing quivers and their representations. Let $Q = (s, t : Q_1 \to Q_0)$ and $Q' = (s', t' : Q'_1 \to Q'_0)$ be two quivers.

DEFINITION 3.5. For each subset $R \subset Q_0 \times Q'_0$, let $Q \cup_R Q'$ be the **merged quiver** whose

(1) vertex set is the quotient $V := (Q_0 \sqcup Q'_0)/R$,
(2) edge set is the disjoint union $E := Q_1 \sqcup Q'_1$,

and source/target maps are given as follows. Let ρ be the composite $Q_0 \hookrightarrow Q_0 \sqcup Q'_0 \twoheadrightarrow V$, and define $\rho' : Q'_0 \to V$ similarly. If $e \in E$ lies in Q_1, then its source and target vertices are $\rho \circ s(e)$ and $\rho \circ t(e)$; and if e lies in Q'_1, then its source and target vertices are $\rho' \circ s'(e)$ and $\rho' \circ t'(e)$ respectively.

Fix representations $\mathbf{A}_\bullet$ of a quiver Q and $\mathbf{A}'_\bullet$ of another quiver Q'.

DEFINITION 3.6. Let $R \subset Q_0 \times Q'_0$ be any subset for which we have $\mathbf{A}_i = \mathbf{A}'_j$ whenever (i, j) lies in R. The **merged representation** of $\mathbf{A}_\bullet$ and $\mathbf{A}'_\bullet$, denoted $(\mathbf{A} \cup_R \mathbf{A}')_\bullet$, is the following representation of $Q \cup_R Q'$. To vertices, it assigns

$$(\mathbf{A} \cup_R \mathbf{A}')_i := \begin{cases} \mathbf{A}_i & i \in Q_0 \text{ and } (i, j) \notin R \text{ for any } j \in Q'_0 \\ \mathbf{A}'_i & i \in Q'_0 \text{ and } (j, i) \notin R \text{ for any } j \in Q_0 \\ \mathbf{A}_i = \mathbf{A}'_j & \text{if } (i, j) \in R \\ \mathbf{A}_j = \mathbf{A}'_i & \text{if } (j, i) \in R \end{cases}$$

And on edges, we have

$$(\mathbf{A} \cup_R \mathbf{A}')_e := \begin{cases} \mathbf{A}_e & \text{if } e \in Q_1, \\ \mathbf{A}'_e & \text{if } e \in Q'_1. \end{cases}$$

One may embed both $\mathrm{Tot}(\mathbf{A}_\bullet)$ and $\mathrm{Tot}(\mathbf{A}'_\bullet)$ within the total space of the merged representation in the natural way, and it is readily seen that the space of sections of $(\mathbf{A} \cup_R \mathbf{A}')_\bullet$ is (isomorphic to) the intersection $\Gamma(\mathbf{A}_\bullet) \cap \Gamma(\mathbf{A}'_\bullet)$.
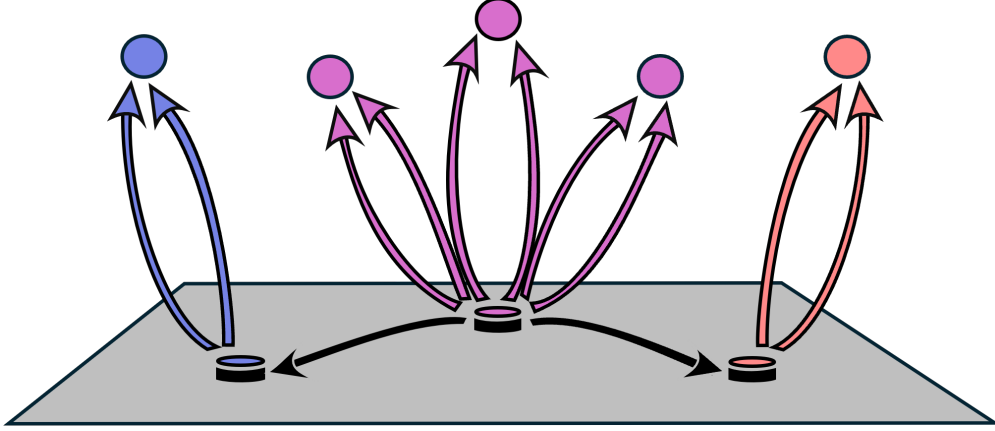

FIGURE 4. Illustration of the merged quiver

Returning to the case of interest, let $\mathbf{S}$ be a feature selector on a finite set X; consider a cover $\mathscr{U}$ of X and a subset $\mathscr{F} \subset X^*$. Let $\widehat{Q}^{\mathscr{U}}$ be the quiver given by the disjoint union of florets $\mathscr{I}^\sigma$ indexed by vertices $\sigma \in Q_0^{\mathscr{U}}$, and let $\widehat{\mathbf{A}}_\bullet^{\mathbf{S},\mathscr{F}}$ be the representation of $\widehat{Q}^{\mathscr{U}}$ given by the floret functors $\mathscr{B}^\sigma$. We define $R \subset (Q^{\mathscr{U}})_0 \times \widehat{Q}_0^{\mathscr{U}}$ as the collection

$$R := \left\{ (\sigma, \sigma_0) \mid \sigma \in Q_0^{\mathscr{U}} \right\},$$

where σ_0 is the central vertex of $\mathscr{I}^\sigma$.

DEFINITION 3.7. The **combined S-representation** (of $Q^{\mathscr{U}} \cup_R \widehat{Q}^{\mathscr{U}}$, along $\mathscr{F}$) is the merged representation $\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}} := \left(\mathbf{A}^{\mathbf{S},\mathscr{F}} \cup_R \widehat{\mathbf{A}}^{\mathbf{S},\mathscr{F}} \right)_\bullet$.

By design, the sections of $\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}$ correspond to relevant features which are both locally and globally compatible with $\mathscr{U}$. We call these sections the **bicompatible features** of $\mathbf{S}$.

EXAMPLE 3.8. Suppose $\mathscr{U} = \{U_1, U_2\}$ is a cover of X, such that $U_1 \cap U_2 \neq \varnothing$. Define $\tau_1 = \{U_1\}$, $\tau_2 = \{U_2\}$ and $\sigma = \{U_1, U_2\}$. The quiver $Q^{\mathscr{U}}$ is

$$\tau_1 \longleftarrow \sigma \longrightarrow \tau_2.$$

Consider a bicompatible feature, i.e., a section $x \in \Gamma(\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}})$. By definition, we have $\iota_{\tau_2}^{\mathscr{F}}(x_{\tau_2}) = \iota_\sigma^{\mathscr{F}}(x_\sigma) = \iota_{\tau_1}^{\mathscr{F}}(x_{\tau_1})$, hence

$$\Gamma(\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}) \cong \iota^{\mathscr{F}}(\mathbf{S}_{|\tau_1|}^{\mathscr{F}}) \cap \iota^{\mathscr{F}}(\mathbf{S}_{|\sigma|}^{\mathscr{F}}) \cap \iota^{\mathscr{F}}(\mathbf{S}_{|\tau_2|}^{\mathscr{F}}).$$

Thus, bicompatible features are the ones which are relevant for every $U \in \mathscr{U}$.

3.4. Dual Representations. If one is also interested in features which are relevant for some – but not all – $U \in \mathscr{U}$, then it becomes necessary to consider a different notion of bicompatibility. Given a quiver $Q = (s, t : Q_1 \to Q_0)$, its *dual* Q^* is the quiver with the same vertices and edges but with source and target maps interchanged, i.e., $s^* := t$ and $t^* := s$. Every representation $\mathbf{A}_\bullet$ of Q valued in $\mathbf{FdHilb}(\mathbb{K})$ induces a unique *dual representation* $\mathbf{A}_\bullet^*$ of Q^*, which is obtained by taking the adjoint of every edge map: namely, $(\mathbf{A}^*)_e := (\mathbf{A}_e)^*$ for each $e \in Q_1$.

When $\mathbf{A}_\bullet$ is the $\mathbf{S}$-representation (from Definition 2.4) of a feature selector $\mathbf{S}_\bullet$ along some input $\mathscr{F}$, the fact that $\iota_\sigma^\mathscr{F}$ and $\pi_\sigma^\mathscr{F}$ form an adjoint pair for each vertex $\sigma \in Q_0^\mathscr{U}$ implies that for every edge $\sigma \to \tau$ in $Q^\mathscr{U}$, we have

$$(\mathbf{A}^{\mathbf{S},\mathscr{F}})^*_{\tau \to \sigma} = (\pi_\tau^\mathscr{F} \circ \iota_\sigma^\mathscr{F})^* = \pi_\sigma^\mathscr{F} \circ \iota_\tau^\mathscr{F}.$$

DEFINITION 3.9. The **mixed S-representation** (of $(Q^\mathscr{U})^* \cup_R \widehat{Q}^\mathscr{U}$, along $\mathscr{F}$) is the merged representation $\mathbf{M}_\bullet^{\mathbf{S},\mathscr{F}} := \left((\mathbf{A}^{\mathbf{S},\mathscr{F}})^* \cup_R \widehat{\mathbf{A}}^{\mathbf{S},\mathscr{F}} \right)_\bullet$.

Let us revisit Example 3.8 with a view towards understanding the sections of $\mathbf{M}_\bullet^{\mathbf{S},\mathscr{F}}$. These sections are given by the direct sum

$$\Gamma(\mathbf{M}_\bullet^{\mathbf{S},\mathscr{F}}) \cong (\iota_{\tau_1}(\mathbf{S}_{|\tau_1|}^\mathscr{F}) \cap \iota_\sigma(\mathbf{S}_{|\sigma|}^\mathscr{F})^\perp) \oplus (\iota_{\tau_2}(\mathbf{S}_{|\tau_2|}^\mathscr{F}) \cap \iota_\sigma(\mathbf{S}_{|\sigma|}^\mathscr{F})^\perp) \oplus \Gamma(\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}).$$

In this case, $\iota_{\tau_1}(\mathbf{S}_{|\tau_1|}^\mathscr{F}) \cap \iota_\sigma(\mathbf{S}_{|\sigma|}^\mathscr{F})^\perp$ corresponds to features that are relevant to U_1 but not to $U_1 \cap U_2$. Note that

$$(\iota_{\tau_1}(\mathbf{S}_{|\tau_1|}^\mathscr{F}) \cap \iota_\sigma(\mathbf{S}_{|\sigma|}^\mathscr{F})^\perp) \cap (\iota_{\tau_2}(\mathbf{S}_{|\tau_2|}^\mathscr{F}) \cap \iota_\sigma(\mathbf{S}_{|\sigma|}^\mathscr{F})^\perp)$$

might not be trivial, i.e. there could be a feature relevant to both U_1 and U_2 but not to $U_1 \cap U_2$, and hence this feature corresponds to two independent sections.

4. The Quiver Laplacian and Approximate Sections

So far, we have reframed compatibility-testing of feature selectors in terms of finding the sections of a quiver representation. However, the space of sections might be trivial [30, Sec 5.1]. Rather than seeking exact solutions, we endeavour to find approximate sections.

4.1. The Quiver Laplacian. Let $\mathbf{FdHilb}(\mathbb{K})$ denote the category of finite-dimensional Hilbert spaces over the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$; every morphism $A : U \to V$ in this category admits an *adjoint* morphism $A^* : V \to U$ characterised by $\langle Au, v \rangle = \langle u, A^* v \rangle$ for all (u, v) in $U \times V$. Here we fix a finite quiver $Q = (s, t : Q_1 \to Q_0)$ and consider a representation $\mathbf{A}_\bullet$ of Q valued in $\mathbf{FdHilb}(\mathbb{K})$. We recall the total space from (1) and define *target space* of $\mathbf{A}_\bullet$ as the direct product

$$\mathrm{Tar}(\mathbf{A}_\bullet) := \prod_{e \in Q_1} \mathbf{A}_{t(e)}.$$

Since we have assumed that Q is finite, both these spaces inherit a Hilbert space structure from the individual factors of the form $\mathbf{A}_v$. We will denote the identity map on $\mathbf{A}_v$ by id_v rather than $\mathrm{id}_{\mathbf{A}_v}$.

DEFINITION 4.1. Given an **FdHilb**$(\mathbb{K})$-valued quiver representation $\mathbf{A}_\bullet$ of Q, its **boundary operator**

$$B_{\mathbf{A}} : \mathrm{Tot}(\mathbf{A}_\bullet) \to \mathrm{Tar}(\mathbf{A}_\bullet)$$

is the linear map given in component form by

$$(B_{\mathbf{A}})_{e,v} = \begin{cases} \mathbf{A}_e - \mathrm{id}_v & \text{if } s(e) = t(e) = v, \\ \mathbf{A}_e & \text{if } s(e) = v \text{ and } t(e) \neq v, \\ -\mathrm{id}_v & \text{if } s(e) \neq v \text{ and } t(e) = v, \\ 0 & \text{otherwise.} \end{cases}$$

If the quiver Q is finite, then it follows from a simple calculation that $\ker B_{\mathbf{A}}$ is isomorphic (as a Hilbert space) to the space $\Gamma(Q; \mathbf{A}_\bullet)$ of $\mathbf{A}_\bullet$'s sections. From the boundary operator of a quiver representation, we can construct another associated operator, the Laplacian, which will allow us to compute approximate sections.

DEFINITION 4.2. The **Laplacian** of $\mathbf{A}_\bullet$ is the endomorphism $L_{\mathbf{A}} : \mathrm{Tot}(\mathbf{A}_\bullet) \to \mathrm{Tot}(\mathbf{A}_\bullet)$ given by

$$L_{\mathbf{A}} = B_{\mathbf{A}}^* B_{\mathbf{A}}.$$

In component form, we have

$$(L_{\mathbf{A}})_{v,v} = \sum_{\substack{e \in Q_1 \\ s(e)=v}} \mathbf{A}_e^* \mathbf{A}_e + \sum_{\substack{e \in Q_1 \\ t(e)=v}} \mathrm{id}_v - \sum_{\substack{e \in Q_1 \\ s(e)=t(e)=v}} [\mathbf{A}_e^* + \mathbf{A}_e]$$

and for $u \neq v$

$$(L_{\mathbf{A}})_{u,v} = - \sum_{\substack{e \in Q_1 \\ s(e)=u \\ t(e)=v}} \mathbf{A}_e^* - \sum_{\substack{e \in Q_1 \\ s(e)=v \\ t(e)=u}} \mathbf{A}_e.$$

Since $\ker L_{\mathbf{A}} = \ker B_{\mathbf{A}}$, the kernel of $L_{\mathbf{A}}$ also computes the sections of $\mathbf{A}_\bullet$. The main advantage of considering $L_{\mathbf{A}}$ rather than $B_{\mathbf{A}}$ in this context is that it is a Hermitian and positive semi-definite matrix. As such, it enjoys favourable spectral properties; in particular, all of its eigenvalues are all real and non-negative. We shall order these eigenvalues in increasing fashion, with multiplicity:

$$\lambda_1(L_{\mathbf{A}}) \leq \cdots \leq \lambda_n(L_{\mathbf{A}})$$

where $n := \dim \mathrm{Tot}(\mathbf{A}_\bullet)$. If $\varphi : \mathbf{A}_\bullet^1 \to \mathbf{A}_\bullet^2$ is an isomorphism of quiver representations with each φ_v a unitary map then

$$L_{\mathbf{A}^1} = \Phi \circ L_{\mathbf{A}^2} \circ \Phi^*,$$

where $\Phi = \prod_{v \in Q_0} \varphi_v$, thus we may speak unambiguously of the eigenvalues $\lambda_i(\mathbf{A}_\bullet) := \lambda_i(L_{\mathbf{A}})$ of a quiver representation $\mathbf{A}_\bullet$.

4.2. Approximate Sections. To measure the distance in $\mathrm{Tot}(\mathbf{A}_\bullet)$ between an arbitrary vector x and the space $\Gamma(\mathbf{A}_\bullet)$ of sections, we consider the **Dirichlet energy**:

$$\mathcal{E}_\mathbf{A}(x) := \langle x, L_\mathbf{A} x \rangle = \sum_{e \in Q_1} \|\mathbf{A}_e(x_{s(e)}) - x_{t(e)}\|_{\mathbf{A}_{t(e)}}^2 \in \mathbb{R}_{\geq 0}$$

Here $\langle -, - \rangle$ is the induced inner product on $\mathrm{Tot}(\mathbf{A}_\bullet)$, while and x_v denotes the component of x in $\mathbf{A}_v$, and $\| - \|$ denotes the norm induced by the inner product. As $L_\mathbf{A}$ is Hermitian, we can form an orthonormal eigenbasis $e_1, \ldots, e_n$ where e_i corresponds to $\lambda_i(\mathbf{A}_\bullet)$. Writing x in this basis, we have

$$\mathcal{E}_\mathbf{A}(x) = \mathcal{E}_\mathbf{A}\left(\sum_i x_i e_i\right) = \sum_i \lambda_i(\mathbf{A}_\bullet) x_i^2.$$

Thus, $\sqrt{\mathcal{E}_\mathbf{A}(x)}$ is the $\lambda_\bullet$-weighted distance from x to $\Gamma(\mathbf{A}_\bullet)$; and moreover,

$$\sum_{i : \lambda_i(\mathbf{A}) = 0} x_i e_i$$

is the closest vector to x in $\Gamma(\mathbf{A}_\bullet)$. This motivates the following definition. (A similar notion for cellular sheaves was introduced in [20]).

DEFINITION 4.3. For a quiver Q and an **FdHilb**$(\mathbb{K})$-valued representation $\mathbf{A}_\bullet$ of Q, and ε-*approximate section* of $\mathbf{A}_\bullet$ is an $x \in \mathrm{Tot}(\mathbf{A}_\bullet)$ such that $\|x\| = 1$ and $\mathcal{E}_\mathbf{A}(x) \leq \varepsilon$.

By Theorem B.1, if $x \in \mathrm{span}\{e_1, \ldots, e_k\}$ and $\|x\| = 1$, then x is a λ_k-approximate section. Note this is not exhaustive: there can exist λ_k-approximate sections of $\mathbf{A}_\bullet$ that are not in the span of the first k eigenvectors, e.g. $\lambda_1 x_1 e_1 + \lambda_{k+1} x_{k+1} e_{k+1}$ for small enough x_{k+1}. Regardless, to compute approximate sections of a quiver representation one can compute eigenvectors of $L_\mathbf{A}$ with small eigenvalues. As we have connected approximate sections with the spectra of the quiver representation, we now want to understand how the spectra changes if we perturb the quiver representation (§5), or the underlying quiver (§6).

5. Spectral Stability of Feature Selectors

Let Q be a quiver and suppose $\mathbf{A}_\bullet$ and $\widehat{\mathbf{A}}_\bullet$ are two representations of Q. Suppose τ is a collection $\{(\tau_v : \mathbf{A}_v \to \widehat{\mathbf{A}}_v) : v \in Q_0\}$ of linear maps. Define the *defect* of τ at an edge $e \in Q_1$ to be

$$\partial(\tau, e) = \left\|\widehat{\mathbf{A}}_e \tau_{s(e)} - \tau_{t(v)} \mathbf{A}_e\right\|,$$

Here, and henceforth, the expression $\| - \|$ when applied to a linear map between Hilbert spaces will always indicate the operator norm — for $M : V \to W$, the norm $\|M\|$ equals the supremum of $\|Mx\|_W$ over unit vectors $x \in V$. Now the *total defect* of τ is:

$$\partial(\tau) = \left(\sum_{e \in Q_1} \partial(\tau, e)^2\right)^{\frac{1}{2}}.$$

The defect of τ is zero if and only if when τ prescribes a morphism of quiver representations. Given bases of $\mathrm{Tot}(\mathbf{A}_\bullet)$ and $\mathrm{Tot}(\widehat{\mathbf{A}}_\bullet)$, we view τ as a block diagonal matrix. Let

τ^+ denote its Moore-Penrose pseudoinverse. The values $\|\tau\|$ and $\|\tau^+\|$ are unique up to unitary transforms of $\mathrm{Tot}(\mathbf{A}_\bullet)$ and $\mathrm{Tot}(\widehat{\mathbf{A}}_\bullet)$. Finally, we let

$$\kappa(\tau) := \|\tau\|\,\|\tau^+\|.$$

be the *generalised condition number* of τ.

We seek to describe how the existence of τ forces a relationship between the Laplacian spectra of $\mathbf{A}_\bullet$ and $\widehat{\mathbf{A}}_\bullet$. The first step in this direction is the following lemma.

LEMMA 5.1. *For a matrix τ, if $x \in (\ker \tau)^\perp$ then*

$$\|\tau x\| \geq \|\tau^+\|^{-1}\|x\|.$$

PROOF. Let $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ be left and right singular vectors of τ respectively. Let $x = \sum_{i=1}^n x_i v_i$. Then $\tau x = \sum_{i=1}^{\min\{n,m\}} \sigma_i x_i u_i$, but as $x \in \ker \tau^\perp$ we have that $x_j = 0$ when $\sigma_j = 0$ or $j > \min\{m, n\}$. Thus

$$\|\tau x\| \geq \left\| \sum_{i=1}^{\min\{n,m\}} \sigma_+ x_i u_i \right\| = \sigma_+ \|x\| = \|\tau^+\|^{-1}\|x\|$$

where σ_+ is the smallest non-zero singular value of τ.   □

The next theorem describes how the existence of a transformation between quiver representations constrains their eigenvalues. In the statement below (and henceforth), we use $\mathrm{null}(f)$ as a shorthand for the dimension of the kernel of a linear map $f : U \to V$ of finite-dimensional Hilbert spaces.

THEOREM 5.2. *Suppose $\mathbf{A}_\bullet$ and $\widehat{\mathbf{A}}_\bullet$ are two representations of a quiver Q whose total spaces have dimensions n and m respectively. Let $\tau = \{\tau_v : \mathbf{A}_v \to \widehat{\mathbf{A}}_v \mid v \in Q_0\}$ be a collection of linear maps, viewed as a single map $\mathrm{Tot}(\mathbf{A}_\bullet) \to \mathrm{Tot}(\widehat{\mathbf{A}}_\bullet)$. Set $q := \mathrm{null}(\tau)$; for every integer k such that $1 \leq k \leq n - q$, we have*

$$\lambda_k(\widehat{\mathbf{A}}) \leq \kappa(\tau)^2 \lambda_{k+q}(\mathbf{A}) + \partial(\tau)\|\tau^+\| \left[2\kappa(\tau)\lambda_{k+q}(\mathbf{A})^{\frac{1}{2}} + \partial(\tau)\|\tau^+\| \right].$$

In particular, if τ is a morphism of quiver representations, then $\partial(\tau)$ vanishes and hence

$$\lambda_k(\widehat{\mathbf{A}}) \leq \kappa(\tau)^2 \lambda_{k+q}(\mathbf{A}).$$

PROOF. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be eigenvectors of $\mathbf{A}_\bullet$ and $\widehat{\mathbf{A}}_\bullet$ respectively, ordered by increasing eigenvalue. Suppose k is an integer such that $1 \leq k \leq n - q$. Let $l = k + q$ and define $X_l = \mathrm{span}\{x_1, \ldots, x_l\}$ and $Y^k = \mathrm{span}\{y_k, \ldots, y_m\}$. As $l = k + q$ we have that $\dim \tau(X_l) \geq k$ thus $\dim \tau(X_l) + \dim Y^k$ is at least $m + 1$ hence $\tau(X_l)$ and Y^k intersect non-trivially. Thus there exists some non-zero $y \in \tau(X_l) \cap Y^k$, and as $y \in Y^k$, by Theorem B.1 it holds that

$$\lambda_k(\widehat{\mathbf{A}}) \leq \frac{1}{\|y\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e y_{s(e)} - y_{t(e)} \right\|^2.$$

Now we shall bound the right-hand side of the above. Since $y \in \tau(X_l)$ there exists some $x \in X_l \cap (\ker \tau)^\perp$ such that $\tau x = y$. Using Lemma 5.1 have that

$$\frac{1}{\|\tau x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2 \leq \frac{\|\tau^+\|^2}{\|x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2 .$$

Then applying the triangle inquality gives us

$$\frac{1}{\|\tau x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2$$

$$\leq \frac{\|\tau^+\|^2}{\|x\|^2} \sum_{e \in Q_1} \left\| \tau_{t(e)} \mathbf{A}_e x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2$$

$$+ 2 \frac{\|\tau^+\|^2}{\|x\|^2} \sum_{e \in Q_1} \left\| \tau_{t(e)} \mathbf{A}_e x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\| \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} \mathbf{A}_e x_{s(e)} \right\|$$

$$+ \frac{\|\tau^+\|^2}{\|x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} \mathbf{A}_e x_{s(e)} \right\|^2$$

and then with the Cauchy-Schwarz inequality we have

$$\frac{1}{\|\tau x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2$$

$$\leq \kappa(\tau)^2 \frac{1}{\|x\|^2} \sum_{e \in Q_1} \left\| \mathbf{A}_e x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2$$

$$+ \kappa(\tau) \|\tau^+\| \left[\frac{1}{\|x\|^2} \sum_{e \in Q_1} \left\| \mathbf{A}_e x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2 \right]^{\frac{1}{2}} \partial(\tau) + \|\tau^+\|^2 \partial(\tau)^2$$

and finally since $x \in X_l$, we may apply Theorem B.1 to obtain

$$\frac{1}{\|\tau x\|^2} \sum_{e \in Q_1} \left\| \widehat{\mathbf{A}}_e \tau_{s(e)} x_{s(e)} - \tau_{t(e)} x_{t(e)} \right\|^2$$

$$\leq \kappa(\tau)^2 \lambda_l(\mathbf{A}) + 2\kappa(\tau) \|\tau^+\| \lambda_l(\mathbf{A})^{\frac{1}{2}} \partial(\tau) + \|\tau^+\|^2 \partial(\tau)^2$$

completing the proof. □

We shall now describe how Theorem 5.2 gives us a bound on the spectral pseudo-metric between quiver representations. If the total spaces of $\mathbf{A}$ and $\widehat{\mathbf{A}}_\bullet$ have the same dimension, then we can directly compare the associated Laplacian spectra. Following [14], we use the Wasserstein metric to perform such comparisons in order to account for the case where total spaces have different dimensions.

DEFINITION 5.3. For each real number $r \geq 0$, let $B(Q, r)$ be the set of representations $\mathbf{A}_\bullet$ of Q which satisfy $\max_{e \in Q_1} \|A_e\| \leq r$.

By design, if $\mathbf{A}_\bullet \in B(Q, r)$, then its eigenvalues are bounded by $|Q_1|(r+1)^2$.

Given a representation $\mathbf{A}_\bullet$ of Q with total dimension n, its *spectral measure* is

$$\mu_\mathbf{A} = \sum_{i=1}^{n} \delta_{\lambda_i(\mathbf{A})}$$

where $\delta_{\lambda_i(\mathbf{A})}$ is the Dirac measure concentrated at $\lambda_i(\mathbf{A})$. Given two probability measures μ_1, μ_2 on $\mathbb{R}_{\geq 0}$, a *coupling* γ of μ_1 and μ_2 is any probability measure on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ whose marginals on each factor are μ_1 and μ_2 respectively. Denote by $\Gamma(\mu_1, \mu_2)$ the set of all such couplings. Then the Wasserstein p-metric between μ_1 and μ_2 is defined as

$$W_p(\mu_1, \mu_2) = \inf_{\gamma \in \Gamma(\mu_1, \mu_2)} \left(\int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} \|x - y\|^p \, \mathrm{d}\gamma(x, y) \right)^{\frac{1}{p}}$$

Adapting the proof of Theorem 6.3 in [**14**] we have the following:

THEOREM 5.4. *Suppose* $\mathbf{A}_\bullet, \widehat{\mathbf{A}}_\bullet \in B(Q, r)$ *with total dimensions n and m respectively, where* $n \geq m$. *Suppose τ is a non-zero transformation from $\mathbf{A}_\bullet$ to $\widehat{\mathbf{A}}_\bullet$. If* $\mathrm{rank}(\eta) \geq \mathrm{null}(\eta^*)$ *then*

$$W_1(\mu_\mathbf{A}, \mu_{\widehat{\mathbf{A}}}) \leq C$$

where C depends on τ, n and r. Moreover, if $\|\tau\|, \|\tau^+\| \leq 1 + \varepsilon$ and $\partial(\tau), \partial(\tau^*) \leq \varepsilon$ for some $\varepsilon > 0$, *then we have*

$$\lim_{\varepsilon \to 0} C = \frac{2 \, \mathrm{null}(\tau) + 2 \, \mathrm{null}(\tau^*)}{n} |Q_1|(r+1)^2,$$

which, if $\mathrm{null}(\tau) + \mathrm{null}(\tau^*) < n/2$, *will be less than naïve bound of* $|Q_1|(r+1)^2$.

PROOF. Define $c_\tau = \mathrm{null}(\tau) + \mathrm{null}(\tau^*)$. We will now define a coupling γ between $\mu_\mathbf{A}$ and $\mu_{\widehat{\mathbf{A}}}$. For each positive integer p, let $[p]$ denote the set $\{1, 2, \ldots, p-1, p\}$. Setting $\ell = \mathrm{null}(\tau^*)$ and and $u = n - \mathrm{null}(\tau)$, define the following partition of $[n] \times [m]$:

$$\begin{aligned}
P_1 &= \{(i, i) : i \in [u] \setminus [\ell]\} \\
P_2 &= \{(i, j) : i \in [\ell] \text{ and } j \in [\ell]\} \\
&\quad \cup \{(i, j) : i \in [\ell] \text{ and } j \in [m] \setminus [u]\} \\
&\quad \cup \{(i, j) : i \in [n] \setminus [u] \text{ and } j \in [\ell]\} \\
&\quad \cup \{(i, j) : i \in [n] \setminus [u] \text{ and } j \in [m] \setminus [u]\} \\
P_3 &= \{(i, j) : i \in [\ell] \text{ and } j \in [u] \setminus [\ell]\} \\
&\quad \cup \{(i, j) : i \in [n] \setminus [u] \text{ and } j \in [u] \setminus [\ell]\} \\
P_4 &= ([n] \times [m]) \setminus (P_1 \cup P_2 \cup P_3).
\end{aligned}$$

Now define the desired coupling γ as

$$\gamma = \sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} w_{ij} \delta_{(\lambda_i(\mathbf{A}), \lambda_j(\widehat{\mathbf{A}}))} \qquad \text{where} \qquad w_{ij} = \begin{cases} \frac{1}{n} & (i,j) \in P_1, \\ \frac{1}{m(n-c_\tau)} & (i,j) \in P_2, \\ \left(\frac{1}{m} - \frac{1}{n}\right)\frac{1}{n-c_\tau} & (i,j) \in P_3, \\ 0 & (i,j) \in P_4. \end{cases}$$

Set $R = |Q_1|(r+1)^2$, which is an upper bound for the eigenvalues of $\mathbf{A}_\bullet$. Then using Theorem 5.2 for τ and τ^* and substituting R gives us

$$\kappa(\tau)^{-2}\lambda_{i-\text{null}(\tau^*)}(\mathbf{A}_\bullet) - a_{\tau^*,R} \le \lambda_i(\widehat{\mathbf{A}}_\bullet) \le \kappa(\tau)^2\lambda_{i+\text{null}(\tau)}(\mathbf{A}_\bullet) + b_{\tau,R}$$

with the constants

$$a_{\tau^*,R} = 2\kappa(\tau)^{-1}\|\tau^+\|\partial(\tau^*)\sqrt{R} + \kappa(\tau)^{-2}\|\tau^+\|^2\partial(\tau^*)^2, \text{ and}$$

$$b_{\tau,R} = 2\kappa(\tau)\|\tau^+\|\partial(\tau)\sqrt{R} + \|\tau^+\|^2\partial(\tau)^2.$$

Computing cost of γ gives us

$$W_1(\mu_\mathbf{A}, \mu_{\widehat{\mathbf{A}}}) \le \frac{1}{n}\left[\sum_{k=\text{null}(\tau^*)+1}^{n-\text{null}(\tau)} |\lambda_k(\mathbf{A}_\bullet) - \lambda_k(\widehat{\mathbf{A}}_\bullet)| + c_\tau R\right]$$

$$\le \frac{1}{n}\left[\sum_{k=\text{null}(\tau^*)+1}^{n-\text{null}(\tau)} (\kappa(\tau)^2\lambda_{k+\text{null}(\tau)}(\mathbf{A}_\bullet) - \kappa(\tau)^{-2}\lambda_{k-\text{null}(\tau^*)}(\mathbf{A}))\right.$$

$$\left. + (n - c_\tau)(a_{\tau^*,R} + b_{\tau,R}) + c_\tau R\right].$$

Rearranging the sum on the right side yields

$$\frac{1}{n}\left[c_\tau\kappa(\tau)^2 R + (n - 2c_\tau)(\kappa(\tau)^2 - \kappa(\tau)^{-2})R + (n - c_\tau)(a_{\tau^*,R} + b_{\tau,R}) + c_\tau R\right],$$

which is the desired C. $\qquad\square$

Suppose A, B are a pair of orthogonal matrices with dimensions $n \times k$ and $n \times l$ respectively. Then the matrix A^*B is an orthogonal projection $\text{img } B \to \text{img } A$, where $\text{img } A$ denotes the image of A. Let the singular values of A^*B be $\sigma_1, \ldots, \sigma_r$ where $r = \min\{k, l\}$. Then the principal angles between $\text{img } A$ and $\text{img } B$ are defined by $\cos \theta_i = \sigma_i$ and the Grassmannian metric between $\text{img } A$ and $\text{img } B$ as defined by [37] is

$$d_{\mathbf{Gr}_\infty}(\text{img } A, \text{img } B) := \left(\sum_{i=1}^r \theta_i^2\right)^{\frac{1}{2}}.$$

COROLLARY 5.5. *Suppose* $\mathbf{S}, \mathbf{T} : \text{Sub}(X^*) \to \mathbf{Gr}_\infty(X^*)$ *are two feature selectors on* X *and suppose* $\mathscr{U}$ *is a cover of* X *and* $\mathscr{F} \subset X^*$. *Suppose that*

$$\max_{\sigma \in Q_0^{\mathscr{U}}} d_{\mathbf{Gr}_\infty}(\mathbf{S}_{|\sigma|}^{\mathscr{F}}, \mathbf{T}_{|\sigma|}^{\mathscr{F}}) \le \varepsilon < \frac{\pi}{2}$$

for some $\varepsilon > 0$, and define

$$c = \sum_{\sigma \in Q_0^{\mathscr{U}}} |\dim \mathbf{S}_{|\sigma|}^{\mathscr{F}} - \dim \mathbf{T}_{|\sigma|}^{\mathscr{F}}|.$$

Then, $W_1(\mu_{\mathbf{A}^{\mathbf{S},\mathscr{F}}}, \mu_{\mathbf{A}^{\mathbf{T},\mathscr{F}}})$ is no larger than

$$\frac{4|Q_1|}{n}\left[(1 + \cos^2 \varepsilon)c + \left(2\sec\varepsilon\tan\varepsilon + (1+\sec^2\varepsilon)\tan^2\varepsilon + \sec^2\varepsilon - \cos^2\varepsilon\right)(n - c)\right]$$

where $n = \max\{\dim \mathrm{Tot}(\mathbf{A}_{\bullet}^{\mathbf{S},\mathscr{F}}), \dim \mathrm{Tot}(\mathbf{A}_{\bullet}^{\mathbf{T},\mathscr{F}})\}$.

PROOF. Assume that $\dim \mathbf{S}_{\bullet}^{\mathscr{F}} \geq \dim \mathbf{T}_{\bullet}^{\mathscr{F}}$. Define the transformation η with components $\eta_\sigma = \pi_\sigma^{\mathbf{T}} \iota_\sigma^{\mathbf{S}}$. Then $\cos\varepsilon \leq \|\eta\| \leq 1$ and $1 \leq \|\eta^+\| \leq (\cos\varepsilon)^{-1}$. In order to bound $\partial(\eta)$ note that

$$\|\iota_\sigma^{\mathbf{S}}\pi_\sigma^{\mathbf{S}} - \iota_\sigma^{\mathbf{T}}\pi_\sigma^{\mathbf{T}}\|_2 = \sin\theta_{\sigma,\max} \leq \sin\varepsilon,$$

where $\theta_{\sigma,\max}$ is the largest principal angle between $\mathbf{S}_{|\sigma|}$ and $\mathbf{T}_{|\sigma|}$. This is known as the projection metric [37] where the equality can be seen from Theorem 2.6.1 in [13]. This then implies that, for each edge $\sigma \to \tau$, we have $\partial(\eta, \sigma \to \tau) \leq 2\sin\varepsilon$ thus $\partial(\eta) \leq 2\sqrt{|Q_1|}\sin\varepsilon$. Similarly, $\partial(\eta^*) \leq 2\sqrt{|Q_1|}\sin\varepsilon$. Applying Theorem 5.4 together with these bounds provide the result. $\square$

6. Eigenvalue Inequalities for Quiver Operations

So far, we have described our problem as computing sections of a quiver representation. Here we seek to simplify the representation and the underlying quiver whilst preserving its space of sections and bounding the change in its eigenvalues. Here we will establish similar eigenvalue inequalities when the quiver itself is deformed via certain natural operations. The most convenient framework for extracting such inequalities is to upgrade the given quiver representation to a sheaf, as described below.

6.1. Sheaves over Quivers. Fix a quiver $Q = (s, t : Q_1 \to Q_0)$.

DEFINITION 6.1. An **FdHilb**$(\mathbb{K})$-valued **sheaf** $\mathscr{A}_\bullet$ over Q consists of:

- A finite-dimensional Hilbert space $\mathscr{A}_v$ for each vertex $v \in Q_0$,
- a finite-dimensional Hilbert space $\mathscr{A}_e$ for each edge $e \in Q_1$, and
- for each for each edge $e \in Q_1$, a pair of linear maps

$$\mathscr{A}_{s(e)\to e} : \mathscr{A}_{s(e)} \to \mathscr{A}_e \quad \text{and} \quad \mathscr{A}_{e\leftarrow t(e)} : \mathscr{A}_{t(e)} \to \mathscr{A}_e.$$

Such sheaves are closely related to representations of quivers. Every representation $\mathbf{A}_\bullet$ over Q induces a sheaf $\mathscr{A}_\bullet$ over Q by setting $\mathscr{A}_{s(e)\to e} = \mathbf{A}_e$ and $\mathscr{A}_{e\leftarrow t(e)} = \mathrm{id}_{t(e)}$. Conversely, every sheaf evidently induces a representation on the subdivision of Q where every edge $e : u \to v$ is decomposed into the zigzag $u \to e \leftarrow v$. Thus, there is a natural notion of Laplacian for sheaves over a quiver, a special case[4] of which appears in [16].

---

[4]The only difference between our quiver sheaf Laplacian and the sheaf Laplacians of [16] over one-dimensional cell complexes is that we allow edges to have the same source and target vertex. On the other hand, cellular sheaves and their Laplacians are also defined over regular cell complexes of dimension > 1.

DEFINITION 6.2. Suppose $\mathscr{A}_\bullet$ is a **FdHilb**$(\mathbb{K})$-valued sheaf over a quiver Q. Define $\mathrm{Tot}(\mathscr{A}_\bullet) = \prod_{v \in Q_0} \mathscr{A}_v$ and $\mathrm{Tar}(\mathscr{A}_\bullet) = \prod_{e \in Q_1} \mathscr{A}_e$.

(1) The **boundary operator** of $\mathbf{A}_\bullet$ is the linear map $B_\mathscr{A} : \mathrm{Tot}(\mathscr{A}_\bullet) \to \mathrm{Tar}(\mathscr{A}_\bullet)$ of $\mathscr{A}_\bullet$ given in component form by

$$(B_\mathscr{A})_{e,v} = \begin{cases} \mathscr{A}_{v \to e} - \mathscr{A}_{e \leftarrow v} & \text{if } s(e) = t(e) = v, \\ \mathscr{A}_{v \to e} & \text{if } s(e) = v \text{ and } t(e) \neq v, \\ -\mathscr{A}_{e \leftarrow v} & \text{if } s(e) \neq v \text{ and } t(e) = v, \\ 0 & \text{otherwise.} \end{cases}$$

(2) The **Laplacian** of $\mathscr{A}_\bullet$ is then defined as

$$L_\mathscr{A} = B_\mathscr{A}^* B_\mathscr{A}.$$

(3) Finally, the **Dirichlet energy** for $\mathscr{A}$ is the function $\mathcal{E}_\mathscr{A} : \mathrm{Tot}(\mathscr{A}_\bullet) \to \mathbb{R}$ given by

$$\mathcal{E}_\mathscr{A}(x) = \sum_{e \in Q_1} \left\| \mathscr{A}_{s(e) \to e}(x_{s(e)}) - \mathscr{A}_{e \leftarrow t(e)}(x_{t(e)}) \right\|_{\mathscr{A}_e}^2.$$

If one constructs a sheaf $\mathscr{A}_\bullet$ from a quiver representation $\mathbf{A}_\bullet$ of Q as described above, then the sheaf boundary operator $B_\mathscr{A}$ coincides exactly with the quiver boundary operator $B_\mathbf{A}$; therefore the Laplacians of $\mathbf{A}$ and $\mathscr{A}$, as well as their spectra, also coincide in this case.

6.2. Operations on the Quiver. This section describes how the spectrum of a sheaf, or indeed a representation, over a quiver changes under various combinatorial operations on the quiver. Removing an edge or a vertex gives bounds on the change in spectrum, as Proposition A.1 describes, however these may reduce the dimension of the space of sections. Propositions 6.5, 6.7, and 6.9 describe operations on the quiver, dependent on the sheaf, that preserve the space of sections. These are described for sheaves over a quiver but all these operations also preserve quiver representations. These intermediate results are used to simplify quivers (while bounding the change in eigenvalues) in our Proofs of Theorems 7.1 and 7.2.

DEFINITION 6.3. For a quiver Q, define an *edge pairing* as a collection P of pairs of distinct neighbouring edges $(e_1^1, e_2^1), \ldots (e_1^m, e_2^m)$ of Q such no two edges are repeated, i.e. $e_j^i \neq e_l^k$ unless $i = k$ and $j = l$. To simplify notation assume that each pair (e_1^i, e_2^i), the edges e_1^i and e_2^i are directed away from their common vertex u^i. Label the other vertices for e_1^i and e_2^i, possibly the same vertex, as v^i and w^i respectively.
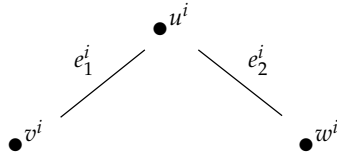
We then define Q^P as the *P-homotopic quiver* to Q where Q^P has the same nodes and edges as Q except with each edge e_2^i replaced by an edge e_3^i between v^i and w^i.

$$\bullet\, u^i$$

$$e_1^i$$

$$e_3^i$$

$$\bullet\, v^i \quad\underline{\hspace{5cm}}\quad \bullet\, w^i$$
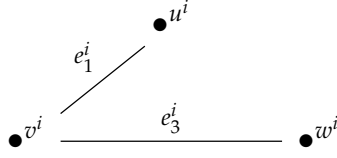
For $\mathscr{A}_\bullet$ is a sheaf over Q and a pairing P of Q, say that $\mathscr{A}_\bullet$ is compatible with P if for each pair $(e_1^i, e_2^i) \in P$ we have that $\mathscr{A}_{e_1^i} = \mathscr{A}_{e_2^i}$ and $\mathscr{A}_{u^i \to e_1^i} = \mathscr{A}_{u^i \to e_2^i}$. Such a compatible sheaf induces a sheaf on Q^P as follows: define the sheaf $\mathscr{A}_\bullet^P$ on Q^P as the same as $\mathscr{A}_\bullet$ on the common edges of Q and Q^P and otherwise by $\mathscr{A}_{e_3^i}^P = \mathscr{A}_{e_2^i}$, $\mathscr{A}_{e_3^i \leftarrow e_3^i}^P = \mathscr{A}_{e_1^i \leftarrow e_1^i}$ and $\mathscr{A}_{w^i \to e_3^i}^P = \mathscr{A}_{e_2^i \leftarrow w^i}$ for each i. The next proposition shows the i-th eigenvalue of $\mathscr{A}_\bullet^P$ is bounded by a scalar multiple of the i-th eigenvalue of $\mathscr{A}_\bullet$. First we will need a lemma.

LEMMA 6.4. *Suppose $X, Y, Z \in \mathbb{R}^n$. Then*

$$\|X - Z\|^2 + \|Z - Y\|^2 \leq \varphi^2(\|X - Y\|^2 + \|Y - Z\|^2)$$

where φ is the golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$. This is also true for any metric space (M, d) and can be written as

$$d(X, Z)^2 \leq \varphi^2 d(X, Y)^2 + \varphi d(Y, Z)^2.$$

PROOF. Let us abbreviate

$$x = \|Y - Z\| \qquad y = \|X - Z\| \qquad z = \|X - Y\|.$$

If any of x, y or z are zero, then the proposition is immediately true, so assume that x, y and z are strictly positive. By the triangle inequality, we have

$$\frac{x^2 + y^2}{x^2 + z^2} \leq \frac{x^2 + (x+z)^2}{x^2 + z^2}.$$

We then want to know when the expression on the right side is maximised. As $x, z > 0$, there exists $u > 0$ for which $x = uz$. Substituting this into the right side gives

$$\frac{2u^2 + 2u + 1}{u^2 + 1},$$

which is maximised when $u = \varphi$. Thus, the maximum value of the ratio is φ^2. If X, Y, Z are colinear with Y between X and Z and $\|Y - Z\| = \varphi\|X - Y\|$, then the maximum is attained, which shows that this inequality is sharp. $\qquad\square$

PROPOSITION 6.5. *For Q a quiver and P an edge pairing of Q, suppose that $\mathbf{A}_\bullet$ is a sheaf on Q compatible with P. Then the eigenvalues of $\mathscr{A}_\bullet$ and $\mathscr{A}_\bullet^P$ are related by*

$$\varphi^{-2}\lambda_j(\mathscr{A}_\bullet) \leq \lambda_j(\mathscr{A}_\bullet^P) \leq \varphi^2\lambda_j(\mathscr{A}_\bullet)$$

and vice-versa for $j = 1, \ldots, n$ where n is the total dimension of $\mathbf{A}_\bullet$.

PROOF. Using Lemma 6.4 we have

$$\mathcal{E}_{\mathscr{A}}(x) = \sum_{(e_1^i, e_2^i) \in P} \left[\left\| \mathscr{A}_{e_1^i \leftarrow v^i} x_{v^i} - \mathscr{A}_{u^i \rightarrow e_1^i} x_{u^i} \right\|_{\mathscr{A}_{e_1^i}}^2 + \left\| \mathscr{A}_{u^i \rightarrow e_2^i} x_{u^i} - \mathscr{A}_{e_2^i \leftarrow w^i} x_{w^i} \right\|_{\mathscr{A}_{e_2^i}}^2 \right]$$

$$+ \sum_{\text{unpaired } e \in Q_1} \left\| \mathscr{A}_{s(e) \rightarrow e} x_{s(e)} - \mathscr{A}_{e \leftarrow t(e)} x_{t(e)} \right\|_{\mathscr{A}_{s(e)}}^2$$

$$\leq \varphi^2 \sum_{(e_1^i, e_2^i) \in P} \left[\left\| \mathscr{A}_{u^i \rightarrow e_1^i} x_{u^i} - \mathscr{A}_{e_1^i \leftarrow v^i} x_{v^i} \right\|_{\mathscr{A}_{e_1^i}}^2 + \left\| \mathscr{A}_{e_3^i \leftarrow v^i}^P x_{v^i} - \mathscr{A}_{w^i \rightarrow e_3^i}^P x_{w^i} \right\|_{\mathscr{A}_{e_3^i}}^2 \right]$$

$$+ \sum_{\text{unpaired } e \in Q_1} \left\| \mathscr{A}_{s(e) \rightarrow e} x_{s(e)} - \mathscr{A}_{e \leftarrow t(e)} x_{t(e)} \right\|_{\mathscr{A}_{s(e)}}^2$$

$$\leq \varphi^2 \mathcal{E}_{\mathscr{A}^P}(x).$$

Then by Theorem B.2

$$\lambda_i(\mathscr{A}_\bullet) = \min_{\dim V = i} \left[\max_{x \in V \setminus 0} \frac{\mathcal{E}_{\mathscr{A}}(x)}{\langle x, x \rangle} \right] \leq \min_{\dim V = i} \left[\max_{x \in V \setminus 0} \frac{\varphi^2 \mathcal{E}_{\mathscr{A}^P}(x)}{\langle x, x \rangle} \right] = \varphi^2 \lambda_i(\mathscr{A}_\bullet^P).$$

proving the proposition. □

We note that Proposition 6.5 directly applies to graphs. Although we make no use of the following corollary, we hope that it will be of independent interest to spectral graph theorists.

COROLLARY 6.6. *Suppose $G = (V, E)$ is a multigraph with n vertices. Suppose we have an edge pairing $P = (e_1^1, e_2^1), \ldots (e_1^m, e_2^m)$ on G as in Definition 6.3. Let $\widehat{G}$ with the same nodes and edges as G except with each edge e_2^i replaced by an edge e_3^i between v^i and w^i. Then the eigenvalues of $L_{\widehat{G}}$ and L_G are related by*

$$\varphi^{-2} \lambda_j(L_G) \leq \lambda_j(L_{\widehat{G}}) \leq \varphi^2 \lambda_j(L_G)$$

and vice-versa for $j = 1, \ldots, n$.

PROPOSITION 6.7. *Suppose $\mathscr{A}_\bullet$ is a sheaf over a quiver Q with total dimension n. Given a subquiver Q', let $\mathscr{A}_\bullet'$ denote the restriction of $\mathscr{A}_\bullet$ to Q'.*

6.7(a) *If $e \in Q_1$ is an edge such that $s(e) = t(e)$ and $\mathscr{A}_{s(e)} = \mathscr{A}_{t(e)} = \mathrm{id}$ then $\lambda_i(\mathscr{A}_\bullet') = \lambda_i(\mathscr{A}_\bullet)$ for $i = 1, \ldots, n$ where $\mathscr{A}_\bullet'$ is $\mathscr{A}_\bullet$ restricted to the subquiver Q' of Q with the edge e removed.*

6.7(b) *Suppose we have an edge pairing $P = (e_1^1, e_2^1), \ldots (e_1^m, e_2^m)$ of Q such that for each i, the edges e_1^i, e_2^i have the same source and target vertices. Assume that for each i, there exists a linear map $M_i : \mathscr{A}_{e_2^i} \rightarrow \mathscr{A}_{e_1^i}$ satisfying both*

$$\mathscr{A}_{s(e_1^i) \rightarrow e_1^i} = M_i \mathscr{A}_{s(e_2^i) \rightarrow e_2^i} \qquad and \qquad \mathscr{A}_{e_1^i \leftarrow t(e_1^i)} = M_i \mathscr{A}_{e_2^i \leftarrow t(e_2^i)}$$

Let Q' be the subquiver of Q with all the edges $e_2^1, \ldots, e_2^n$ removed, and let $\mathscr{A}'_\bullet$ be the restriction of $\mathscr{A}_\bullet$ to Q'. Then, we have

$$(1 + \Gamma_P)^{-1} \lambda_i(\mathscr{A}_\bullet) \leq \lambda_i(\mathscr{A}'_\bullet) \leq \lambda_i(\mathscr{A}_\bullet)$$

for all $i = 1, \ldots, n$, where $\Gamma_P := \max_i \{\|M_i\|^2\}$.

PROOF. For 6.7(a), observe that $\mathcal{E}_\mathscr{A}(x) = \mathcal{E}_{\mathscr{A}'}(x)$ and thus the result follows from Theorem B.2. For 6.7(b), we have

$$\mathcal{E}_\mathscr{A}(x) \leq \sum_{(e_1^i, e_2^i) \in P} (1 + \|M_i\|^2) \left\| \mathscr{A}_{s(e_1^i) \to e_1^i} x_{s(e_1^i)} - \mathscr{A}_{e_1^i \leftarrow t(e_1^i)} x_{t(e_1^i)} \right\|^2$$

$$+ \sum_{\text{unpaired } e \in Q_1} \left\| \mathscr{A}_{s(e) \to e} x_{s(e)} - \mathscr{A}_{e \leftarrow t(e)} x_{t(e)} \right\|^2_{\mathscr{A}_{s(e)}}$$

$$\leq (1 + \Gamma_P) \mathcal{E}_{\mathscr{A}'}(x).$$

and the result follows from Theorem B.2.   □

In the following, at the risk of causing ambiguity with self-loops, we will ignore the direction of edges in and use the relation $v \sim e$ if either $s(e) = v$ or $t(e) = v$.

DEFINITION 6.8. Let Q be a quiver and let V be a subset of vertices such that there are no edges in Q between vertices of V. The **V-reduced quiver** Q^V is defined as follows. For each vertex $v \in V$,

(1) first remove v and any associated edges from Q;
(2) then, for each pair of edges (p, q) in Q such that $u \sim p \sim v \sim q \sim w$ for vertices $u, w \in Q_0 \setminus V$, add a new edge pq between u and w.

(We note in the second step above that if $u = w$, then pq forms a self-loop.) The resulting quiver is the desired Q^V.

FIGURE 5. An example of the reduction of a quiver Q with $V = \{v\}$.
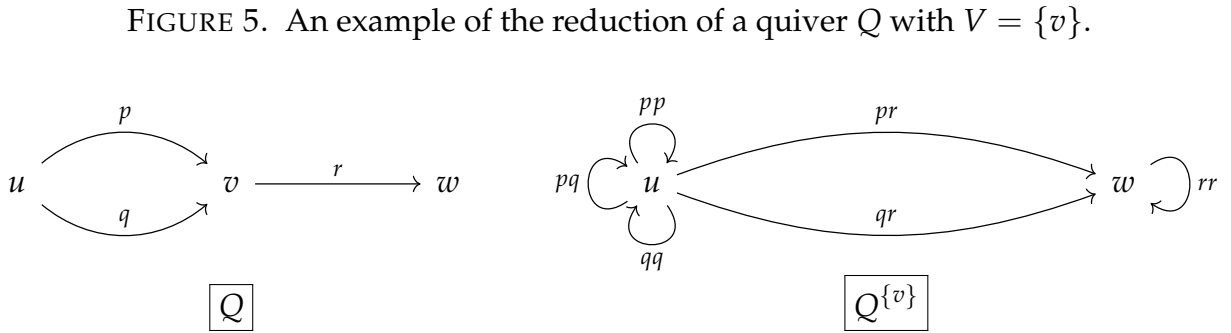
We now describe an analogue of the Kron reduction of a graph [7] for sheaves over a quiver. As discussed in [16], Kron reduction does not work in general for sheaves over a (loopless) graph. However, as the next proposition shows, if one restricts the type of

vertex removed, dependent on the sheaf, and moreover allows for loops in the underlying quiver, then Kron reduction is possible.

PROPOSITION 6.9. *Given a sheaf $\mathscr{A}_\bullet$ over a quiver Q, assume that $V \subset Q_0$ is a subset of vertices such that there are no edges in Q between vertices of V. Further suppose that for each $v \in V$ there exists some real number $\alpha_v > 0$ such that for each edge $e \sim v$, we have*

$$\text{either} \quad \mathscr{A}^*_{v\to e}\mathscr{A}_{v\to e} = \alpha_v\mathrm{id}_v \quad \text{or} \quad \mathscr{A}^*_{e\leftarrow v}\mathscr{A}_{e\leftarrow v} = \alpha_v\mathrm{id}_v$$

as appropriate. Then there exists a sheaf $\mathscr{A}^V_\bullet$ over Q^V such that $\Gamma(\mathscr{A}^V_\bullet) \cong \Gamma(\mathscr{A}_\bullet)$ and

$$\lambda_i(\mathscr{A}_\bullet) \le \lambda_i(\mathscr{A}^V_\bullet) \le \lambda_{i+r}(\mathscr{A}_\bullet)$$

for $i = 1, \ldots, n - r$ where n is the total dimension of $\mathscr{A}_\bullet$ and $r = \sum_{v \in V} \dim \mathscr{A}_v$.

PROOF. Consider the Laplacian $L_\mathscr{A}$. It can be written in a block form as

$$L_\mathscr{A} = \begin{pmatrix} X & Y^* \\ Y & D \end{pmatrix}\begin{matrix} Q_0 \setminus V \\ V \end{matrix} \quad \begin{matrix} Q_0 \setminus V & V \end{matrix}.$$

As there are no edges between vertices of V, D is a diagonal matrix where the (v, v) block is $d_v\alpha_v\mathrm{id}$ where d_v is the degree of v. Thus D is invertible, and we can form the Schur complement

$$L_\mathscr{A}/D := X - Y^*D^{-1}Y.$$

A vector $\begin{bmatrix} x & z \end{bmatrix}^T$ satisfies

$$\begin{bmatrix} X & Y^\dagger \\ Y & D \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

if and only if $Xx + Y^\dagger z = 0$ and $Yx + Dz = 0$. As D is invertible this can be rearranged to $Xx - YD^{-1}Y^\dagger x = 0$ hence $\ker(L_A/D) \cong \ker L_A$. For the interlacing, we can directly apply Theorem B.5.

We will now define the sheaf $\mathscr{A}^V_\bullet$. As in Definition 6.8, we will ignore orientation of the edges and write $u \sim e$ if either $s(e) = v$ or $t(e) = v$ and $\mathscr{A}_{v\sim e}$ for the corresponding assigned linear map. If $p \ne q$, define $\mathscr{A}^V_{u\sim pq} = (d_v\alpha_v)^{-1/2}\mathscr{A}^*_{v\sim p}\mathscr{A}_{u\sim p}$ and likewise for w. Otherwise, if $p = q$, define $\mathscr{A}^V_{u\to pp} = Z_p\mathscr{A}_{u\sim p}$ and $\mathscr{A}^V_{pp\leftarrow u} = \mathbf{0}$ for a linear map Z_p which we will now define. First observe that for every non-zero vector x, we have

$$\frac{1}{\alpha_v}x^*\mathscr{A}_{v\sim p}\mathscr{A}^*_{v\sim p}x \le \frac{1}{\alpha_v}\|\mathscr{A}^*_{v\sim p}\|^2\|x\|^2 = \frac{1}{\alpha_v}\|\mathscr{A}^*_{v\sim p}\mathscr{A}_{v\sim p}\|\|x\|^2 = \|x\|^2$$

thus there exists some Z_p such that

$$\mathrm{id} - \frac{1}{\alpha_v}\mathscr{A}_{v\sim p}\mathscr{A}^*_{v\sim p} = Z^*_pZ_p.$$

since the left-hand side is positive semi-definite.

Now we claim that $L_{\mathscr{A}^V} = L_{\mathscr{A}}/D$. To do so, consider vertices $u \in Q_0 \setminus V$ and $v \in V$. For each edge $u \sim p \sim v$, the contribution to the diagonal block indexed by (u, u) of $L_{\mathscr{A}}/D$ is

$$\mathscr{A}^*_{u \sim p} \mathscr{A}_{u \sim p} - \frac{1}{d_v \alpha_v} \sum_{u \sim q \sim v} \mathscr{A}^*_{u \sim p} \mathscr{A}_{v \sim p} \mathscr{A}^*_{v \sim q} \mathscr{A}_{u \sim q}.$$

We need to show this is equal to the contribution from the new edges of $\mathscr{A}^Q_\bullet$. For each vertex $u \in Q_0 \setminus V$, the contribution to the diagonal block is

$$\sum_{v \in V} \sum_{u \sim p \sim v} \left[\mathscr{A}^*_{u \sim p} \mathscr{A}_{u \sim p} - \frac{1}{d_v \alpha_v} \sum_{u \sim q \sim v} \mathscr{A}^*_{u \sim p} \mathscr{A}_{v \sim p} \mathscr{A}^*_{v \sim q} \mathscr{A}_{u \sim q} \right]$$

$$= \sum_{v \in V} \left[\sum_{u \sim p \sim v} \left[\mathscr{A}^*_{u \sim p} \mathscr{A}_{u \sim p} - \frac{d^u_v}{d_v \alpha_v} \mathscr{A}^*_{u \sim p} \mathscr{A}_{v \sim p} \mathscr{A}^*_{v \sim p} \mathscr{A}_{u \sim p} \right] \right.$$

$$\left. + \sum_{\substack{u \sim p \sim v \\ u \sim q \sim v \\ p \neq q}} \frac{1}{d_v \alpha_v} (\mathscr{A}^*_{v \sim p} \mathscr{A}_{u \sim p} - \mathscr{A}^*_{v \sim q} \mathscr{A}_{u \sim q})^* (\mathscr{A}^*_{v \sim p} \mathscr{A}_{u \sim p} - \mathscr{A}^*_{v \sim q} \mathscr{A}_{u \sim q}) \right]$$

where d^u_v is the number of edges between u and v. The term on the last line corresponds to the new self-loops on u given by v. By the defining property of Z_p, the term

$$\sum_{u \sim p \sim v} \left[\mathscr{A}^*_{u \sim p} \mathscr{A}_{u \sim p} - \frac{d^u_v}{d_v \alpha_v} \mathscr{A}^*_{u \sim p} \mathscr{A}_{v \sim p} \mathscr{A}^*_{v \sim p} \mathscr{A}_{u \sim p} \right]$$

equals

$$\sum_{u \sim p \sim v} \left[\frac{d_v - d^u_v}{d_v \alpha_v} \mathscr{A}^*_{u \sim p} \mathscr{A}_{v \sim p} \mathscr{A}^*_{v \sim p} \mathscr{A}_{u \sim p} + \mathscr{A}^*_{u \sim p} Z^*_p Z_p \mathscr{A}_{u \sim p} \right]$$

and thus corresponds with the contribution of the self-loop pp and the new edges to different vertices, as required. $\square$

7. Reduction of the Feature Selector Problem

Using the quiver operations in §6, we can reduce the number of vertices in the quiver whilst preserving the space of sections; this allows us to reduce the difficulty of computing approximate sections.

THEOREM 7.1. *Let **S** be a feature selector on X, $\mathscr{U}$ a cover of X such that $Q^{\mathscr{U}}$ is (weakly) connected, and $\mathscr{F} \subset X^*$. Fix a vertex $\sigma_0 \in Q^{\mathscr{F}}_0$. Then there exists a quiver Q' and a representation $\mathbf{A}'_\bullet$ of Q' satisfying the following properties:*

(1) *The quiver Q' has only one vertex and $|Q^{\mathscr{F}}_0|$ edges.*
(2) *The Laplacian of $\mathbf{A}'_\bullet$ is:*

$$\sum_{\tau \in Q^{\mathscr{U}}_0} \left[\mathrm{id}_{\sigma_0} - \pi^{\mathscr{F}}_{\sigma_0} \iota^{\mathscr{F}}_\tau \pi^{\mathscr{F}}_\tau \iota^{\mathscr{F}}_{\sigma_0} \right]$$

(3) There exist constants $C_1, C_2 > 0$ that depend on $\mathbf{S}, \mathscr{U}$ and $\mathscr{F}$ such that

$$C_1 \lambda_i(\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}) \leq \lambda_i(\mathbf{A}_\bullet') \leq C_2 \lambda_{i+k}(\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}})$$

for $i = 1, \ldots, n - k$.

Here n is the total dimension of $\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}$ and $k := n - \dim \mathbf{A}_{\sigma_0}^{\mathbf{S},\mathscr{F}}$.

PROOF. We view $\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}$ as a sheaf $\mathscr{A}_\bullet$ over a quiver Q, defined as follows:

$$Q_0 = \left\{ v_\sigma : \sigma \in Q_0^{\mathscr{U}} \right\} \cup \left\{ v_{\sigma\tau} : (\sigma \to \tau) \in Q_1^{\mathscr{U}} \right\},$$

$$Q_1 = \left\{ e_{\sigma\tau}, r_{\sigma\tau}^{DL}, r_{\sigma\tau}^{LD} : (\sigma \to \tau) \in Q_1^{\mathscr{U}} \right\},$$

$$s(e_{\sigma\tau}) = v_\sigma, \quad t(e_{\sigma\tau}) = v_\tau,$$

$$s(r_{\sigma\tau}^{DL}) = s(r_{\sigma\tau}^{LD}) = v_\sigma, \quad t(r_{\sigma\tau}^{DL}) = t(r_{\sigma\tau}^{LD}) = v_{\sigma\tau}.$$

This is precisely the merged quiver on which $\mathbf{N}_\bullet^{\mathbf{S},\mathscr{F}}$ is defined, except that we have dispensed with identity self-loops in light of Proposition 6.7(a). The desired sheaf $\mathscr{A}_\bullet$ assigns the following data to vertices and edges of Q for each $\sigma \in Q_0^{\mathscr{U}}$:

$$\mathscr{A}_{v_\sigma \to e_{\sigma\tau}} = \pi_\tau^{\mathscr{F}} \circ \iota_\sigma^{\mathscr{F}}, \qquad \mathscr{A}_{e_{\sigma\tau} \leftarrow v_\tau} = \mathrm{id}, \qquad \mathscr{A}_{v_\sigma \to r_{\sigma\tau}^{DL}} = \iota_\tau^{\mathscr{F}} \circ \pi_\tau^{\mathscr{F}} \circ \iota_\sigma^{\mathscr{F}},$$

$$\mathscr{A}_{r_{\sigma\tau}^{DL} \leftarrow v_\sigma} = \mathrm{id}, \qquad \mathscr{A}_{v_\sigma \to r_{\sigma\tau}^{LD}} = \iota_\sigma^{\mathscr{F}}, \qquad \mathscr{A}_{r_{\sigma\tau}^{DL} \leftarrow v_\sigma} = \mathrm{id}.$$

Use Proposition 6.7(b) to get a new copy $e_{\sigma\tau}^2$ of the edge $e_{\sigma\tau}$ with the maps multiplied by $\iota_\tau^{\mathscr{F}}$. Use this edge with Proposition 6.5 to move the edge $r_{\sigma\tau}^{DL}$ to an edge between v_τ and $v_{\sigma\tau}$, then using the edge $r_{\sigma\tau}^{LD}$ with Proposition 6.5 move this edge between v_τ and $v_{\sigma\tau}$ to an edge between v_σ and v_τ which has maps ι_σ and ι_τ respectively. Call this edge $p_{\sigma\tau}$. Using Proposition 6.7(b) can remove both $e_{\sigma\tau}$ and $e_{\sigma\tau}^2$ since $\pi_\tau^{\mathscr{F}}$ is the left inverse of $\iota_\tau^{\mathscr{F}}$. Finally, we can remove the edge $r_{\sigma\tau}^{LD}$ with Proposition 6.9 as $(\iota_\sigma^{\mathscr{F}})^* \iota_\sigma^{\mathscr{F}} = \mathrm{id}$. Now choose some vertex σ_0. As the nerve is connected, and each edge from each vertex has the same restriction maps to that vertex we can keep applying Proposition 6.5 until all edges have σ_0 as a boundary. There might be double edges, but these can be removed with Proposition 6.7(b). Then we apply Proposition 6.9 to reduce to just the vertex σ_0. The Laplacian is now

$$\sum_{\tau \in Q_0^{\mathscr{U}}} \left[\mathrm{id}_{\sigma_0} - \pi_{\sigma_0}^{\mathscr{F}} \iota_\tau^{\mathscr{F}} \pi_\tau^{\mathscr{F}} \iota_{\sigma_0}^{\mathscr{F}} \right]$$

and the eigenvalue relationship follows by tracking all the uses of Propositions 6.7(a), 6.7(b), 6.5, 6.9. $\qquad \square$

THEOREM 7.2. *Let $\mathbf{S}$ be a feature selector on a set X equipped with a cover $\mathscr{U}$. For each subset $\mathscr{F} \subset X^*$, here exists a sheaf $\mathscr{A}'$ over a quiver Q' with Laplacian*

$$(L_{\mathscr{A}'})_{\sigma,\tau} = \begin{cases} \sum_{\tau \subset \sigma} (2\mathrm{id}_\sigma - \pi_\sigma^{\mathscr{F}} \iota_\tau^{\mathscr{F}} \pi_\tau^{\mathscr{F}} \iota_\sigma^{\mathscr{F}}) + \sum_{\sigma \subset \gamma} \pi_\sigma^{\mathscr{F}} \iota_\gamma^{\mathscr{F}} \pi_\gamma^{\mathscr{F}} \iota_\sigma^{\mathscr{F}} & \text{if } \sigma = \tau, \\ -\pi_\sigma^{\mathscr{F}} \iota_\tau^{\mathscr{F}} & \text{if } \tau \subset \sigma \text{ or } \sigma \subset \tau, \\ 0 & \text{otherwise.} \end{cases}$$

such that there is an isomorphism $\Gamma(\mathscr{A}'_\bullet) \cong \Gamma(\mathbf{M}^{\mathbf{S},\mathscr{F}}_\bullet)$; and moreover, we have

$$\lambda_i(\mathbf{M}^{\mathbf{S},\mathscr{F}}_\bullet) \leq \lambda_i(\mathscr{A}'_\bullet) \leq \lambda_{i+k}(\mathbf{N}^{\mathbf{S},\mathscr{F}}_\bullet)$$

for all $i \in \{1, \dots, n-k\}$. Here $\mathbf{M}^{\mathbf{S},\mathscr{F}}_\bullet$ is the mixed representation of Definition 3.9, while n is its total dimension and $k := n - \sum_{\sigma \in Q_0} \dim \mathbf{A}^{\mathbf{S},\mathscr{F}}_\sigma$.

PROOF. Consider the representation $(\mathbf{A}^{\mathbf{S},\mathscr{F}}_\bullet)^* \sqcup_R \widehat{\mathbf{A}}^{\mathbf{S},\mathscr{F}}_\bullet$ as defined in Definition 3.9, and view it as a sheaf $\mathscr{A}_\bullet$ over a quiver Q. We can immediately remove identity self-loops using Proposition 6.7(a). The quiver Q is as follows:

$$Q_0 = \left\{ v_\sigma : \sigma \in Q^{\mathscr{U}}_0 \right\} \cup \left\{ v_{\sigma\tau} : (\sigma \to \tau) \in Q^{\mathscr{U}}_1 \right\},$$
$$Q_1 = \left\{ e_{\sigma\tau}, r^{DL}_{\sigma\tau}, r^{LD}_{\sigma\tau} : (\sigma \to \tau) \in Q^{\mathscr{U}}_1 \right\},$$
$$s(e_{\sigma\tau}) = v_\tau, \quad t(e_{\sigma\tau}) = v_\sigma,$$
$$s(r^{DL}_{\sigma\tau}) = s(r^{LD}_{\sigma\tau}) = v_\sigma, \quad t(r^{DL}_{\sigma\tau}) = t(r^{LD}_{\sigma\tau}) = v_{\sigma\tau}.$$

The sheaf $\mathscr{A}_\bullet$ consists of the following for each $\sigma \in Q^{\mathscr{U}}$:

$$\mathscr{A}_{v_\tau \to e_{\sigma\tau}} = \pi^{\mathscr{F}}_\sigma \circ \iota^{\mathscr{F}}_\tau, \qquad\qquad \mathscr{A}_{e_{\sigma\tau} \leftarrow v_\sigma} = \mathrm{id},$$
$$\mathscr{A}_{v_\sigma \to r^{DL}_{\sigma\tau}} = \iota^{\mathscr{F}}_\tau \circ \pi^{\mathscr{F}}_\tau \circ \iota^{\mathscr{F}}_\sigma, \qquad\qquad \mathscr{A}_{r^{DL}_{\sigma\tau} \leftarrow v_\sigma} = \mathrm{id},$$
$$\mathscr{A}_{v_\sigma \to r^{LD}_{\sigma\tau}} = \iota^{\mathscr{F}}_\sigma, \qquad\qquad \mathscr{A}_{r^{DL}_{\sigma\tau} \leftarrow v_\sigma} = \mathrm{id}.$$

We can use Proposition 6.9 to remove each $v_{\sigma\tau}$. In doing so, we add three new self-loops to v_σ; two identities and a self loop with maps $\iota^{\mathscr{F}}_\sigma$ and $\iota^{\mathscr{F}}_\tau \circ \pi^{\mathscr{F}}_\tau \circ \iota^{\mathscr{F}}_\sigma$. The two identity self loops can be removed with Proposition 6.7(a). □

7.1. Remarks on computation. In Theorems 7.1 and 7.2, we have assumed that L is Hermitian with respect to some orthonormal basis. Often, however, the bases we would prefer to use are not orthogonal with respect to our chosen inner product. We will see such an example in §8, where computing inner products is straightforward whilst finding an orthonormal basis is computationally infeasible. It is then advantageous then to compute eigenvectors of a Laplacian without such a change of basis. Suppose, for each $\sigma \in \mathscr{U}$, we have some basis $\mathscr{B}_\sigma$ of $\mathbf{S}^{\mathscr{F}}_{|\sigma|}$ and a Hermitian positive definite matrix $M^{\mathbf{S},\mathscr{F}}_\sigma$ representing the inner product in this basis. As $M^{\mathbf{S},\mathscr{F}}_\sigma$ is positive definite there exists an invertible W_σ such that $M^{\mathbf{S},\mathscr{F}}_\sigma = W^*_\sigma W_\sigma$. Then, in this basis, we now have for $\sigma, \tau \in \mathscr{U}$

$$\pi^{\mathscr{F}}_\sigma \iota^{\mathscr{F}}_\tau = W^{-*}_\sigma M^{\mathbf{S},\mathscr{F}}_{\sigma\tau} W^{-1}_\tau,$$
$$\pi^{\mathscr{F}}_\sigma \iota^{\mathscr{F}}_\tau \pi^{\mathscr{F}}_\tau \iota^{\mathscr{F}}_\sigma = W^{-*}_\sigma M^{\mathbf{S},\mathscr{F}}_{\sigma\tau} (M^{\mathbf{S},\mathscr{F}}_\tau)^{-1} M^{\mathbf{S},\mathscr{F}}_{\tau\sigma} W^{-1}_\sigma$$

where $M^{\mathbf{S},\mathscr{F}}_{\sigma\tau}$ is the matrix of inner products between $\mathscr{B}_\sigma$ and $\mathscr{B}_\tau$. Now computing eigenvectors of a Laplacian L in this basis can be realised as the generalised eigenvector problem

$$Lx = \lambda M^{\mathbf{S},\mathscr{F}}_{\mathscr{U}} x$$

where $M^{\mathbf{S},\mathscr{F}}_{\mathscr{U}}$ is the block diagonal matrix with blocks $M^{\mathbf{S},\mathscr{F}}_\sigma$ for $\sigma \in \mathscr{U}$.

As the matrices $M_\tau^{\mathbf{S},\mathscr{F}}$ are Hermitian and positive-definite, the products

$$M_{\sigma\tau}^{\mathbf{S},\mathscr{F}} (M_\tau^{\mathbf{S},\mathscr{F}})^{-1} M_{\tau\sigma}^{\mathbf{S},\mathscr{F}}$$

can be quickly computed using the Cholesky decomposition of $M_\tau^{\mathbf{S},\mathscr{F}}$. Furthermore, as L is Hermitian and usually sparse, there exist computationally efficient methods for calculating its smallest eigenvalues and their eigenvectors, e.g. variations on the Lanczos method [21], combined with the Cholesky decomposition (see e.g. [18]) of $L + \alpha \mathbf{I}$ for some small $\alpha > 0$.

8. Single-Cell Chromatin Accessibility

Chromatin is a macromolecule consisting of DNA and certain proteins, primarily histones [4]. Histones facilitate the dense compaction of DNA, which allows it to fit inside limited space within cells. The fundamental unit of chromatin is a *nucleosome*, which consists of approximately 147 *base-pairs* (namely, the usual A, C, G and T nucleotides) wound around histone molecules. The density of nucleosomes varies considerably across the genome; low density regions are more accessible to transcription factors and other binding proteins. Thus, understanding the relative accessibility of various parts of the genome is important in the study of gene regulation and its variation across different cell types.

Among the most promising experimental frameworks for measuring chromatin accessibility is **ATAC-seq** [1], which has been extended to single cell sequencing **scATAC-seq** (see e.g. [28]). The output of scATAC-seq may be viewed from a mathematical perspective as follows. Consider a set X of cells $\{x_1, \ldots, x_N\}$, with N being on the order of thousands. Each cell x_i has a genome consisting of M base pairs, where M is approximately three billion. The output of ATAC-seq on X is an $N \times M$ integer matrix. For each $1 \le i \le N$ and $1 \le j \le M$, the entry $\alpha_{i,j}$ is a number in $\{0, 1, 2\}$ representing the accessibility (for cell x_i at the j-th genomic position). Such data is high-dimensional, noisy and sparse, and hence warrants feature selection. This is accomplished via a class of techniques called **peak-calling** algorithms [36]. The goal of these methods is to identify specific locations along the genome, called *summits*, which are significantly accessible across the cells in X — see for instance the MACS2 algorithm [39].

REMARK 8.1. We are often interested in quantifying accessibility when one restricts to various sub-populations $Y \subset X$ of cell types. However, the peaks of various Y may be lost when calling peaks on X, particularly when the cardinalities of the X under consideration are unbalanced. The standard remedy is to first cluster the data and then call peaks on each cluster [28]. Subsequently, peaks from all the clusters are combined, with various rules beings applied for grouping and extending peaks which overlap. We show below how the quiver Laplacian provides a different mechanism for achieving similar results.

8.1. Pre-processing the data. We use the single-cell ATAC-seq data of [28], which consist of 28,274 tumour-infiltrating T cells. The pre-processing steps are the same as those of ibid:

(1) We align fragments to the GRCH37/HG19 assembly.
(2) We filter sample profiles to keep those with at least 1,000 fragments and a transcription start site (TSS) enrichment score of 8.
(3) We create a count matrix by tiling the genome with 2,500 base-pair windows and counting overlaps of endpoints of fragments with the windows.
(4) Next, we binarise this matrix and retain only the top 20,000 most accessible sites; call the resulting $28,724 \times 20,000$ boolean matrix B.
(5) We then create a frequency-inverse document frequency matrix, apply truncated singular value decomposition to obtain the 2nd to 25th singular vectors. This results in a reduced matrix E of size 28724×24.

8.2. Generating the quiver. We build an undirected graph G from the reduced matrix E as follows. Its vertices are the cells C_i, which correspond to rows of E. Treating each such row as a vector in $\mathbb{R}^{24}$, let $\delta_{i,j} \geq 0$ be the Euclidean distance between rows i and j for each $i < j$. We construct the 15 nearest neighbour graph generated by the $\delta_{i,j}$ distances. Our analysis now deviates from [28]. Using the Leiden algorithm [33] the vertex set V of G is partitioned into finitely many communities, $V = \coprod_i V_i$. As in ibid, communities of size < 200 are excluded. We then generate a cover $\mathscr{U}$ of X from the surviving communities as follows: each cover element $U_i \in \mathscr{U}$ has the form $V_i \cup W_i$, where $W_i \subset (V - V_i)$ consists of those vertices which share an edge with some vertex of V_i. The resulting quiver $Q^{\mathscr{U}}$ has 87 vertices, and is depicted in Figure 6A. Similarly, Figure 6B displays the contents of each vertex relative to the overall distribution, using the cell types identified in [28].

8.3. Building the representation. Having constructed $Q^{\mathscr{U}}$, we apply our feature selector **S** – which in this case is the MACS2 peak-calling algorithm [39] – to the fragment profiles of cells corresponding to each individual vertex. In the notation of Section 2, the input $\mathscr{F}$ is given by the columns of the binarised matrix B; explicitly, the map $f_j : X \to \mathbb{R}$ associated to the j-th column of E sends each x_i to the entry $B_{ij} \in \{0,1\}$. Calling peaks on this set of features produces a vector space $\mathbf{S}_{|\sigma|} \subset \mathbf{V}(\mathscr{F})$ for every vertex σ of $Q^{\mathscr{U}}$. We used the same parameters as in [28], and obtained several hundred thousand summits for each vertex as a result. For the purposes of this demonstration, we retained only the most significant 1,000 summits per vertex. Thus the total dimension of the selected vector spaces is 87,000. Since peak-calling on different sub-populations may change the true summit location on the genome slightly, we follow [38] and use a Gaussian kernel between summits (with bandwidth equal to 1,000 base-pairs) as the inner product on $\mathbf{V}(\mathscr{F})$. This allows us to build the mixed representation $\mathbf{M}_{\bullet}^{\mathbf{S},\mathscr{F}}$ (see Definition 3.9).

8.4. Results. We computed the first 3,000 eigenvectors of the Laplacian of $\mathbf{M}_{\bullet}^{\mathbf{S},\mathscr{F}}$. Since the total space admits a distinguished basis whose vectors correspond to summits (i.e., certain genomic positions $\{p_k\}$), we may associate to each eigenvector $v = \sum_k \alpha_k p_k$ the subset of supported genomic positions $P_v := \{k \mid \alpha_k \neq 0\}$[5]. Figure 6D shows that for

---

[5]Or more precisely, $|\alpha_k| > 10^{-5}$, the tolerance parameter for ARPACK-NG [22] in our experiment.

an overwhelming majority of eigenvectors v, the diameter of P_v for most v approximately equals 147 base-pairs, which is the average length of DNA around a nucleosome. A small fraction of the eigenspaces had dimension > 1; this resulted in a mixing of genomic positions, as shown by the blue outliers in Figure 6D. Nevertheless, in almost all cases, we observed that **the genomic supports of eigenvectors are tightly aligned across the vertices** of $Q^{\mathcal{U}}$.

We then performed a more detailed analysis of the sparsity pattern of the 3,000 eigenvectors. The outcome of this analysis has been summarised in Figure 6E, which plots the 3,000 eigenvectors against all 87 vertices of $Q^{\mathcal{U}}$ and highlights those vertices along which a given eigenvector has a non-zero component. We observe that the eigenvectors fall broadly into two classes. The first class, which is considerably larger, consists of those **eigenvectors whose genomic support contains summits which are shared by most, if not all vertices**: see, for instance, the components of eigenvector 11 in Figure 7. The second class contains **eigenvectors whose genomic support contains summits which are shared by a small fraction of the vertices**. A prime example of this type is eigenvector 2702, which has nonzero components along only 3 of the 87 vertices, and has also been depicted in Figure 7.

Finally, we note that (at least for these two examples) **the size of the support of the eigenvector relates to biological function**. Eigenvector 11, for instance, is concentrated in a regulatory region a few hundred base pairs downstream of the gene *FAM72D*. This gene is involved in the cell cycle [9], and is therefore relevant to most subpopulations of cells. On the other hand, eigenvector 2702 lies just downstream of the gene *STAT3*, which is a transcription factor known to be an important regulator for differentiation and proliferation in Th1, Th17, and Treg cells [27]. We also note that this eigenvector was only supported on the vertices 8, (8, 11), and 11 (Figure 7). These vertices are particularly enriched in Th1, Th17 and Treg cells (see Figure 6B). Indeed, Figure 6E reveals that many other eigenvectors are also uniquely supported on the vertices 8, (8, 11), 11 and (1, 11). Since vertex 11 is relatively disconnected from the other vertices, our analysis highlights it as an unusual sub-population of X.

Code and data availability

Code produced for the analysis will be available at

https://github.com/osumray/harmonic_feature_selection.

The data from [28] is available from the Gene Expression Omnibus with accession number GSE129785.

Appendix A. Removing Edges and Vertices

The following extends known results from graphs, e.g. see [34] and [35] respectively.

PROPOSITION A.1. *Let $\mathscr{A}_\bullet$ be a sheaf over a quiver Q, and let n be its total dimension.*

A.1(a) Let $R \subset Q_1$ be a subset of edges of Q. Let Q' be the quiver with edges $Q_1 \setminus R$ and $\mathscr{A}'$ the restriction of $\mathscr{A}$ to Q'. Then

$$\lambda_{k+i}(\mathscr{A}_\bullet) \leq \lambda_{k+i}(\mathscr{A}_\bullet') \leq \lambda_{k+r+i}(\mathscr{A}_\bullet)$$

for $i = 1, \ldots, n - k - r$ where $k = \dim \Gamma(\mathscr{A}_\bullet)$ and $r = \sum_{e \in R} \dim \mathscr{A}_e$.

A.1(b) Suppose $V \subset Q_0$ is a subset of vertices of Q. Let Q' be the quiver with vertices $Q_0 \setminus V$ and all edges $e \in Q_1$ such that either $s(e) \in V$ or $t(e) \in V$, and $\mathscr{A}_\bullet'$ the restriction of $\mathscr{A}_\bullet$ to Q'. Then

$$\lambda_i(\mathscr{A}_\bullet) - w_1 \leq \lambda_i(\mathscr{A}_\bullet') \leq \lambda_{k+i}(\mathscr{A}_\bullet) - w_2$$

for $i = 1, \ldots, n - k$ where $k = \sum_{u \in Q_0 \setminus V} \dim \mathscr{A}_u$ and

$$w_1 = \max_{v \in V} \left\{ \sum_{\substack{s(e)=v \\ t(e) \notin V}} \sigma_{\max}(\mathscr{A}_{v \to e})^2 + \sum_{\substack{t(e)=v \\ s(e) \notin V}} \sigma_{\max}(\mathscr{A}_{e \leftarrow v})^2 \right\},$$

$$w_2 = \min_{v \in V} \left\{ \sum_{\substack{s(e)=v \\ t(e) \notin V}} \sigma_{\min}(\mathscr{A}_{v \to e})^2 + \sum_{\substack{t(e)=v \\ s(e) \notin V}} \sigma_{\min}(\mathscr{A}_{e \leftarrow v})^2 \right\}.$$

PROOF OF PROPOSITION A.1(A). Define $L_{\mathscr{A},\mathrm{op}} = B_\mathscr{A} B_\mathscr{A}^*$. Let $k = \dim \ker L_\mathscr{A}$ and $c = \dim \ker L_{\mathscr{A},\mathrm{op}}$. Then for $i = 1, \ldots, n - k$, we have that

$$\lambda_{k+i}(L_\mathscr{A}) = \lambda_{c+i}(L_{\mathscr{A},\mathrm{op}}).$$

Now, removing the set of edges $R \subset Q_1$ corresponds to the $r \times r$ principal submatrix of $L_{\mathscr{A},\mathrm{op}}$, where $r := \sum_{e \in R} \dim \mathscr{A}_e$. By Theorem B.3 we have

$$\lambda_i(L_{\mathscr{A},\mathrm{op}}) \leq \lambda_i(L_{\mathscr{A}',\mathrm{op}}) \leq \lambda_{i+r}(L_{\mathscr{A},\mathrm{op}})$$

where $i \in \{1, \ldots, m - r\}$ thus

$$\lambda_{k+i}(\mathscr{A}_\bullet) \leq \lambda_{k+i}(\mathscr{A}_\bullet') \leq \lambda_{k+r+i}(\mathscr{A}_\bullet)$$

for $i \in \{1, \ldots, n - k - r\}$. □

PROOF OF PROPOSITION A.1(B). The boundary matrix $B_\mathscr{A}$ has block form

$$\begin{bmatrix} B_{11} & \mathbf{0} \\ B_{21} & B_{22} \end{bmatrix}$$

where $B_{11} = B_{\mathscr{A}'}$. Here, $B_\mathscr{A}$ has row blocks indexed by edges between vertices in V and column blocks indexed by vertices in V. Then Laplacian $L_\mathscr{A} = B_\mathscr{A}^* B_\mathscr{A}$ has block form

$$\begin{bmatrix} B_{11}^* B_{11} + B_{21}^* B_{21} & B_{21}^* B_{22} \\ B_{22}^* B_{21} & B_{22}^* B_{22} \end{bmatrix}.$$

Define $L_{Q^0 \setminus V} = B_{11}^* B_{11} + B_{21}^* B_{21}$ which is a principal submatrix of $L_\mathscr{A}$. Note that $B_{21}^* B_{21}$ is a diagonal matrix: The v, v block is

$$\sum_{\substack{s(e)=v \\ t(e) \neq V}} \mathscr{A}_{v \to e}^* \mathscr{A}_{v \to e} + \sum_{\substack{t(e)=v \\ s(e) \neq V}} \mathscr{A}_{e \leftarrow v}^* \mathscr{A}_{e \leftarrow v}.$$

We now have that

$$
\begin{aligned}
\lambda_i(\mathscr{A}'_\bullet) = \lambda_i(B_{11}^* B_{11}) &= \lambda_i(L_{Q_0 \setminus V} - B_{21}^* B_{21}) \leq \lambda_i(L_{Q_0 \setminus V}) + \lambda_{n-k}(-B_{21}^* B_{21}) \\
&= \lambda_i(L_{Q_0 \setminus V}) - \lambda_1(B_{21}^* B_{21}) \\
&= \lambda_i(L_{Q_0 \setminus V}) - \min_{v \in V} \lambda_1 \Big(\sum_{\substack{s(e)=v \\ t(e) \neq V}} \mathscr{A}_{v \to e}^* \mathscr{A}_{v \to e} + \sum_{\substack{t(e)=v \\ s(e) \neq V}} \mathscr{A}_{e \leftarrow v}^* \mathscr{A}_{e \leftarrow v} \Big) \\
&\leq \lambda_i(L_{Q_0 \setminus V}) - \min_{v \in V} \Big\{ \sum_{\substack{s(e)=v \\ t(e) \neq V}} \lambda_1(\mathscr{A}_{v \to e}^* \mathscr{A}_{v \to e}) + \sum_{\substack{t(e)=v \\ s(e) \neq V}} \mathscr{A}_{e \leftarrow v}^* \mathscr{A}_{e \leftarrow v} \Big\} \\
&= \lambda_i(L_{Q_0 \setminus V}) - \min_{v \in V} \Big\{ \sum_{\substack{s(e)=v \\ t(e) \neq V}} \sigma_{\min}(\mathscr{A}_{v \to e})^2 + \sum_{\substack{t(e)=v \\ s(e) \neq V}} \sigma_{\min}(\mathscr{A}_{e \leftarrow v})^2 \Big\} \\
&\leq \lambda_{i+k}(L_{\mathscr{A}}) - \min_{v \in V} \Big\{ \sum_{\substack{s(e)=v \\ t(e) \neq V}} \sigma_{\min}(\mathscr{A}_{v \to e})^2 + \sum_{\substack{t(e)=v \\ s(e) \neq V}} \sigma_{\min}(\mathscr{A}_{e \leftarrow v})^2 \Big\}.
\end{aligned}
$$

The first two inequalities are from Theorem B.4 and the final inequality is due to Theorem B.3. The right inequality is proven similarly. $\square$

Appendix B. Classical Results on Eigenvalues of Hermitian Matrices

In the following, for an $n \times n$ Hermitian matrix A, let $\lambda_1(A), \ldots, \lambda_n(A)$ be the eigenvalues ordered in increasing fashion.

THEOREM B.1 (Rayleigh, see [18] Theorem 4.2.2). *Let A be a Hermitian matrix. Then for integers $1 \leq i_1, \ldots, i_k \leq n$ let $x_{i_1}, \ldots, x_{i_k}$ be orthonormal, where x_{i_j} is an eigenvector corresponding to $\lambda_{i_j}(A)$. Define $S = \mathrm{span}\{x_{i_1}, \ldots, x_{i_k}\}$. Then for any $x \in S$*

$$
\lambda_{i_1}(A) \leq \frac{x^* A x}{\|x\|^2} \leq \lambda_{i_k}(A).
$$

THEOREM B.2 (Courant-Fischer, see [18] Theorem 4.2.6). *Let A be Hermitian matrix. Then*

$$
\lambda_k(A) = \min_{\dim V = k} \left[\max_{x \in V \setminus 0} \frac{\langle x, Ax \rangle}{\langle x, x \rangle} \right].
$$

THEOREM B.3 (Cauchy's interlacing theorem, see [18] Theorem 4.3.17). *Let A be a Hermitian matrix of size $n \times n$ and B an $m \times m$ principal submatrix of A. Then*

$$
\lambda_i(A) \leq \lambda_i(B) \leq \lambda_{i+n-m}(A)
$$

for $i = 1, \ldots, m$.

THEOREM B.4 (Weyl's inequalities, see [18] Theorem 4.3.1). *For A, B Hermitian matrices of size $n \times n$ the following inequalities hold:*

$$
\lambda_i(A + B) \leq \lambda_{i+j}(A) + \lambda_{n-j}(B)
$$

s for $j = 0, \ldots, n-1$ and

$$\lambda_i(A + B) \geq \lambda_{i-j+1}(A) + \lambda_j(B)$$

for $j = 1, \ldots, i$.

THEOREM B.5 (See [**31**] Theorem 5). *For H a Hermitian matrix of dimension n given in block form as*

$$H = \begin{bmatrix} A & B^* \\ B & D \end{bmatrix}$$

where D is non-singular let $H/D = A - BD^{-1}B^*$ be the Schur complement of D. Then the eigenvalues of H and H/D are interlaced:

$$\lambda_i(H) \leq \lambda_i(H/D) \leq \lambda_{i+r}(H)$$

for $i = 1, \ldots, n-r$ where r is the rank of D.

References

[1] J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods*, 10(12):1213–1218, 2013.

[2] F. R. Chung. The Laplacian of a hypergraph. *Expanding graphs*, 10:21–36, 1992.

[3] F. R. Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[4] G. Cooper and K. Adams. *The cell: a molecular approach*. Oxford University Press, 2022.

[5] J. Curry, R. Ghrist, and V. Nanda. Discrete Morse theory for computing cellular sheaf cohomology. *Foundations of Computational Mathematics*, 16:875–897, 2016.

[6] J. M. Curry. *Sheaves, cosheaves and applications*. University of Pennsylvania, 2014.

[7] F. Dorfler and F. Bullo. Synchronization of power networks: Network reduction and effective resistance. *IFAC Proceedings Volumes*, 43(19):197–202, 2010.

[8] S. Eilenberg and N. Steenrod. *Foundations of algebraic topology*, volume 2193. Princeton University Press, 2015.

[9] M. Fischer, P. Grossmann, M. Padi, and J. A. DeCaprio. Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic acids research*, 44(13):6070–6086, 2016.

[10] T. Gao, J. Brodzki, and S. Mukherjee. The geometry of synchronization problems and learning group actions. *Discrete & Computational Geometry*, 65(1):150–211, 2021.

[11] R. Ghrist and H. Riess. Cellular sheaves of lattices and the Tarski Laplacian. *arXiv preprint arXiv:2007.04099*, 2020.

[12] T. E. Goldberg. *Combinatorial Laplacians of Simplicial Complexes*. Senior thesis, Bard College, 2002.

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 3rd ed edition, 1996.

[14] J. Gu, B. Hua, and S. Liu. Spectral distances on graphs. *Discrete Applied Mathematics*, 190:56–74, 2015.

[15] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46:389–422, 2002.

[16] J. Hansen and R. Ghrist. Toward a Spectral Theory of Cellular Sheaves. *Journal of Applied and Computational Topology*, 3(4), 2019.

[17] A. N. Hirani. *Discrete exterior calculus*. California Institute of Technology, 2003.

[18] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, second edition, corrected reprint edition, 2017.

[19] I. T. Jolliffe. *Principal Components Analysis, 2nd Ed.* Springer, 2002.

[20] C. A. Joslyn, L. Charles, C. DePerno, N. Gould, K. Nowak, B. Praggastis, E. Purvine, M. Robinson, J. Strules, and P. Whitney. A sheaf theoretical approach to uncertainty quantification of heterogeneous geolocation information. *Sensors*, 20(12):3418, 2020.

[21] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *United States Governm. Press Office Los Angeles, CA*, 1950.

[22] R. Lehoucq, K. Maschhoff, D. Sorensen, C. Yang, F. Houssen, S. Ledru, et al. ARPACK-NG: Large scale eigenvalue problem solver. *Astrophysics Source Code Library*, pages ascl–2306, 2023.

[23] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[24] T. Muller, A. Seigal, and V. Nanda. Multilinear hyperquiver representations. *arXiv:2305.05622v2 [math.AG]*, 2023.

[25] A. Parada-Mayorga, H. Riess, A. Ribeiro, and R. Ghrist. Quiver signal processing (QSP). *arXiv preprint arXiv:2010.11525*, 2020.

[26] K. Pearson F.R.S. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[27] C. Rébé and F. Ghiringhelli. STAT3, a master regulator of anti-tumor immune response. *Cancers*, 11(9):1280, 2019.

[28] A. T. Satpathy, J. M. Granja, K. E. Yost, Y. Qi, F. Meschi, G. P. McDermott, B. N. Olsen, M. R. Mumbach, S. E. Pierce, M. R. Corces, P. Shah, J. C. Bell, D. Jhutty, C. M. Nemec, J. Wang, L. Wang, Y. Yin, P. G. Giresi, A. L. S. Chang, G. X. Y. Zheng, W. J. Greenleaf, and H. Y. Chang. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nature Biotechnology*, 37(8):925–936, 2019.

[29] R. Schiffler. *Quiver Representations*. Number 184 in CMS Books in Mathematics. Springer, 2014.

[30] A. Seigal, H. A. Harrington, and V. Nanda. Principal components along quiver representations. *Foundations of Computational Mathematics*, 23(4):1129–1165, 2023.

[31] R. L. Smith. Some interlacing properties of the Schur complement of a Hermitian matrix. *Linear Algebra and its Applications*, 177:137–144, 1992.

[32] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.

[33] V. A. Traag, L. Waltman, and N. J. Van Eck. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):5233, 2019.

[34] J. Van Den Heuvel. Hamilton cycles and eigenvalues of graphs. *Linear algebra and its applications*, 226:723–730, 1995.

[35] B. Wu, J. Shao, and X. Yuan. Deleting vertices and interlacing Laplacian eigenvalues. *Chinese Annals of Mathematics, Series B*, 31(2):231–236, 2010.

[36] F. Yan, D. R. Powell, D. J. Curtis, and N. C. Wong. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome biology*, 21:1–16, 2020.

[37] K. Ye and L.-H. Lim. Schubert varieties and distances between subspaces of different dimensions. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1176–1197, 2016.

[38] J. Yu, D. Sun, Z. Hou, and L.-Y. Wu. Single-cell ATAC-seq analysis via network refinement with peaks location information. *bioRxiv*, page 2022.11.18.517159, 2022.

[39] Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, et al. Model-based analysis of ChIP-Seq (MACS). *Genome biology*, 9(9):1–9, 2008.

FIGURE 6. Overview of Laplacian from 7.2 applied to tumour-infiltrating T cells. (A) Quiver constructed from cover, with vertex size corresponding to the cardinality of the covering set. (B) Log likelihood of each cell type label within each vertex. (C) First 3,000 eigenvalues. (D) Range of genomic positions on each chromosome corresponding to non-zero components of the first 3,000 eigenvectors. Blue points correspond to positions across multiple chromosomes. (E) Binary matrix corresponding to the existence of a non-zero component in each vertex for the first 3,000 eigenvectors.
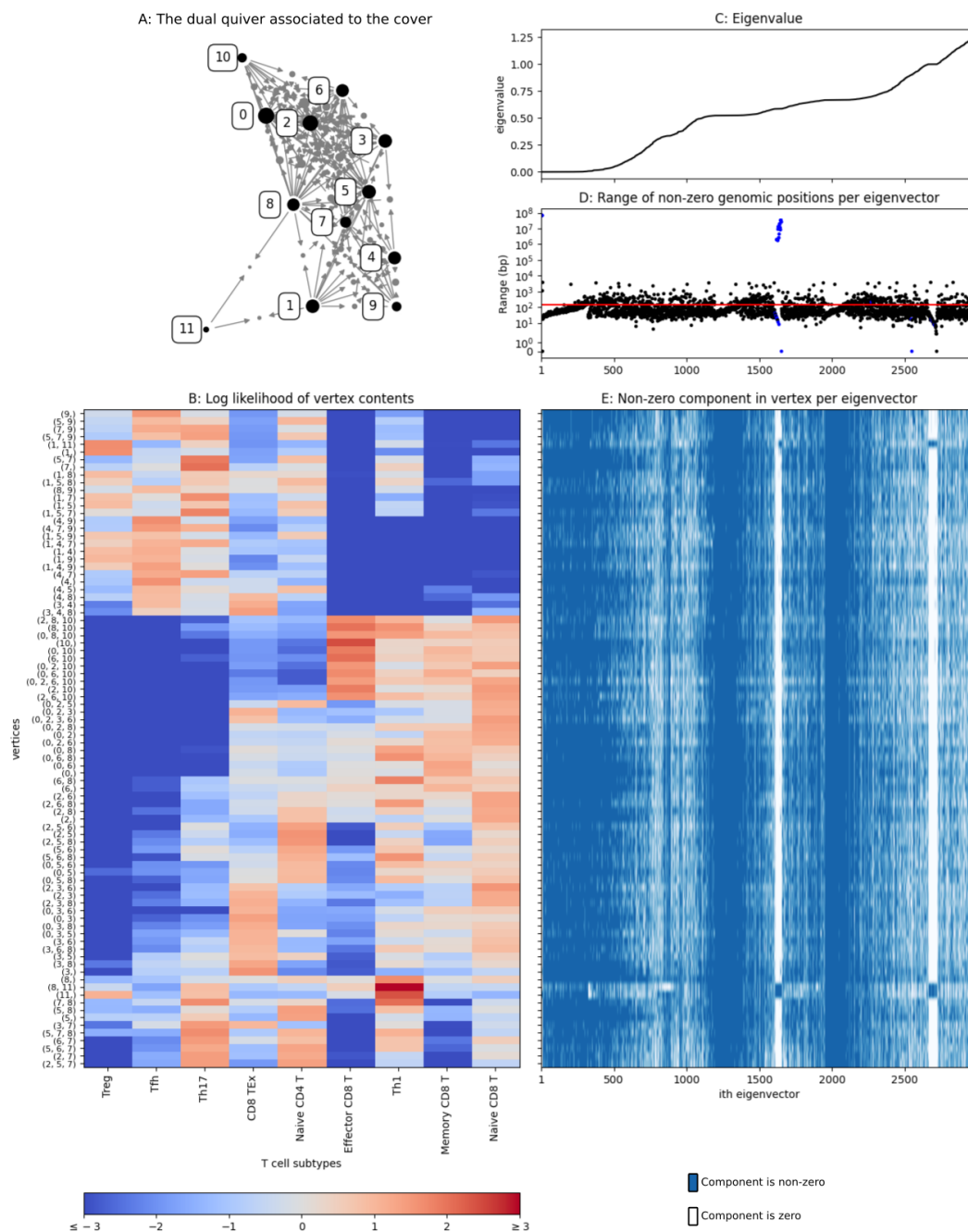
FIGURE 7. Examples of eigenvectors. Eigenvector #11 has a non-zero component at each vertex, whilst eigenvector #2702 has only non-zero components in vertices (8,), (11,) and (8, 11).