Machinery for Proving Sum of Squares Lower Bounds on Planted Problems

Aaron Potechin

February 5, 2020

Abstract

In this manuscript, we generalize the machinery used to analyze the moment matrix for planted clique.

Note: This machinery was originally supposed to be part of the paper "The Power of Sum of Squares on Hidden Substructures". However, appropriately generalizing the planted clique machinery and ironing out the technical details turned out to be quite involved, so this generalization is in this separate manuscript.

Acknowledgement: Section 1 of this manuscript was written by Sam Hopkins, Pravesh Kothari, and Aaron Potechin. Thanks to Sam Hopkins and Goutham Rajendran for helpful comments on this manuscript .

1 Informal Statement of the Explicit Coefficient Lower Bounds

1.1 Introduction

In this section and those which follow, we study symmetric matrix-valued functions on subset-indexed product spaces which are *almost positive semidefinite*. To warm up, here is a special case of the setting we study. Let $n, d, k \in \mathbb{N}$. We consider functions $\Lambda : \{\pm 1\}^{\binom{n}{k}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$ which map the product space $\{\pm 1\}^{\binom{n}{k}}$ to real symmetric matrices.

Notice that $\{\pm 1\}^{\binom{n}{k}}$ is indexed by *subsets* of $[n]$, and $\mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$ is indexed by pairs of subsets of $[n]$. The permutation group S_n has a natural action on both $\{\pm 1\}^{\binom{n}{k}}$ and $\mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$, and hence also it acts on Λ: suppose that Λ is invariant under this action. That is, for every $\pi \in S_n$ it holds that $\pi \cdot \Lambda = \Lambda$. For instance, if $k = 2$ and $d = 1$, the function $\Lambda(x) = x + \mathrm{Id}$ is such a function, where Id is the $d \times d$ identity matrix.

Symmetric functions Λ of this type arise as the main players in the construction of SoS lower bounds for planted problems, where the goal is often to show that Λ is positive semi-definite with high probability:

$$\Lambda(x) \succeq 0 \text{ with probability at least } 1 - o(1) \text{ when } x \sim \{\pm 1\}^{\binom{n}{k}} \text{ uniformly.} \tag{1}$$

For instance, if $k = 2$, then $x \sim \{\pm 1\}^{\binom{n}{2}}$ chosen uniformly at random is in fact a sample from $G(n, 1/2)$; this recovers settings in which one might study optimization problems over random graphs, such as the planted clique problem. In the example case $k = 2, d = 1$, if $\Lambda(x) = x + 10\sqrt{n} \cdot \mathrm{Id}$ then (1) holds for Λ, since with high probability the $n \times n$ random matrix x has spectral norm at most $2\sqrt{n}$. When $k > 2$, the space $\{\pm 1\}^{\binom{n}{k}}$ is the set of k-tensors with ± 1 entries, which puts us in the setting of our SoS lower bound for tensor principal component analysis.

Our main result in this section is a sufficient condition on Λ for (1) to hold. Together with the pseudocalibration method (which supplies the function Λ we employ to prove SoS lower bounds for planted problems), this condition suffices to prove our exponential lower bounds for componenent analysis problems.

Setup overview The sufficient condition we describe is Fourier-theoretic: that is, it concerns the pattern of decay in the coefficients of a Fourier expansion of Λ. Stating this condition, and hence our main theorem, requires some setup, as we have to establish enough definitions to do Fourier analysis on S_n-symmetric matrix-valued subset-indexed functions Λ. Our first task is to describe an explicit orthonormal basis for the set of functions $\{\Lambda : \{\pm 1\}^{\binom{n}{k}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}} \; S_n\text{-symmetric}\}$. We establish a description of these basis functions in terms of combinatorial objects we call *ribbons* and *shapes* – this will ultimately allow us to turn Fourier-theoretic statements into combinatorial ones, a connection which will be quite useful in our proofs.

The basis functions we construct are of course themselves maps $M : \{\pm 1\}^{\binom{n}{k}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$. The basis functions have a factorization structure: that is, some basis functions M_α can be (approximately) factored into a product of two (or more) other basis functions: $M_\alpha(x) \approx M_\beta(x) M_\gamma(x)$, where we emphasize that $M_\beta(x) M_\gamma(x)$ are multiplied *as matrices in* $\mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$. To state our main theorem we have to describe some basics of this factorization structure.

Organization We will give definitions and theorem statements in two rounds. As warm-ups, in the remainder of this section we offer definitions and a theorem statement which are simplified in two ways with respect to our main theorem: they concern just the case that the underlying product space is $\{\pm 1\}^{\binom{n}{2}}$, and they leave out some of the quantitative bounds necessary to apply the main theorem and prove SoS lower bounds. In the following section we make more general definitions and state our main theorem. The subsequent sections prove the main theorem.

On generality Since we will not be able to make a full statement of our main theorem until after some setup, we describe a couple features of the general setting which we have avoided in the special case considered so far. The challenges presented by these generalizations are primarily technical and definitional; the general setting is required, however, to obtain a theorem which can be used for SoS lower bounds outside boolean settings and for less symmetric problems.

First of all, we do not require that the underlying product space is $\{\pm 1\}^{\binom{n}{k}}$. Instead, under mild conditions on a set Ω and a measure μ on it, our theorem applies to Λ defined on $\Omega^{\binom{n}{k}}$. For instance, $\Omega = \mathbb{R}$ and $\mu = \mathcal{N}(0,1)$ is another common setting. Second, if $n_1, \ldots, n_k \in \mathbb{N}$, our theorem applies also if the underlying product space is not indexed by subsets of $[n]$ but rather by tuples in $[n_1] \times [n_2] \times \ldots [n_k]$. In the example case $k = 2$, this allows $\Omega^{n_1 \times n_2}$ to be the set of rectangular matrices with entries in Ω. Then the required symmetry group of Λ is $S_{n_1} \times \ldots \times S_{n_k}$ rather than S_n.

Finally, we note that as far as we are aware simpler approaches to Fourier analysis of $\Lambda : \{\pm 1\}^{\binom{n}{k}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}$ than we follow here do not suffice to prove results like our main theorem. It is tempting, for instance, to do more traditional Fourier analysis of scalar-valued functions on each entry of Λ. Unfortunately, this approach seems not to adequately take advantage of the matrix structure of Λ, and hence we suspect it cannot be used to prove spectral statements like $\Lambda(x) \succeq 0$.

1.2 Fourier analysis for matrix-valued functions: ribbons and shapes

In this section we construct an explicit orthonormal basis for $\{\Lambda : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}\ S_n\text{-symmetric}\}$.

1.2.1 Ribbons

The construction starts by lifting the usual Fourier basis for $\{f : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}\}$ to the matrix-valued functions: we call the resulting object a *ribbon*.

Becase $\{\pm 1\}^{\binom{n}{2}}$ is (isomorphic to) the set of graphs on n vertices, we denote its elements by G. Also, since $\binom{n}{d}$ is (isomorphic to) the set of multilinear monomials of degree d in variables $x_1, \ldots, x_n$, we interchangebly consider Λ to be indexed by $A \subseteq [n]$ or by monomials $x^A = \prod_{i \in A} x_i$.

We start by recalling the usual Fourier characters for functions on the hypercube.

Definition 1.1. *For $n \in \mathbb{N}$ and $E \subseteq \binom{n}{2}$, let $\chi_E : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}$ be given by $\chi_E(G) = \prod_{(i,j) \in E} G_{ij}$.*

The functions $\{\chi_E\}_{E \subseteq \binom{n}{2}}$ form an orthonormal basis for $\{f : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}\}$ with respect to the uniform distribution on $\{\pm 1\}^{\binom{n}{2}}$ – see e.g. [**?**].

Definition 1.2 (Simplified – see Definition **??**). *Let $n \in \mathbb{N}$. A ribbon R is a tuple (E, A, B) where $E \subseteq \binom{n}{2}$ and $A, B \subseteq [n]$. To distinguish the components of one ribbon from those of another, we often write $E = E(R)$ and $A = A_R, B = B_R$. R thus specifies:*

1. *A Fourier character $\chi_{E(R)}$.*
2. *Row and column indices $x_{A_R} := \prod_{i \in A_R} x_i$ and $x_{B_R} := \prod_{i \in B_R} x_i$.*

We think of R as a graph with vertices

$$V(R) = \{ \text{ endpoints of } (i,j) \in E(R) \} \cup A_R \cup B_R$$

and edges $E(R)$, where A_R, B_R are distinguished sets of vertices.

Definition 1.3 (Matrix-valued function for a ribbon R). *Given a ribbon R, we define the function $M_R : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{|A_R|} \times \binom{n}{|B_R|}}$ to have entries $M(A, B) = \chi_{E(R)}$ and $M(A', B') = 0$ whenever $A' \neq A$ or $B' \neq B$ Notice that if $|A_R|$ or $|B_R$ is not equal to d, the function M_R assumes values in matrices of dimensions other than $\binom{n}{d} \times \binom{n}{d}$. We will need to work with matrices of dimensions $\binom{n}{d'} \times \binom{n}{d''}$ because they will appear in factorizations of $\binom{n}{d} \times \binom{n}{d}$ matrices.*

The following proposition captures the main property of the matrix-valued functions M_R – they are an orthonormal basis (though not our final one). We leave the proof to the reader.

Proposition 1.4. *The matrix-valued functions $\{M_R : R \text{ is a ribbon with } |A_R| = |B_R| = d\}$ form an orthonormal basis for the vector space $\{M : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}\}$ of matrix valued functions with respect to the inner product*

$$\langle M, M' \rangle = \mathop{\mathbb{E}}_{G \sim \{\pm 1\}^{\binom{n}{2}}} [\operatorname{Tr} M(G)(M'(G))^\top].$$

1.2.2 Shapes and Graph Matrices

We have defined an orthonormal basis for $\{M : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{d} \times \binom{n}{d}}\}$. Our aim is an explicit orthonormal basis for the subset of those functions which are symmetric with respect to the action of S_n. We will obtain such a basis by projecting the M_R functions to the span of the symmetric functions; this just requires averaging each M_R over its orbit under S_n (then checking that the projected functions are still orthonormal, once we remove duplicates). As with ribbons, we will make use of a combinatorial description of the resulting matrix-valued functions; we call the combinatorial objects we use *shapes*.

Definition 1.5 (Simplified – see Definition **??**). *Informally, a shape α is just a ribbon R where the vertices are specified by variables rather than having specific values in $[1, n]$. More precisely, a shape $\alpha = (V(\alpha), E(\alpha), U, V)$ is a graph on vertices $V(\alpha)$, with*

1. *edges $E(\alpha) \subseteq \binom{V(\alpha)}{2}$*
2. *Distinguished tuples of vertices $U = U_\alpha = (u_1, u_2, \dots)$ and $V = V_\alpha = (v_1, v_2, \dots)$, where $u_i, v_i \in V(\alpha)$.*

(Note that $V(\alpha)$ and V_α are not the same object!)

Now we can define our basis functions, called *graph matrices*, for the symmetric matrix-valued functions, by projecting the functions M_R.

Definition 1.6 (Graph matrices)**.** *Let α be a shape. The graph matrix $M_\alpha : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{|U_\alpha|}\times\binom{n}{|V_\alpha|}}$ is defined to be the matrix-valued function with A, B-th entry*

$$M_\alpha(A,B) = \sum_{\substack{R \text{ s.t. } A_R=A, B_R=B \\ \exists \phi: V(\alpha)\to[1,n]: \\ \phi \text{ is injective}, \phi(\alpha)=R}} \chi_{E(R)}$$

That is, $M_\alpha = \sum_R M_R$ where the sum is over ribbons which can be obtained by assigning each vertex in $V(\alpha)$ a label from $[1,n]$. Up to scaling, M_α is the result of averaging each M_R over the orbit of R.

Example 1.7. *Suppose $d = 1$, so that $M_\alpha : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{n\times n}$. Let α be the following shape. $V(\alpha) = \{u_1, v_1\}$ is a 2-element set. $E(\alpha) = \{\{u_1, v_1\}\}$ contains a single edge between vertices u_1, v_1. The tuples U_α, V_α are $(u_1), (v_1)$, respectively. Then M_α has entries $(M_\alpha)_{i,j}(G) = G_{ij}$ if $i \neq j$ and $(M_\alpha)_{i,i} = 0$. If $G \in \{\pm 1\}^{\binom{n}{2}}$ is thought of as a graph, then M_α is its ± 1 adjacency matrix with zeros on the diagonal.*

Example 1.8. *If $V(\alpha) = \{u\}$ is a singleton, $E(\alpha) = \emptyset$, and $U_\alpha = V_\alpha = (u)$, then $M_\alpha(G)$ is a constant function with value always equal to the $n \times n$ identity matrix.*

1.3 Factoring graph matrices and decomposing shapes

A crucial idea in our application of Fourier analysis to study matrix-valued function Λ are approximate factorizations of graph matrices M_α into products of other graph matrices.

Informal example: lines of lengths 1 and 2 To get a sense of the kind of factorizations we are interested in, consider the following informal example. Let α_1 be the shape (consisting of the one-edge graph) described in Example 1.7, where $M_{\alpha_1}(G)$ for $G \in \{\pm 1\}^{\binom{n}{2}}$ is just the ± 1 adjacency matrix of G.

Let α_2 be a length-two path. More formally, let $V(\alpha_2) = \{u, v, w\}$ be a set of 3 vertices. Let $E(\alpha_2) = \{\{u, w\}, \{w, v\}\}$ be an (undirected) path $u \to w \to v$. Let $U_{\alpha_2} = (u)$ and $V_{\alpha_2} = (v)$ be length-1 tuples. Then for $i \neq j$, the matrix-valued function M_{α_2} has entries $(M_{\alpha_2})_{ij}(G) = \sum_{k\neq i,j} G_{ik}G_{jk}$. That is, the ij-th entry roughly corresponds to a centered count of the number of length-2 paths from i to j in G, considered as a graph.

Observe that $M_{\alpha_1}^2$, the square of the function M_{α_1}, *also has entries which correspond to counts of such paths*. That is, $M_{\alpha_2} \approx M_{\alpha_1}^2$. Unfortunately, $(M_{\alpha_1}^2)_{ii} \neq (M_{\alpha_2})_{ii}$ – that is, equality does not hold on the diagonal – so factorization of M_{α_2} holds only approximately.

The relationship between α_1 and α_2 that this factorization relies on is that

1. The vertex w in α_2 is a vertex separator in the graph $E(\alpha_2)$ between the sets of vertices $U_{\alpha_2} = \{u\}$ and $V_{\alpha_2} = \{v\}$, and

2. If the graph $E(\alpha_2)$ is "split" along this vertex separator w, the result is two copies of the one-edge graph $E(\alpha_1)$.

For any shape α we can hope for similar factorizations $M_\alpha \approx M_\beta M_\gamma$, if β, γ are the result of "splitting" the graph $E(\alpha)$ along some set of vertices which separate U_α and V_α. Any such vertex separator corresponds to a potential factorization of M_α, but some will be more useful to us than others. In what follows, we identify for each α some canonical vertex separators whose corresponding factorizations will be useful in our arguments.

1.3.1 Leftmost and Rightmost Minimum Vertex Separators and Decomposition of Shapes into Left, Middle, and Right Parts

For each shape α we will identify three other shapes, which (for reasons we will see soon) we denote by σ, τ, σ'^T and call *left, middle, and right parts of* α, respectively. The idea is that $M_\alpha \approx M_\sigma M_\tau M_{\sigma'^T} \approx M_\sigma M_\tau M_{\sigma'}^T$. (We have not yet described what the transpose operation in σ'^T means for shapes.) Following the strategy outlined above, σ, τ, and $(\sigma')^T$ arise by splitting the shape α along some well-chosen vertex separators.

Definition 1.9 (Vertex Separators)**.** *We say that a set of vertices S is a vertex separator of α if every path from U_α to V_α in the graph $E(\alpha)$ (including paths of length 0) intersects S.*

We note one basic property of vertex separators; the proof is straightforward.

Proposition 1.10. *For any vertex separator S of α, $U_\alpha \cap V_\alpha \subseteq S$*

To obtain σ, τ, σ' from α we rely on *leftmost and rightmost minimum-weight* vertex separators:

Definition 1.11 (Minimum Vertex Separators)**.** *We say that S is a mimumum vertex separator of α if S is a vertex separator of α and for any other vertex separator S' of α, $|S| \leq |S'|$.*

Definition 1.12 (Leftmost and Rightmost Minimum Vertex Separators)**.**

1. *We say that S is the leftmost minum vertex separator of α if S is a minimum vertex separator of α and for every other minimum vertex separator S' of α, every path from U_α to S' inersects S.*

2. *We say that T is the rightmost minum vertex separator of α if T is a minimum vertex separator of α and for every other minimum vertex separator S' of α, every path from S' to V_α inersects T.*

It requires some straightforward combinatorial arguments to show that leftmost and rightmost minimum vertex separators exist and are unique.

Proposition 1.13 (See (*** to be added))**.** *For every shape α there is a unique leftmost vertex separator and a unique rightmost vertex separator.*

We need one more notational convenience before we define formally the operation of splitting α along its minimum leftmost and rightmost vertex separators.

Definition 1.14 (Shape transposes)**.** *Given a shape α, we define α^T to be the shape α with U_α and V_α swapped i.e. $U_{\sigma^T} = V_\sigma$ and $V_{\sigma^T} = U_\sigma$. Note that $M_{\alpha^T} = M_\alpha^T$, where M_α^T is the usual transpose of the matrix-valued function M_α.*

We are finally ready to describe the left, middle, and right part decomposition of a shape α.

Definition 1.15 (Decomposition Into Left, Middle, and Right Parts)**.** *Given a shape α, letting S and T be the leftmost and rightmost minimum vertex separators of α, we decompose α into left, middle, and right parts. We will have $V(\sigma), V(\tau), V(\sigma') \subseteq V(\alpha)$ and $E(\sigma), E(\tau), E(\sigma') \subseteq E(\alpha)$, but the distinguished vertices U_σ, V_σ and so on will differ from U_α, V_α.*

1. *The left part σ of α is the part of α reachable from U_α without intersecting S more than once. It includes S but excludes all edges which are entirely within S). More formally, let*

   $E(\sigma) = \{\{u, v\} \in E(\alpha) \,:\, \{u, v\}$ *lies on a path from U to S in $E(\alpha)$ intersecting S in at most one of* $u, v\}$

   We let $U_\sigma = U_\alpha$ and $V_\sigma = S$. Finally $V(\sigma) \subseteq V(\alpha)$ contains all the endpoints of $E(\sigma)$ together with U_σ, V_σ.

2. *The right part ${\sigma'}^T$ of α is like the left part with the roles of S and T interchanged.*

3. *The middle part τ of α is, informally, the part of α between S and T (including S and T and all edges which are entirely within S or within T). More formally, let $U_\tau = S, V_\tau = V$, and $E(\tau) = E(\alpha) \setminus (E(\sigma) \cup E(\sigma'))$ be all those edges of $E(\alpha)$ which do not appear in $E(\sigma)$ or $E(\sigma')$. Then $V(\tau)$ is all those vertices indicent to edges in $E(\tau)$ together with S, T.*

Because of minimality and leftmost/rightmost-ness of the vertex separators S, T used to define σ, τ, σ', the shapes σ, τ, σ' have some special combinatorial structure, which we capture in the following proposition; the proof is until we state a generalized version.

Proposition 1.16 (See generalization for proof – **??**)**.** *σ, τ, and ${\sigma'}^T$ have the following properties:*

1. *$V_\sigma = S$ is the unique minimum vertex separator of σ.*
2. *S and T are the leftmost and rightmost minimum vertex separators of τ.*
3. *$T = U_{{\sigma'}^T}$ is the unique minimum vertex separator of ${\sigma'}^T$.*

Based on this, we define sets of shapes we can appear as left, middle, or right parts.

Definition 1.17 (Left, Middle, and Right Parts)**.** *Let α be a shape.*

1. *We say that α is a left part if V_α is the unique minimum vertex separator of α and $E(\alpha)$ has no edges which are entirely contained in V_α.*
2. *We say that α is a proper middle part if U_α is the leftmost minimum vertex separator of α and V_α is the rightmost minimum vertex separator of α*
3. *We say that α is a right part if U_α is the unique minimum vertex separator of α and $E(\alpha)$ has no edges which are entirely contained in U_α.*

Remark 1.18. *For technical reasons, later on we will need to consider improper middle parts τ where U_τ and V_τ are not the leftmost and rightmost minimum vertex separators of τ, which is why we make this distinction here.*

The following proposition is also straightforward from the definitions.

Proposition 1.19. *A shape σ is a left part if and only if σ^T is a right part*

Examples should be added.

1.3.2 Concatenation of Ribbons and Shapes and Products of Graph Matrices

Having discussed how to take a shape α apart into left, middle, and right parts, now we set about putting them back together again. The following definitions work towards a concatenation operation on shapes, with the idea that if α, β are shapes, then the matrix product $M_\alpha M_\beta$ should correspond to another graph matrix M_γ; we will identify γ as the concatenation of α and β, denoted $\alpha \cup \beta$.

Putting this together with the left/middle/right decomposition, eventually we aim for the approximations and equalities:

$$M_\alpha \approx M_\sigma M_\tau M_{(\sigma')^T} \approx M_{\sigma \cup \tau \cup (\sigma')^T} = M_\alpha$$

but it will take some work to get there.

The starting point will be an analogous concatenation operation on ribbons.

Definition 1.20 (Ribbon Concatenation, Simplified – see **??**). *If R_1 and R_2 are two ribbons such that $V(R_1) \cap V(R_2) = B_{R_1} = A_{R_2}$ and either R_1 or R_2 contains no edges entirely within $B_{R_1} = A_{R_2}$ then we define $R_1 \cup R_2$ to be the ribbon formed by glueing together R_1 and R_2 along $B_{R_1} = A_{R_2}$. In other words,*

1. $V(R_1 \cup R_2) = V(R_1) \cup V(R_2)$
2. $E(R_1 \cup R_2) = E(R_1) \cup E(R_2)$
3. $A_{R_1 \cup R_2} = A_{R_1}$ *and* $B_{R_1 \cup R_2} = B_{R_2}$.

The following proposition is easy to check.

Proposition 1.21. *Whenever R_1, R_2 are ribbons such that $R_1 \cup R_2$ is defined, $M_{R_1} M_{R_2} = M_{R_1 \cup R_2}$*

We have an analogous definition for concatenating shapes:

Definition 1.22 (Shape Concatenation, Simplified – see **??**). *If α_1 and α_2 are two shapes such that $V(\alpha_1) \cap V(\alpha_2) = V_{\alpha_1} = U_{\alpha_2}$ and either α_1 or α_2 contains no edges entirely within $V_{\alpha_1} = U_{\alpha_2}$ then we define $\alpha_1 \cup \alpha_2$ to be the shape formed by glueing together α_1 and α_2 along $V_{\alpha_1} = U_{\alpha_2}$. In other words,*

1. $V(\alpha_1 \cup \alpha_2) = V(\alpha_1) \cup V(\alpha_2)$
2. $E(\alpha_1 \cup \alpha_2) = E(\alpha_1) \cup E(\alpha_2)$
3. $U_{\alpha_1 \cup \alpha_2} = U_{\alpha_1}$ *and* $V_{\alpha_1 \cup \alpha_2} = V_{\alpha_2}$.

The next proposition, again easy to check, shows that the shape concatenation operation respects the left/middle/right part decomposition.

Proposition 1.23. *If σ, τ, σ'^T are the left, middle, and rights parts for α then $\alpha = \sigma \cup \tau \cup \sigma'^T$.*

On intersection terms As mentioned before, we will make use of the approximate factorization $M_\alpha = M_{\sigma\cup\tau\cup\sigma'^T} \approx M_\sigma M_\tau M_{\sigma'^T}$. In what sense is this factorization approximate?

Consider the difference $M_\sigma M_\tau M_{\sigma'^T} - M_{\sigma\cup\tau\cup\sigma'^T}$. The graph matrix $M_{\sigma\cup\tau\cup\sigma'^T}$ decomposes (by definition) into a sum over injective maps $\phi : V(\sigma \cup \tau \cup \sigma'^T) \to [n]$. Also by expanding definitions, the product $M_\sigma M_\tau M_{\sigma'^T}$ expands into a sum over triples of injective maps (ϕ_1, ϕ_2, ϕ_3), where $\phi_1 : V(\sigma) \to [n], \phi_2 : V(\tau) \to [n], \phi_3 : V(\sigma') \to [n]$, and ϕ_1 and ϕ_2 agree on $V_\sigma = U_\tau$ and ϕ_2 and ϕ_3 agree on $V_\tau = U_{\sigma'^T}$.

If they are combined into one map $\phi : V(\sigma\cup\tau\cup\sigma') \to [n]$, the resulting ϕ may not be injective because $\phi_1(V(\sigma)), \phi_2(V(\tau)), \phi_3(V(\sigma'^T))$ may have nontrivial intersection (beyond $\phi_1(V_\sigma)$ and $\phi_2(V_\tau)$). We call the resulting terms *intersection terms* and handling them properly is a major part of the technical analysis.

1.4 Comparing Fourier Coefficients via Coefficient Matrices

Our goal remains to set up enough Fourier analysis for the set of functions $\{\Lambda : \{\pm 1\}^{\binom{n}{2}} \to \mathbb{R}^{\binom{n}{d}\times\binom{n}{d}}$ to state a simplified version of our main theorem. We have established enough definitions to know that each such Λ has a unique expansion $\Lambda = \sum_\alpha \lambda_\alpha M_\alpha$, where the sum is over shapes, M_α is a graph matrix, and $\lambda_\alpha \in \mathbb{R}$.

The hypotheses of our main theorem concern the pattern of decay among Fourier coefficients λ_α of Λ. Capturing such a "decay pattern" requires a formalism to compare groups of Fourier coefficients to each other. The next family of definitions concerns objects we call *coefficient matrices* which are the main tool to describe the patterns of decay among the coefficients λ_α.

Definition 1.24 (Coefficient Matrices, Simplified – see **??**). *Given $\Lambda = \sum_\alpha \lambda_\alpha M_\alpha$, for each proper middle shape τ we define the coefficient matrix H_τ to be the matrix such that*

1. *The rows of H_τ are indexed by left parts σ where $V_\sigma = U_\tau$.*
2. *The columns of H_τ are indexed by left parts σ' where $V_{\sigma'} = V_\tau$.*
3. *H_τ has entries $H_\tau(\sigma, \sigma') = \lambda_{\sigma\cup\tau\cup\sigma'^T}$.*

An important special case is when α has a unique minimum vertex separator so the middle part τ is trivial.

Definition 1.25. *Define Id_r to be the shape with no edges where $U_{\mathrm{Id}_r} = V_{\mathrm{Id}_r}$ has r vertices. Note that M_{Id_r} is the constant function with value always equal to the $r \times r$ identity matrix.*

Definition 1.26. *We say that a shape α is trivial if α has no edges and $U_\alpha = V_\alpha$ as sets (so their ordering may be different). Otherwise, we say that α is non-trivial.*

Example 1.27. *H_{Id_r} has entries $H_{\mathrm{Id}_r}(\sigma, \sigma') = \lambda_{\sigma\cup\mathrm{Id}_r\cup\sigma'^T} = \lambda_{\sigma\cup\sigma'^T}$*

At a high level, our main theorem makes hypotheses on the spectra of coefficient matrices – for instance, $H_{\mathrm{Id}_r} \succeq 0$ – and concludes properties of the spectrum of Λ, as in (1). Note that a spectral condition on a coefficient matrix, such as $H_{\mathrm{Id}_r} \succeq 0$, already makes a comparison between Fourier coefficients of Λ. For instance it implies that for all $\sigma' \neq \sigma$,

$$|\lambda_{\sigma\cup\mathrm{Id}_r\cup\sigma'^T}| \cdot |\lambda_{\sigma'\cup\mathrm{Id}_r\cup\sigma^T}| \leq \lambda_{\sigma\cup\mathrm{Id}_r\cup\sigma^T} \cdot \lambda_{\sigma'\cup\mathrm{Id}_r\cup\sigma'^T}.$$

We need some terminology to express a more complex spectral condition on the coefficient matrices, comparing different H_τ's to each other. We call this *the* $-\gamma, \gamma$ *operation.*

Definition 1.28. *Given a coefficient matrix* H *and a left part* γ*, we define the coefficient matrix* $H^{-\gamma,\gamma}$ *to be the matrix with entries* $H^{-\gamma,\gamma}(\sigma,\sigma') = H(\sigma \cup \gamma, \sigma' \cup \gamma)$

1.5 Informal Theorem Statement

We are now ready to state a simplified version of our main theorem. For readability, in this statement, we have left out:

- Quantitative bounds needed for the applications to SoS lower bounds (in particular, quantitative expressions for $f(\gamma)$ and $f(\tau)$.
- Generalization beyond the setting $k = 2$ or to the case that the underlying product space is $\Omega^{\binom{n}{k}}$ for som $\Omega \neq \{\pm 1\}$.
- Generalization to the asymmeric setting that the underlying product space is $\Omega^{n_1 \times n_2 \times \ldots \times n_k}$.

Theorem 1.29. *There exist functions* $f(\tau)$: *middle parts* $\rightarrow \mathbb{R}$ *and* $f(\gamma)$: *left parts* $\rightarrow \mathbb{R}$ *depending on* n *and other parameters such that if* $\Lambda = \sum_\alpha \lambda_\alpha M_\alpha$ *and the following conditions hold:*

1. *For all* $r \in [0, \frac{d}{2}]$*,* $H_{Id_r} \succeq 0$
2. *For all* $r \in [0, \frac{d}{2}]$ *and all proper, non-trivial middle shapes* τ *whose left and right sides have* r *vertices,*
$$\begin{bmatrix} H_{Id_r} & f(\tau)H_\tau \\ f(\tau)H_\tau^T & H_{Id_r} \end{bmatrix} \succeq 0$$
If the matrices H_τ *are always symmetric, this reduces to the condition that* $f(\tau)H_\tau \preceq H_{Id_r}$
3. *For all non-trivial left parts* γ*,* $H^{-\gamma,\gamma}_{Id_{|V_\gamma|}} \preceq f(\gamma) H_{Id_{|U_\gamma|}}$

then with probability at least $1 - o(1)$ *over* $G \sim \{\pm 1\}^{\binom{n}{2}}$ *it holds that* $\Lambda(G) \succeq 0$*.*

Before we move on to further definitions needed for a more complete statement of the main theorem, we present an informal example

Example 1.30. *When the pseudocalibration method is applied to prove an SoS lower bound for the planted clique problem in* n *node graphs with clique size* $\omega = \omega(n)$*, as in [**?**], the matrix-valued function which results is* $\Lambda = \sum_{\alpha : |V(\alpha)| \leq t} \left(\frac{\omega}{n}\right)^{|V(\alpha)|} M_\alpha$ *where* $t \approx \log(n)$*. One may then compute that the matrices* H_r *and* H_τ *are as follows (at least so long as* $|V(\sigma)|, |V(\tau)|, |V(\sigma')| \ll t$*; we ignore this detail for now). For all* $r \in [0, \frac{d}{2}]$*,*

1. $H_r(\sigma,\sigma') = \left(\frac{\omega}{n}\right)^{|V(\sigma)|+|V(\sigma')|-r}$
2. *For all proper, non-trivial middle shapes* τ *such that* $|U_\tau| = |V_\tau| = r$*,*
$$H_\tau(\sigma,\sigma') = \left(\frac{\omega}{n}\right)^{|V(\sigma)|+|V(\sigma')|+|V(\tau)|-2r}$$

Defining v_r to be the vector such that $v_r(\sigma) = \left(\frac{\omega}{n}\right)^{|V(\sigma)|-\frac{r}{2}}$, we have that

1. *$H_r = v_r v_r^T$*
2. *For all proper, non-trivial middle shapes τ such that $|U_\tau| = |V_\tau| = r$, $H_\tau = \left(\frac{\omega}{n}\right)^{|V(\tau)|-r} v_r v_r^T$*
3. *For all left parts γ, $H^{-\gamma,\gamma}_{Id_{|V_\gamma|}} = \left(\frac{\omega}{n}\right)^{2|V(\gamma)|-|U_\gamma|-|V_\gamma|} v_{|U_\gamma|} v^T_{|U_\gamma|}$*

It turns out in this setting that we can take $f(\tau)$ to be $\tilde{O}(n^{\frac{|V(\tau)|-r}{2}})$ and $f(\gamma)$ to be $\tilde{O}(n^{|V(\gamma)\setminus U_\gamma|})$. Thus, as long as $\omega \ll \sqrt{n}$,

1. *For all τ such that $V_\tau \neq U_\tau$, $f(\tau)H_\tau \preceq H_{Id_r}$.*
2. *For all non-trivial left parts γ, $H^{-\gamma,\gamma}_{Id_{|V_\gamma|}} \preceq f(\gamma) H_{Id_{|U_\gamma|}}$*

Thus the conditions of the main theorem are satisfied for this function Λ.

Remark 1.31. *It should be noted that this theorem as stated (even with $f(\tau), f(\gamma)$ made quantitative) is not quite enough to prove the planted clique SoS lower bound of [**?**]. The reason is that there are τ such that $V_\tau = U_\tau$ but which are non-trivial because $E(\tau) \neq \emptyset$. In order to prove the planted clique lower bound, all of the τ where $V_\tau = U_\tau$ have to be grouped together into the indicator function for whether $V_\tau = U_\tau$ is a clique.*

2 Technical Definitions and Theorem Statement

2.1 Section Introduction

In this section, we make our definitions and results more precise. We also generalize our definitions and results to handle problems where one or more of the following is true:

1. The input entries correspond to hyperedges rather than edges.
2. We have different types of indices.
3. Ω is a more complicated distribution than $\{-1, +1\}$.
4. We have to consider matrix indices which are not multilinear.

Throughout this section and the remainder of this manuscript, we give the reader a choice for the level of generality of this machinery. In particular, we will first recall our definition for the simpler case when our input is $\{-1, +1\}^{\binom{n}{2}}$ and we only consider multilinear indices. We will then discuss how this simpler definition generalizes. We denote these generalizations with an asterix $*$.

2.1.1 Additional Parameters for the General Case*

In the general case we will need a few additional parameters which we define here.

Definition 2.1.

1. *We define k to be the arity of the hyperedges corresponding to the input.*
2. *We define t_{max} to be the number of different types of indices. We define n_i to be the number of possibilities for indicies of type i and we define $n = \max\{n_i : i \in [t_{max}]\}$.*

2.2 Indices, Input Entries, Vertices, and Edges

Note: For this section, we use X to denote the input, we use x to denote entries of the input and we use y to denote solution variables.

Definition 2.2 (Vertices: Simplified Case)**.** *When the input and solution variables are indexed by one type of index which takes values in $[n]$ then we represent the index i by a vertex labeled i.*

If we want to leave an index unspecified, we instead represent it by a vertex labeled with a variable (we will generally use u, v, or w for these variables).

Definition 2.3 (Vertices: General Case*)**.** *When the input and solution variables are indexed by several types of indices where indices of type t take values in $[n_t]$, we represent an index of type t with value i as a vertex labeled by the tuple (t, i). We say that such a vertex has type t.*

If we want to leave an index of type t unspecified, we instead represent it by a vertex labeled with a tuple $(t, ?)$ where $?$ is a variable (which will generally be u, v, or w).

Definition 2.4 (Edges: Simplified Case)**.** *When the input is $X \in \{-1, +1\}^{\binom{n}{2}}$, we represent the entries of the input by the undirected edges $\{(i, j) : i < j \in [n]\}$. Given an edge $e = (i, j)$, we take $x_e = x_{ij}$ to be the input entry corresponding to e.*

Definition 2.5 (Edges: General Case*)**.** *In general, we represent the entries of the input by hyperedges whose form depends on nature of the input. We still take x_e to be the input entry corresponding to e.*

Example 2.6. *If the input is an $n_1 \times n_2$ matrix X then we will have two types of indices, one for the row and one for the column. Thus, we will have the vertices $\{(1, i) : i \in [n_1]\} \cup \{(2, j) : j \in [n_2]\}$. In this case, we have an edge $((1, i), (2, j))$ for each entry x_{ij} of the input.*

Example 2.7. *If the input is an $n \times n$ matrix X which is not symmetric then we only need the indices $[n]$. In this case, we have a directed edge (i, j) for each entry x_{ij} where $i \neq j$. If the entries x_{ii} are also part of the input than we also have loops (i, i) for these entries.*

Example 2.8. *If our input is a symmetric $n \times n \times n$ tensor X (i.e. $x_{ijk} = x_{ikj} = x_{jik} = x_{jki} = x_{kij} = x_{kji}$) and $x_{ijk} = 0$ whenever i, j, k are not distinct then we only need the indices $[n]$. In this case, we have an undirected hyperedge $e = (i, j, k)$ for each entry $x_e = x_{ijk}$ of the input where i, j, k are distinct.*

Example 2.9. *If the input is an $n_1 \times n_2 \times n_3$ tensor X then we will have three types of indices. Thus, we will have the vertices $\{(1, i) : i \in [n_1]\} \cup \{(2, j) : j \in [n_2]\} \cup \{(3, k) : k \in [n_3]\}$. In this case, we have a hyperedge $e = ((1, i), (2, j), (3, k))$ for each entry $x_e = x_{ijk}$ of the input.*

2.3 Matrix Indices and Monomials

In this subsection, we discuss how our matrices are indexed and how we associate matrix indices with monomials. We also describe the automorphism groups of matrix indices.

Definition 2.10 (Matrix Indices: Simplified Case)**.** *If there is only one type of index and we have the constraints $y_i^2 = 1$ or $y_i^2 = y_i$ on the solution variables then we define a matrix index A to be a tuple of indices $(a_1, \ldots, a_{|A|})$. We make the following definitions about matrix indices:*

1. *We associate the monomial* $\prod_{j=1}^{|A|} y_{a_j}$ *to* A.

2. *We define* $V(A)$ *to be the set of vertices* $\{a_i : i \in [|A|]\}$. *For brevity, we will often write* A *instead of* $V(A)$ *when it is clear from context that we are referring to* A *as a set of vertices rather than a matrix index.*

3. *We take the automorphism group of* A *to be* $Aut(A) = S_{|A|}$ *(the permutations of the elements of* A*)*

Example 2.11. *The matrix index* $A = (4, 6, 1)$ *represents the monomial* $y_4y_6y_1 = y_1y_4y_6$ *and* $Aut(A) = S_3$

Remark 2.12. *We take* A *to be an ordered tuple rather than a set for technical reasons.*

In general, we need a more intricate definition for matrix indices. We start by defining matrix index pieces

Definition 2.13 (Matrix Index Piece Definition*)**.** *We define a matrix index piece* $A_i = ((a_{i1}, \ldots, a_{i|A_i|}), t_i, p_i)$ *to be a tuple of indices* $(a_{i1}, \ldots, a_{i|A_i|})$ *together with a type* t_i *and a power* p_i. *We make the following definitions about matrix index pieces:*

1. *We associate the monomial* $p_{A_i} = \prod_{j=1}^{|A_i|} y_{t_ij}^{p_i}$ *with* A_i.

2. *We define* $V(A_i)$ *to be the set of vertices* $\{(t_i, a_{ij}) : j \in [|A_i|]\}$.

3. *We take the automorphism group of* A_i *to be* $Aut(A_i) = S_{|A_i|}$

4. *We say that* A_i *and* A_j *are disjoint if* $V(A_i) \cap V(A_j) = \emptyset$ *(i.e.* $t_i \neq t_j$ *or* $\{a_{i1}, \ldots, a_{i|A_i|}\} \cap \{a_{j1}, \ldots, a_{j|A_j|}\} = \emptyset$*)*

Definition 2.14 (General Matrix Index Definition*)**.** *We define a matrix index* $A = \{A_i\}$ *to be a set of disjoint matrix index pieces. We make the following definitions about matrix indices:*

1. *We associate the monomial* $p_A = \prod_{A_i \in A} p(A_i)$ *with* A.

2. *We define* $V(A)$ *to be the set of vertices* $\cup_{A_i \in A} V(A_i)$. *For brevity, we will often write* A *instead of* $V(A)$ *when it is clear from context that we are referring to* A *as a set of vertices rather than a matrix index.*

3. *We take the automorphism group of* A *to be* $Aut(A) = \prod_{A_i \in A} Aut(A_i)$

Example 2.15 (*)**.** *If* $A_1 = ((2), 1, 1)$, $A_2 = ((3, 1), 1, 2)$, *and* $A_3 = ((1, 2, 3), 2, 1)$ *then* $A = \{A_1, A_2, A_3\}$ *represesents the monomial* $p = y_{12}y_{13}^2y_{11}^2y_{21}y_{22}y_{23}$ *and we have* $Aut(A) = S_1 \times S_2 \times S_3$

2.4 Fourier Characters and Ribbons

A key idea is to analyze Fourier characters of the input.

Definition 2.16 (Simplified Fourier Characters)**.** *If the input distribution is $\Omega = \{-1, 1\}$ then given a multi-set of edges E, we define $\chi_E(X) = \prod_{e \in E} x_e$.*

Example 2.17. *If the input is a graph $G \in \{-1, 1\}^{\binom{n}{2}}$ and E is a set of potential edges of G (with no multiple edges) then $\chi_E(G) = (-1)^{|E \setminus E(G)|}$.*

In general, the Fourier characters are somewhat more complicated.

Definition 2.18 (Orthonormal Basis for Ω*)**.** *We define the polynomials $\{h_i : i \in \mathbb{Z} \cap [0, |supp(\Omega)| - 1]\}$ to be the unique polynomials (which can be found through the Gram-Schmidt process) such that*

1. *$\forall i, E_\Omega[h_i^2(x)] = 1$*
2. *$\forall i \neq j, E_\Omega[h_i(x)h_j(x)] = 0$*
3. *For all i, the leading coefficient of $h_i(x)$ is positive.*

Example 2.19. *If Ω is the normal distribution then the polynomials $\{h_i\}$ are the Hermite polynomials with the appropriate normalization so that for all i, $E_\Omega[h_i^2(x)] = 1$. In particular, $h_0(x) = 1$, $h_1(x) = x$, $h_2(x) = \frac{x^2-1}{\sqrt{2!}}$, $h_3(x) = \frac{x^3-3x}{\sqrt{3!}}$, etc.*

Definition 2.20 (General Fourier Characters*)**.** *Given a multi-set of hyperedges E, each of which has a label $l(e) \in [|support(\Omega)| - 1]$ (or $\mathbb{N}$ if Ω has infinite support), we define $\chi_E = \prod_{e \in E} h_{l(e)}(X_e)$.*

We say that such a multi-set of hyperedges E is proper if it contains no duplicate hyperedges, i.e. it is a set (though the labels on the hyperedges can be arbitrary non-negative integers). Otherwise, we say that E is improper.

Remark 2.21. *The Fourier characters are $\{\chi_E : E \text{ is proper}\}$. For improper E, χ_E can be decomposed as a linear combination of χ_{E_j} where each E_j is proper. We allow improper E because it is sometimes more convenient to have improper E in the middle of the analysis and then do this decomposition at the end.*

Definition 2.22 (Ribbons)**.** *A ribbon R is a tuple (H_R, A_R, B_R) where H_R is a multi-graph (*or multi-hypergraph with labeled edges in the general case) whose vertices are indices of the input and A_R and B_R are matrix indices such that $V(A_R) \subseteq V(H_R)$ and $V(B_R) \subseteq V(H_R)$. We make the following definitions about ribbons:*

1. *We define $V(R) = V(H_R)$ and $E(R) = E(H_R)$*
2. *We define $\chi_R = \chi_{E(R)}$.*
3. *We define M_R to be the matrix such that $(M_R)_{A_R B_R} = \chi_R$ and $M_{AB} = 0$ whenever $A \neq A_R$ or $B \neq B_R$.*

We say that R is a proper ribbon if H_R contains no isolated vertices outside of $A_R \cup B_R$ and $E(R)$ is proper. If there is an isolated vertex in $(V(R) \setminus A_R) \setminus B_R$ or $E(R)$ is improper then we say that R is an improper ribbon.

Proper ribbons are useful because they give an orthonormal basis for the space of matrix valued functions.

Definition 2.23 (Inner products of matrix functions)**.** *For a pair of real matrices M_1, M_2 of the same dimension, we write $\langle M_1, M_2\rangle = tr(M_1 {M_2}^T)$ (i.e. $\langle M_1, M_2\rangle$ is the entrywise dot product of M_1 and M_2). For a pair of matrix-valued functions M_1, M_2 (of the same dimensions), we define*

$$\langle M_1, M_2\rangle = E_X\left[\langle M_1(X), M_2(X)\rangle\right]$$

Proposition 2.24. *If R and R' are two proper ribbons then $\langle M_R, M_{R'}\rangle = 1$ if $R = R'$ and is 0 otherwise.*

2.5 Shapes

In this subsection, we describe a basis for S-invariant matrix valued functions where each matrix in this basis can be described by a relatively small *shape* α. The fundamental idea behind shapes is that we keep the structure of the objects we are working with but leave the elements of the object unspecified.

2.5.1 Simplified Index Shapes

Definition 2.25 (Simplified Index shapes)**.** *With our simplifying assumptions, an index shape U is a tuple of unspecified indices $(u_1, \cdots, u_{|U|})$. We make the following definitions about index shapes:*

1. *We define $V(U)$ to be the set of vertices $\{u_i : i \in [|U|]\}$. For brevity, we will often write U instead of $V(U)$ when it is clear from context that we are referring to U as a set of vertices rather than an index shape.*
2. *We define the weight of U to be $w(U) = |U|$.*
3. *We take the automorphism group of U to be $Aut(U) = S_{|U|}$ (the permutations of the elements of U)*

Definition 2.26. *We say that a matrix index $A = (a_1, \ldots, a_{|A|})$ has index shape $U = (u_1, \ldots, u_{|U|})$ if $|U| = |A|$. Note that in this case, if we take the map $\phi : \{u_j : j \in [|U|]\} \to [n]$ where $\phi(u_j) = a_j$ then $\phi(U) = (\phi(u_1), \ldots, \phi(u_{|U|})) = (a_1, \ldots, a_{|A|}) = A$*

Definition 2.27. *We say that index shapes $U = (u_1, \ldots, u_{|U|})$ and $V = (v_1, \ldots, v_{|V|})$ are equivalent (which we write as $U \equiv V$) if $|U| = |V|$. If $U \equiv V$ then we can set $U = V$ by setting $v_j = u_j$ for all $j \in [|U|]$.*

Example 2.28. *The matrix index $A = \{4, 6, 1\}$ has shape $U = \{u_1, u_2, u_3\}$ which has weight 3.*

2.5.2 General Index Shapes*

In general, we define general index shapes in the same way that we defined general matrix indices (just with unspecified indices)

Definition 2.29 (Index Shape Piece Definition)**.** *We define a index shape piece $U_i = ((u_{i1}, \ldots, u_{i|U_i|}), t_i, p_i)$ to be a tuple of indices $(u_{i1}, \ldots, u_{i|A_i|})$ together with a type t_i and a power p_i. We make the following definitions about index shape pieces:*

1. *We define $V(U_i)$ to be the set of vertices $\{(t_i, u_{ij}) : j \in [|U_i|]\}$.*
2. *We define $w(U_i) = |U_i| log_n(n_{t_i})$*
3. *We take the automorphism group of U_i to be $Aut(U_i) = S_{|U_i|}$*

Definition 2.30 (General Index Shape Definition)**.** *We define an index shape $U = \{U_i\}$ to be a set of index shape pieces such that for all $i' \neq i$, either $t_{i'} \neq t_i$ or $p_{i'} \neq p_i$. We make the following definitions about index shapes:*

1. *We define $V(U)$ to be the set of vertices $\cup_{U_i \in U} V(U_i)$. For brevity, we will often write U instead of $V(U)$ when it is clear from context that we are referring to U as a set of vertices rather than an index shape.*
2. *We define $w(U)$ to be $w(U) = \sum_{U_i \in U} w(U_i)$*
3. *We take the automorphism group of U to be $Aut(U) = \prod_{U_i \in U} Aut(U_i)$*

Remark 2.31. *For technical reasons, we want to ensure that if two index shapes U and U' have the same weight then U and U' have the same number of each type of vertex. To ensure this, we add an infinitesimal perturbation to each n_i if necessary.*

Definition 2.32. *We say that a matrix index A has index shape U if there is an assignment of values to the unspecified indices of U which results in A. More precisely, we say that A has index shape U if there is a map $\phi : \{u_{ij}\} \to \mathbb{N}$ such that if we define $\phi(U_i)$ to be $\phi(U_i) = ((\phi(u_{i1}), \ldots, \phi(u_{i|U_i|})), t_i, p_i)$ then $\phi(U) = \{\phi(U_i)\} = \{A_i\} = A$.*

Definition 2.33. *If U and V are two index shapes, we say that U is equivalent to V (which we write as $U \equiv V$) if U and V have the same number of index shape pieces and we can order the index shape pieces of U and V so that writing $U = \{U_i\}$ and $V = \{V_i\}$ where $U_i = ((u_{i1}, \ldots, u_{i|U_i|}), t_i, p_i)$ and $V_i = ((v_{i1}, \ldots, v_{i|V_i|}), t'_i, p'_i)$, we have that for all i, $|V_i| = |U_i|$, $t'_i = t_i$, and $p'_i = p_i$. If $U \equiv V$ then we can set $U = V$ by setting $u_{ij} = v_{ij}$ for all i and all $j \in [|U_i|]$.*

2.5.3 Ribbon Shapes

With these definitions, we are now ready to define shapes and the matrices associated to them.

Definition 2.34 (Shapes)**.** *A ribbon shape α (which we call a shape for brevity) is a tuple $\alpha = (H_\alpha, U_\alpha, V_\alpha)$ where H_α is a multi-graph (*or multi-hypergraph with labeled edges in the general case) whose vertices are unspecified distinct indices of the input (*whose type is specified in the general case) and U_α and V_α are index shapes such that $V(U_\alpha) \subseteq V(H_\alpha)$ and $V(V_\alpha) \subseteq V(H_\alpha)$. We make the following definitions about shapes:*

1. *We define $V(\alpha) = V(H_\alpha)$ (note that $V(\alpha)$ and V_α are not the same thing) and we define $E(\alpha) = E(H_\alpha)$.*

2. *We say that a shape α is proper if it contains no isolated vertices outside of $V(U_\alpha) \cup V(V_\alpha)$ and $E(\alpha)$ has no multiple edges/hyperedges. If there is an isolated vertex in $V(\alpha) \setminus V(U_\alpha) \setminus V(V_\alpha)$ or $E(\alpha)$ has a multiple edge/hyperedge then we say that α is an improper shape.*

Note: For brevity, we will often write U_α and V_α instead of $V(U_\alpha)$ and $V(V_\alpha)$ when it is clear from context that we are referring to U_α and V_α as sets of vertices rather than index shapes.

Definition 2.35 (Trivial shapes)**.** *We say that a shape α is trivial if $V(\alpha) = V(U_\alpha) = V(V_\alpha)$ and $E(\alpha) = \emptyset$. Otherwise, we say that α is non-trivial.*

Remark 2.36. *Note that all trivial shapes can do is permute the order of the vertices in $V(U_\alpha) = V(V_\alpha)$.*

Definition 2.37. *Informally, we say that a ribbon R has shape α if replacing the indices in R with unspecified labels results in α. Formally, we say that R has shape α if there is an injective mapping $\phi : V(\alpha) \to [n]$ (*or $[t_{max}] \times [n]$ in the general case) such that $\phi(\alpha) = R$, i.e. $\phi(H_\alpha) = H_R$, $\phi(U_\alpha) = A_R$, and $\phi(V_\alpha) = B_R$*

Definition 2.38. *We say that two shapes α and β are equivalent (which we write as $\alpha \equiv \beta$) if they are the same up to renaming their indices. More precisely, we say that $\alpha \equiv \beta$ if there is a bijective map $\pi : V(H_\alpha) \to V(H_\beta)$ such that $\pi(H_\alpha) = H_\beta$, $\pi(U_\alpha) = U_\beta$, and $\pi(V_\alpha) = V_\beta$.*

Definition 2.39. *Given a shape α and matrix indices A, B of shapes U_α and V_α respectively, we define $\mathcal{R}(\alpha, A, B)$ to be the set of ribbons R such that R* has shape *α, $A_R = A$, and $B_R = B$.*

Definition 2.40. *For a shape α, we define the matrix-valued function M_α to have entries $M_\alpha(A, B)$ given by*

$$(M_\alpha)_{A,B}(X) = \sum_{R \in \mathcal{R}(\alpha,A,B)} \chi_R(X)$$

To be added: (examples of M_α)

Proposition 2.41. *The M_α's for proper shapes α are an orthogonal basis for the S-invariant functions.*[1]

Remark 2.42. *Conceptually, one may think of forming an orthonormal basis for this space with the functions $M_\alpha/\sqrt{\langle M_\alpha, M_\alpha \rangle}$, but for technical reasons it is easiest to work with these functions without normalizing them to 1. By orthogonality and the fact that every Boolean function is a polynomial, any S-invariant matrix-valued function Λ is expressible as*

$$\Lambda = \sum_\alpha \frac{\langle \Lambda, M_\alpha \rangle}{\langle M_\alpha, M_\alpha \rangle} \cdot M_\alpha$$

[1] Because of orthogonality of the underlying Fourier characters, is not hard to check that when $\alpha \neq \alpha'$ and $M_\alpha, M_{\alpha'}$ have the same dimensions, $\langle M_\alpha, M_{\alpha'} \rangle = 0$.

2.6 Composing Ribbons and Shapes

Definition 2.43 (Composing Ribbons)**.** *We say that ribbons R_1 and R_2 are composable if $B_{R_1} = A_{R_2}$. Note that this definition is not symmetric so we may have that R_1 and R_2 are composable but R_2 and R_1 are not composable.*

We say that R_1 and R_2 are properly composable if we also have that $V(R_1) \cap V(R_2) = V(B_{R_1}) = V(A_{R_2})$ (there are no unexpected intersections between R_1 and R_2).

If R_1 and R_2 are composable ribbons then we define the composition of R_1 and R_2 to be the ribbon $R_1 \circ R_2$ such that

1. $A_{R_1 \circ R_2} = A_{R_1}$ *and* $B_{R_1 \circ R_2} = B_{R_2}$
2. $V(R_1 \circ R_2) = V(R_1) \cup V(R_2)$
3. $E(R_1 \circ R_2) = E(R_1) \cup E(R_2)$ *(and thus $\chi_{R_1 \circ R_2} = \chi_{R_1}\chi_{R_2}$)*

We say that ribbons $R_1, \ldots, R_k$ are composable/properly composable if for all $j \in [k-1]$, $R_1 \circ \ldots \circ R_j$ and R_{j+1} are composable/properly composable. If $R_1, \ldots, R_k$ are composable then we define $R_1 \circ \ldots \circ R_k$ to be $R_1 \circ \ldots \circ R_k = (R_1 \circ \ldots \circ R_{k-1}) \circ R_k$

Proposition 2.44. *Ribbon composition is associative, i.e. if R_1, R_2, R_3 are composable/properly composable ribbons then R_2, R_3 are composable/properly composable, $R_1, (R_2 \circ R_3)$ are composable/properly composable, and $R_1 \circ (R_2 \circ R_3) = (R_1 \circ R_2) \circ R_3$*

Proposition 2.45. *If R_1 and R_2 are composable ribbons then $M_{R_1 \cup R_2} = M_{R_1} M_{R_2}$.*

We have similar definitions for composing shapes.

Definition 2.46 (Composing Shapes)**.** *We say that shapes α and β are composable if $U_\beta \equiv V_\alpha$. Note that this definition is not symmetric so we may have that α and β are composable but β and α are not composable.*

If α and β are composable shapes then we define the composition of α and β to be the shape $\alpha \circ \beta$ such that

1. $U_{\alpha \circ \beta} = U_\alpha$ *and* $V_{\alpha \circ \beta} = V_\beta$
2. *After setting $U_\beta = V_\alpha$, we take $V(\alpha \circ \beta) = V(\alpha) \cup V(\beta)$*
3. $E(\alpha \circ \beta) = E(\alpha) \cup E(\beta)$

We say that shapes $\alpha_1, \ldots, \alpha_k$ are composable if for all $j \in [k-1]$, $\alpha_1 \circ \ldots \circ \alpha_j$ and α_{j+1} are composable. If $\alpha_1, \ldots, \alpha_k$ are composable then we define the shape $\alpha_1 \circ \ldots \circ \alpha_k$ to be $\alpha_1 \circ \ldots \circ \alpha_k = (\alpha_1 \circ \ldots \circ \alpha_{k-1}) \circ \alpha_k$

Proposition 2.47. *Shape composition is associative, i.e. if $\alpha_1, \alpha_2, \alpha_3$ are composable shapes then α_2, α_3 are composable, $\alpha_1, (\alpha_2 \circ \alpha_3)$ are composable, and $\alpha_1 \circ (\alpha_2 \circ \alpha_3) = (\alpha_1 \circ \alpha_2) \circ \alpha_3$*

2.7 Decomposition of Shapes into Left, Middle, and Right parts

In this subsection, we describe how shapes can be decomposed into left, middle, and right parts based on the leftmost and rightmost *minimum vertex separators*, which is a crucial idea for our analysis.

Definition 2.48 (Paths)**.** *A* path *in a shape α is a sequence of vertices $v_1, \ldots, v_t$ such that v_i, v_{i+1} are in some edge/hyperedge together. A pair of paths is vertex-disjoint if the corresponding sequences of vertices are disjoint.*

Definition 2.49 (Vertex separators)**.** *Let α be a shape and let U and V be sets of vertices in α. We say that a set of vertices $S \subseteq V(\alpha)$ is a* vertex separator *of U and V if every path in α from U to V contains at least one vertex in S. Note that any vertex separator S of U and V must contain all of the vertices in $U \cap V$.*

As a special case, we say that S is a vertex separator of α if S is a vertex separator of U_α and V_α

We define the weight of a set of vertices $S \subseteq V(\alpha)$ in the same way that weight is defined for index shapes.

Definition 2.50 (Simplified Weight)**.** *When there is only one type of index, the weight of a set of vertices $S \subseteq V(\alpha)$ is simply $|S|$.*

Definition 2.51 (General Weight*)**.** *In general, given a set of vertices $S \subseteq V(\alpha)$, writing $S = \cup_t S_t$ where S_t is the set of vertices of type t in S, we define the weight of S to be $w(S) = \sum_t |S_t| log_n(n_t)$*

Remark 2.52 (*)**.** *Again, if necessary, we add an infinitesimal perturbation to $n_1, n_2, \ldots, n_{t_{max}}$ so that if two separators S and S' have the same weight then S and S' have the same number of each type of vertex.*

Definition 2.53 (Leftmost and rightmost minimum vertex separators)**.** *The* leftmost *minimum vertex separator is the vertex separator S of minimum weight such that for every other minimum-weight vertex separator S', S is a separator of U_α and S'. The* rightmost *minimum vertex separator is the vertex separator T of minimum weight such that for every other minimum-weight vertex separator T', T is a separator of T' and V_α*

The work [1] showed that under our simplifying assumptions, leftmost and rightmost minimum vertex separators are well defined. For a general proof that leftmost and rightmost minimum vertex separators are well defined, see Appendix A.

We now have the following crucial idea. Every shape α can be decomposed into the composition of three composable shapes $\sigma, \tau, {\sigma'}^T$ based on the leftmost and rightmost minimum vertex separators S, T of α together with orderings of S and T.

Definition 2.54 (Simplified Separators With Orderings)**.** *Under our simplifying assumptions, given a set of vertices $S \subseteq V(\alpha)$ and an ordering $O_S = s_1, \ldots, s_{|S|}$ of the vertices of S, we define the index shape (S, O_S) to be $(S, O_S) = (s_1, \ldots, s_{|S|})$.*

Definition 2.55 (General Separators With Orderings*). *In the general case, we need to give an ordering for each type of vertex. Let $S \subseteq V(\alpha)$ be a subset of the vertices of α and write $S = \cup_t S_t$ where S_t is the set of vertices in S of type t. Given $O_S = \{O_t\}$ where $O_t = s_{t1}, \ldots, s_{t|S_t|}$ is an ordering of the vertices of S_t, we define the index shape piece (S_t, O_t) to be $(S_t, O_t) = ((s_{t1}, \ldots, s_{t|S_t|}), t, 1)$ and we define the index shape (S, O_S) to be $(S, O_S) = \{(S_t, O_t)\}$.*

Proposition 2.56. *The number of possible orderings O for S is equal to $|Aut((S, O_S))|$*

Definition 2.57 (Shape transposes). *Given a shape α, we define α^T to be the shape α with U_α and V_α swapped i.e. $U_{\sigma^T} = V_\sigma$ and $V_{\sigma^T} = U_\sigma$.*

Definition 2.58 (Left, middle, and right parts). *Let α be a shape. Let S and T be the leftmost and rightmost minimal vertex separators of α together with orderings O_S, O_T of S and T.*

- *We define the* left part *σ_α of α to be the shape such that*
  1. *H_{σ_α} is the induced subgraph of H_α on all of the vertices of α reachable from U_α without passing through S (note that H_{σ_α} includes the vertices of S) except that we remove any edges/hyperedges which are contained entirely within S.*
  2. *$U_{\sigma_\alpha} = U_\alpha$ and $V_{\sigma_\alpha} = (S, O_S)$*
- *We define the* right part *${\sigma'}_\alpha^T$ of α to be the shape such that*
  1. *$H_{{\sigma'}_\alpha^T}$ is the induced subgraph of H_α on all of the vertices of α reachable from V_α without passing through T (note that $H_{{\sigma'}_\alpha^T}$ includes the vertices of T) except that we remove any edges/hyperedges which are contained entirely within T.*
  2. *$V_{{\sigma'}_\alpha^T} = V_\alpha$ and $U_{{\sigma'}_\alpha^T} = (T, O_T)$*
- *We define the* middle part *τ_α of α to be the shape such that*
  1. *H_{τ_α} is the induced subgraph of H_α on all of the vertices of α which are not reachable from U_α and V_α without touching S and T (note that H_{τ_α} includes the vertices of S and T). H_{τ_α} also includes the hyperedges entirely within S and the hyperedges entirely within T.*
  2. *$U_{\tau_\alpha} = (S, O_S)$ and $V_{\tau_\alpha} = (T, O_T)$*

.

Proposition 2.59. *If $\sigma, \tau, {\sigma'}^T$ are the left, middle, and rights parts for α for given orderings O_S, O_T of S and T then $\alpha = \sigma \circ \tau \circ {\sigma'}^T$.*

Remark 2.60. *One may ask which ordering(s) we should take of S and T. The answer is that we will take all of the possible orderings of S and T simultaneously, giving equal weight to each.*

Based on this decomposition and the following claim, we make the following definitions for what it means for a shape to be a left, middle, or right part.

Claim 2.61 (Proved in Section 6.1 in [1]). [2]

- *Every shape σ which is the left part of some other shape α has that V_σ is its left-most and right-most minimum-weight separator.*
- *Every shape σ^T which is the right part of some other shape α has that U_{σ^T} is its left-most and right-most minimum-weight separator.*
- *Every shape τ which is the middle part of some other shape α has U_τ as its left-most minimum size separator and V_τ as its right-most minimum-weight separator.*

Definition 2.62.

1. *We say that a shape σ is a left shape if σ is a proper shape, V_σ is the left-most and right-most minimum-weight separator of σ, every vertex in $V(\sigma) \setminus V_\sigma$ is reachable from U_σ without touching V_σ, and σ has no hyperedges entirely within V_σ.*
2. *We say that a shape τ is a proper middle shape if τ is a proper shape, U_τ is the left-most minimum-weight separator of τ, and V_τ is the right most minimum-weight separator of τ. In the analysis, we will also need to consider improper middle shapes τ which may not be proper shapes and which may have smaller separators between U_τ and V_τ.*
3. *We say that a shape σ^T is a right shape if σ^T is a proper shape, U_{σ^T} is the left-most and right-most minimum-weight separator of σ^T, every vertex in $V(\sigma^T) \setminus U_{\sigma^T}$ is reachable from V_{σ^T} without touching U_{σ^T}, and σ^T has no hyperedges entirely within U_{σ^T}.*

Proposition 2.63. *For all shapes σ, σ is a left shape if and only if σ^T is a right shape.*

Remark 2.64. *As the reader has likely guessed, throughout this section we use σ to denote left parts and τ to denote middle parts. Instead of having a separate letter for right parts, we express right parts as the transpose of a left part.*

2.8 Coefficient matrices

We will have that $\Lambda = \sum_\alpha \lambda_\alpha M_\alpha$. To analyze Λ, it is extremely useful to express these coefficients in terms of matrices. To do this, we will need a few more definitions. We start by defining the sets of index shapes that can appear when analyzing Λ.

Definition 2.65. *Given a moment matrix Λ, we define the following sets of index shapes.*

1. *We define $\mathcal{I}(\Lambda) = \{U : \exists \text{ matrix index } A : A \text{ is a row index of } \Lambda, A \text{ has shape } U\}$ to be the set of index shapes which describe row and column indices of Λ.*
2. *We define w_{max} to be $w_{max} = \max\{w(U) : U \in \mathcal{I}(\Lambda)\}$.*
3. *With our simplifying assumptions, we define $\mathcal{I}_{mid}$ to be $\mathcal{I}_{mid} = \{U : |U| \leq w_{max}\}$*

[2] The proof in [1] only explicitly treats the case when the shapes α are graphs, but the proof easily generalizes to the case when the α are hypergraphs.

3*. *In general, we define* $\mathcal{I}_{mid}$ *to be* $\mathcal{I}_{mid} = \{U : w(U) \leq w_{max}, \forall U_i \in U, p_i = 1\}$

We also need to define the sets of shapes which can appear when analyzing Λ.

Definition 2.66 (Truncation Parameters). *Given a moment matrix* $\Lambda = \sum_{\alpha} \lambda_{\alpha} M_{\alpha}$*, we define* D_V, D_E *to be the smallest natural numbers such that for all shapes* α *such that* $\lambda_{\alpha} \neq 0$*, decomposing* α *as* $\alpha = \sigma \circ \tau \circ {\sigma'}^T$*,*

1. $|V(\sigma)| \leq D_V$*,* $|V(\tau)| \leq D_V$*, and* $|V(\sigma')| \leq D_V$*.*

2.* *For all edges* $e \in E(\sigma) \cup E(\tau) \cup E(\sigma')$*,* $l_e \leq D_E$*.*

Remark 2.67. *Under our simplifying assumptions, all edges have label* 1 *so we will take* $D_E = 1$ *and ignore conditions involving* D_E*.*

Definition 2.68. *Given a moment matrix* Λ*, we define the following sets of shapes:*

1. $\mathcal{L} = \{\sigma : \sigma \text{ is a left shape}, U_{\sigma} \in \mathcal{I}(\Lambda), V_{\sigma} \in \mathcal{I}_{mid}, |V(\sigma)| \leq D_V, \forall e \in E(\sigma), l_e \leq D_E\}$

2. *Given* $V \in \mathcal{I}_{mid}$*, we define* $\mathcal{L}_V = \{\sigma \in \mathcal{L} : V_{\sigma} \equiv V\}$

3. *Given* $U \in \mathcal{I}_{mid}$*, we define* $\mathcal{M}_U = \{\tau : \tau \text{ is a non-trivial proper middle shape}, U_{\tau} \equiv V_{\tau} \equiv U, |V(\tau)| \leq D_V, \forall e \in E(\tau), l_e \leq D_E\}$

Definition 2.69. *Given a moment matrix* Λ*, we define a* Λ*-coefficient matrix (which we call a coefficient matrix for brevity) to be a matrix whose rows and columns are indexed by left shapes* $\sigma, \sigma' \in \mathcal{L}$*.*

We say that a coefficient matrix H *is SOS-symmetric if* $H(\sigma, \sigma')$ *is invariant under permuting the vertices of* U_{σ} *and permuting the vertices of* $U_{\sigma'}$ *(*more precisely, for the general case we permute the vertices within each index shape piece of* U_{σ} *and permute the vertices within each index shape piece of* $U_{\sigma'}$*).*

Definition 2.70. *Given a shape* τ*, we say that a coefficient matrix* H *is a* τ*-coefficient matrix if* $H(\sigma, \sigma') = 0$ *whenever* $V_{\sigma} \not\equiv U_{\tau}$ *or* $V_{\tau} \not\equiv U_{{\sigma'}^T}$*.*

Definition 2.71. *Given an index shape* U*, we define* Id_U *to be the shape with* $U_{Id_U} = V_{Id_U} = U$*, no other vertices, and no edges.*

Given a shape τ and a τ-coefficient matrix H, we create two different matrix-valued functions, $M^{fact}_{\tau}(H)$ and $M^{orth}_{\tau}(H)$. As we will see, we can express Λ in terms of M^{orth} but to show PSDness we will need to shift to M^{fact}. We analyze the difference betweem M^{fact} and M^{orth} in subsections 3.2, 3.3, and 3.4.

Definition 2.72. *Given a shape* τ *and a* τ*-coefficient matrix* H*, define*

$$M^{fact}_{\tau}(H) = \sum_{\sigma \in \mathcal{L}_{U_{\tau}}, \sigma' \in \mathcal{L}_{V_{\tau}}} H(\sigma, \sigma') M_{\sigma} M_{\tau} M^T_{\sigma'}$$

Proposition 2.73. *For all A and B with shapes in $\mathcal{I}(\Lambda)$,*

$$\left(M_{\tau}^{fact}(H)\right)(A,B) = \sum_{\sigma\in\mathcal{L}_{U_\tau},\sigma'\in\mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1\in\mathcal{R}(\sigma,A,A'),R_2\in\mathcal{R}(\tau,A',B'),\\ R_3\in\mathcal{R}({\sigma'}^T,B',B)}} M_{R_1}(A,A')M_{R_2}(A',B')M_{R_3}(B',B)$$

If R_1, R_2, R_3 are properly composable then $R = R_1 \circ R_2 \circ R_3$ has the expected shape $\sigma \circ \tau \circ {\sigma'}^T$. Otherwise, $R_1 \circ R_2 \circ R_3$ will have a different shape. We define $M_{\tau}^{orth}(H)$ to be the same sum as $M_{\tau}^{fact}(H)$ except that it is restricted to properly composable ribbons R_1, R_2, R_3.

Definition 2.74. *We define $M_{\tau}^{orth}(H)$ so that for all A and B with shapes in $\mathcal{I}(\Lambda)$,*

$$\begin{aligned}
&\left(M_{\tau}^{orth}(H)\right)(A,B) \\
&= \sum_{\sigma\in\mathcal{L}_{U_\tau},\sigma'\in\mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1\in\mathcal{R}(\sigma,A,A'),R_2\in\mathcal{R}(\tau,A',B'),\\ R_3\in\mathcal{R}({\sigma'}^T,B',B),R_1,R_2,R_3 \text{ are properly composable}}} M_{R_1}(A,A')M_{R_2}(A',B')M_{R_3}(B',B) \\
&= \sum_{\sigma\in\mathcal{L}_{U_\tau},\sigma'\in\mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1\in\mathcal{R}(\sigma,A,A'),R_2\in\mathcal{R}(\tau,A',B'),\\ R_3\in\mathcal{R}({\sigma'}^T,B',B),R_1,R_2,R_3 \text{ are properly composable}}} M_{R_1\circ R_2\circ R_3}(A,B)
\end{aligned}$$

It would be nice if we had that $M_{\tau}^{orth}(H) = \sum_{\sigma\in\mathcal{R}_{U_\tau},\sigma'\in\mathcal{R}_{V_\tau}} H(\sigma,\sigma')M_{\sigma\circ\tau\circ{\sigma'}^T}$. However, this is not quite correct because there is an additional term related to automorphism groups.

Definition 2.75. *Given a shape α, define $Aut(\alpha)$ to be the set of mappings from α to itself which keep U_α and V_α fixed.*

Definition 2.76. *Given composable shapes $\sigma, \tau, {\sigma'}^T$, we define*

$$Decomp(\sigma,\tau,\sigma') = Aut(\sigma\circ\tau\circ\sigma')/(Aut(\sigma)\times Aut(\tau)\times Aut({\sigma'}^T))$$

Remark 2.77. *Each element $\pi \in Decomp(\sigma,\tau,\sigma')$ decomposes $\sigma\circ\tau\circ{\sigma'}^T$ into σ, τ, and ${\sigma'}^T$ by specifying copies $\pi(\sigma)$, $\pi(\tau)$, $\pi({\sigma'}^T)$ of σ, τ, and ${\sigma'}^T$ such that $\pi(\sigma)\circ\pi(\tau)\circ\pi({\sigma'}^T) = \pi(\sigma\circ\tau\circ{\sigma'}^T) = \sigma\circ\tau\circ{\sigma'}^T$. Thus, $|Decomp(\sigma,\tau,\sigma')|$ is the number of ways to decompose $\sigma\circ\tau\circ{\sigma'}^T$ into σ, τ, and ${\sigma'}^T$.*

Lemma 2.78.

$$M_{\tau}^{orth}(H) = \sum_{\sigma\in\mathcal{L}_{U_\tau},\sigma'\in\mathcal{L}_{V_\tau}} H(\sigma,\sigma')|Decomp(\sigma,\tau,{\sigma'}^T)|M_{\sigma\circ\tau\circ{\sigma'}^T}$$

Proof sketch. Observe that there is a bijection between ribbons R with shape $\sigma\circ\tau\circ{\sigma'}^T$ together with an element $\pi \in Decomp(\sigma,\tau,\sigma')$ and triples of ribbons (R_1, R_2, R_3) such that

1. R_1, R_2, R_3 have shapes σ, τ, and ${\sigma'}^T$, respectively.
2. $V(R_1)\cap V(R_2) = A_{R_2} = B_{R_1}$, $V(R_2)\cap V(R_3) = A_{R_3} = B_{R_2}$, and $V(R_1)\cap V(R_3) = A_{R_2}\cap B_{R_2}$

To see this, note that given such ribbons R_1, R_2, R_3, the ribbon $R = R_1 \circ R_2 \circ R_3$ has shape $\sigma \circ \tau \circ \sigma'^T$ and the ribbons R_1, R_2, R_3 specify a decomposition of $\sigma \circ \tau \circ \sigma'^T$ into σ, τ, and σ'^T.

Conversely, given R and an element $\pi \in Decomp(\sigma, \tau, \sigma')$, π specifies how to decompose R into ribbons R_1, R_2, R_3 of shapes σ, τ, and σ'^T.

For a more rigorous proof, see Appendix B. □

Remark 2.79. *As this lemma shows, we have to be very careful about symmetry groups in our analysis. For accuracy, it is safest to check that the coefficients for each individual ribbon match.*

Given a matrix-valued function Λ, we can associate coefficient matrices to Λ as follows:

Definition 2.80. *Given a matrix-valued function* $\Lambda = \sum_{\alpha:\alpha \text{ is proper}} \lambda_\alpha M_\alpha$*,*

1. *For each index shape* $U \in \mathcal{I}_{mid}$ *and every* $\sigma, \sigma' \in \mathcal{L}_U$*, we take* $H_{Id_U}(\sigma, \sigma') = \frac{1}{|Aut(U)|}\lambda_{\sigma \circ \sigma'^T}$
2. *For each* $U \in \mathcal{I}_{mid}$*,* $\tau \in \mathcal{M}_U$ *and* $\sigma, \sigma' \in \mathcal{L}_U$*, we take* $H_\tau(\sigma, \sigma') = \frac{1}{|Aut(U_\tau)| \cdot |Aut(V_\tau)|}\lambda_{\sigma \circ \tau \circ \sigma'^T}$

Lemma 2.81. $\Lambda = \sum_{U \in \mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{orth}_\tau(H_\tau)$

Proof. We check that the coefficients for each individual ribbon R match. There are two cases to consider.

If R has shape α where α has a unique minimum vertex separator S, then there is a bijection between orderings O_S for S and pairs of ribbons R_1, R_2 such that $R_1 \circ R_2 = R$ and the shapes σ, σ'^T of R_1, R_2 are left and right shapes respectively.

To see this, observe that when we concatenate R_1 and R_2, this assigns the matrix index $B_{R_1} = A_{R_2}$ to S, which is equivalent to specifying an ordering O_S for S. Conversely, given an ordering O_S for S, we take R_1 to be the part of R between A_R and (S, O_S) and we take R_2 to be the part of R between (S, O_S) and B_R.

From this bijection, it follows that the coefficient of M_R is λ_α on both sides of the equation.

Similarly, if R has shape α where α does not have a unique minimal vertex separator, then there is a bijection between orderings O_S, O_T for the leftmost and rightmost minimum vertex separators S, T of R and triples of ribbons R_1, R_2, R_3 such that $R_1 \circ R_2 \circ R_3 = R$ and the shapes σ, τ, σ'^T of R_1, R_2, R_3 are left, proper middle, and right shapes respectively.

To see this, observe that when we concatenate R_1, R_2, and R_3, this assigns the matrix index $B_{R_1} = A_{R_2}$ to S and assigns the matrix index $B_{R_2} = A_{R_3}$ to T, which is equivalent to specifying orderings O_S, O_T for S, T. Conversely, given orderings O_S, O_T for S, T, we take R_1 to be the part of R between A_R and (S, O_S), we take R_2 to be the part of R between (S, O_S) and (T, O_T), and we take R_2 to be the part of R between (T, O_T) and B_R.

From this bijection, it again follows that the coefficient of M_R is λ_α on both sides of the equation. □

2.9 The $-\gamma, -\gamma$ operation and qualitative theorem statement

In the intersection term anaylsis (see subsections 3.2, 3.3, and 3.4), we will need to further decompose left shapes σ as $\sigma = \sigma_2 \circ \gamma$ where σ_2 and γ are themselves left shapes. Accordingly, we make the following definitions

Definition 2.82. *Given a moment matrix Λ, we define the following sets of left shapes:*

1. $\Gamma = \{\gamma : \gamma$ *is a non-trivial left shape*$, U_\gamma, V_\gamma \in \mathcal{I}_{mid}, |V(\gamma)| \leq D_V, \forall e \in E(\gamma), l_e \leq D_E\}$
2. *Given $U, V \in \mathcal{I}_{mid}$ such that $w(U) > w(V)$, define $\Gamma_{U,V} = \{\gamma \in \Gamma : U_\gamma \equiv U, V_\gamma \equiv V\}$.*
3. *Given $U \in \mathcal{I}_{mid}$, define $\Gamma_{U,*} = \{\gamma \in \Gamma : U_\gamma \equiv U\}$*
4. *Given $V \in \mathcal{I}_{mid}$, define $\Gamma_{*,V} = \{\gamma \in \Gamma : V_\gamma \equiv V\}$*

Remark 2.83. *Under our simplifying assumptions, Γ is the same as $\mathcal{L}$ except that Γ excludes the trivial shapes. In general, while $\mathcal{L}$ requires that $U_\sigma \in \mathcal{I}(\Lambda)$, Γ requires that $U_\gamma \in \mathcal{I}_{mid}$. Note that $\mathcal{I}(\Lambda)$ and $\mathcal{I}_{mid}$ may be incomparable because*

1. *There may be index shapes $U \in \mathcal{I}_{mid}$ such that no matrix index of Λ has shape U.*
2. *All index shape pieces U_i for index shapes $U \in \mathcal{I}_{mid}$ must have $p_i = 1$ while this is not the case for $\mathcal{I}(\Lambda)$.*

We now state our theorem qualitatively after giving one more definition.

Definition 2.84. *Given a shape τ, left shapes $\gamma \in \Gamma_{*,U_\tau}$ and $\gamma' \in \Gamma_{*,V_\tau}$, and a τ-coefficient matrix H, define $H^{-\gamma,\gamma'}$ to be the $(\gamma \circ \tau \circ \gamma'^T)$-coefficient matrix with entries*

1. $H^{-\gamma,\gamma'}(\sigma, \sigma') = H(\sigma \circ \gamma, \sigma' \circ \gamma')$ *if* $|V(\sigma \circ \gamma)| \leq D_V$ *and* $|V(\sigma' \circ \gamma')| \leq D_V$.
2. $H^{-\gamma,\gamma'}(\sigma, \sigma') = 0$ *if* $|V(\sigma \circ \gamma)| > D_V$ *or* $|V(\sigma' \circ \gamma')| > D_V$.

Remark 2.85. *For the theorem, we will only need the case when $\gamma' = \gamma$*

Our qualitative theorem statement is as follows:

Theorem 2.86. *Let $\Lambda = \sum_{U \in \mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{orth}_\tau(H_\tau)$ be an SOS-symmetric matrix valued function.*

There exist functions $f(\tau)$ and $f(\gamma)$ depending on n and other parameters such that if the following conditions hold:

1. *For all $U \in \mathcal{I}_{mid}$, $H_{Id_U} \succeq 0$*
2. *For all $U \in \mathcal{I}_{mid}$ and all $\tau \in \mathcal{M}_U$,*

$$\begin{bmatrix} H_{Id_U} & f(\tau)H_\tau \\ f(\tau)H_\tau^T & H_{Id_U} \end{bmatrix} \succeq 0$$

3. *For all $U, V \in \mathcal{I}_{mid}$ where $w(U) > w(V)$ and all $\gamma \in \Gamma_{U,V}$, $H^{-\gamma,\gamma}_{Id_V} \preceq f(\gamma) H_{Id_U}$*

then with high probability $\Lambda \succeq 0$

Remark 2.87. *Roughly speaking, conditions 1 and 2 give us an approximate PSD decomposition for the moment matrix M. Condition 3 comes from the intersection term analysis, which is the most technically intensive part of the proof.*

2.10 Quantitative theorem statement

To state our theorem quantitatively, we will need a few more things. First, the conditions of the theorem will involve functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$. Roughly speaking, these functions will be used as follows in the analysis:

1. $B_{norm}(\alpha)$ will bound the norms of the matrices M_α
2. $B(\gamma)$ and $N(\gamma)$ will help us bound the intersection terms (see Section 3.4).
3. $c(\alpha)$ will help us sum over the possible γ and τ.

Second, for technical reasons it turns out that comparing $H^{-\gamma,\gamma}_{Id_{V_\gamma}}$ to $H_{Id_{U_\gamma}}$ doesn't quite work. Instead, we comapre $H^{-\gamma,\gamma}_{Id_{V_\gamma}}$ to a matrix H'_γ of our choice where H'_γ is very close to H_{U_γ} (H'_γ will be the same as H_{U_γ} up to truncation error).

Definition 2.88. *Given a function $B_{norm}(\alpha)$, we define the distance $d_\tau(H_\tau, H'_\tau)$ between two τ-coefficient matrices H_τ and H'_τ to be*

$$d_\tau(H_\tau, H'_\tau) = \sum_{\sigma \in \mathcal{L}_{U_\tau}, \sigma' \in \mathcal{L}_{V_\tau}} |H'_\tau(\sigma,\sigma') - H_\tau(\sigma,\sigma')| B_{norm}(\sigma) B_{norm}(\tau) B_{norm}(\sigma')$$

Third, we need an SOS-symmetric analogue of the identity matrix.

Definition 2.89. *We define Id_{Sym} to be the matrix such that*

1. *The rows and columns of Id_{Sym} are indexed by the matrix indices A, B whose index shape is in $\mathcal{I}(\Lambda)$.*
2. *$Id_{Sym}(A,B) = 1$ if $p_A = p_B$ and $Id_{Sym}(A.B) = 0$ if $p_A \neq p_B$.*

Proposition 2.90. *If M has SOS-symmetry and the rows and columns of Id_{Sym} are indexed by matrix indices A, B whose index shape is in $\mathcal{I}(\Lambda)$ then $M \preceq ||M|| Id_{Sym}$*

Corollary 2.91. *For all τ and all SOS-symmetric τ-coefficient matrices H_τ and H'_τ,*

$$M^{fact}_\tau(H'_\tau) + M^{fact}_{\tau^T}(H'_{\tau^T}) - M^{fact}_\tau(H_\tau) - M^{fact}_{\tau^T}(H_{\tau^T}) \preceq 2 d_\tau(H_\tau, H'_\tau) Id_{Sym}$$

Note that if τ, H_τ and H'_τ are all symmetric then

$$M^{fact}_\tau(H'_\tau) - M^{fact}_\tau(H_\tau) \preceq d_\tau(H_\tau, H'_\tau) Id_{Sym}$$

Finally, we need a few more definitions about shapes α.

Definition 2.92 ($\mathcal{M}'$)**.** *We define $\mathcal{M}'$ to be the set of all shapes α such that*

1. *$|V(\alpha)| \leq 3D_V$*

2.* *$\forall e \in E(\alpha), l_e \leq D_E$*

3.* *All edges $e \in E(\alpha)$ have multiplicity at most $3D_V$.*

Definition 2.93 (S_α)**.** *Given a shape α, define S_α to be the leftmost minimum vertex separator of α*

Definition 2.94 (I_α)**.** *Given a shape α, define I_α to be the set of vertices in $V(\alpha)\setminus(U_\alpha \cup V_\alpha)$ which are isolated.*

We can now state our main theorem.

Theorem 2.95. *Given the moment matrix $\Lambda = \sum_{U\in\mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U\in\mathcal{I}_{mid}}\sum_{\tau\in\mathcal{M}_U} M^{orth}_{\tau}(H_\tau)$, for all $\epsilon > 0$, if we take*

1. $q = 3\left\lceil D_V ln(n) + \frac{ln(\frac{1}{\epsilon})}{3} + D_V ln(5) + 3D_V^2 ln(2)\right\rceil$
2. $B_{vertex} = 6D_V\sqrt[4]{2eq}$
3. $B_{norm}(\alpha) = B_{vertex}^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha))+w(I_\alpha)-w(S_\alpha)}{2}}$
4. $B(\gamma) = B_{vertex}^{|V(\gamma)\setminus U_\gamma|+|V(\gamma)\setminus V_\gamma|} n^{\frac{w(V(\gamma)\setminus U_\gamma)}{2}}$
5. $N(\gamma) = (3D_V)^{2|V(\gamma)\setminus V_\gamma|+|V(\gamma)\setminus U_\gamma|}$
6. $c(\alpha) = \frac{1}{100(3D_V)^{|U_\alpha\setminus V_\alpha|+|V_\alpha\setminus U_\alpha|+2|E(\alpha)|}2^{|V(\alpha)\setminus(U_\alpha\cup V_\alpha)|}}$

and we have SOS-symmetric coefficient matrices $\{H'_\gamma : \gamma\in\Gamma\}$ such that the following conditions hold:

1. *For all $U\in\mathcal{I}_{mid}$, $H_{Id_U}\succeq 0$*
2. *For all $U\in\mathcal{I}_{mid}$ and $\tau\in\mathcal{M}_U$,*
$$\begin{bmatrix} \frac{1}{|Aut(U)|c(\tau)}H_{Id_U} & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H_\tau^T & \frac{1}{|Aut(U)|c(\tau)}H_{Id_U}\end{bmatrix}\succeq 0$$
3. *For all $U,V\in\mathcal{I}_{mid}$ where $w(U)>w(V)$ and all $\gamma\in\Gamma_{U,V}$,*
$$c(\gamma)^2N(\gamma)^2B(\gamma)^2H^{-\gamma,\gamma}_{Id_V}\preceq H'_\gamma$$

then with probability at least $1-\epsilon$,

$$\Lambda\succeq\frac{1}{2}\left(\sum_{U\in\mathcal{I}_{mid}}M^{fact}_{Id_U}(H_{Id_U})\right) - 3\left(\sum_{U\in\mathcal{I}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_U}(H'_\gamma,H_{Id_U})}{|Aut(U)|c(\gamma)}\right)Id_{sym}$$

If it is also true that whenever $||M_\alpha||\leq B_{norm}(\alpha)$ for all $\alpha\in\mathcal{M}'$,

$$\sum_{U\in\mathcal{I}_{mid}}M^{fact}_{Id_U}(H_{Id_U})\succeq 6\left(\sum_{U\in\mathcal{I}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_U}(H'_\gamma,H_{Id_U})}{|Aut(U)|c(\gamma)}\right)Id_{sym}$$

then with probability at least $1-\epsilon$, $\Lambda\succeq 0$.

2.10.1 General Main Theorem

Before stating the general main theorem, we need to modify a few definitions for α and give a few definitions for Ω

Definition 2.96 ($S_{\alpha,min}$ and $S_{\alpha,max}$)**.** *Given a shape $\alpha \in \mathcal{M}'$, define $S_{\alpha,min}$ to be the leftmost minimum vertex separator of α if all edges with multiplicity at least 2 are deleted and define $S_{\alpha,max}$ to be the leftmost minimum vertex separator of α if all edges with multiplicity at least 2 are present.*

Definition 2.97 (General I_α)**.** *Given a shape α, define I_α to be the set of vertices in $V(\alpha)\setminus(U_\alpha \cup V_\alpha)$ such that all edges incident with that vertex have multplicity at least 2.*

Definition 2.98 (B_Ω)**.** *We take $B_\Omega(j)$ to be a non-decreasing function such that for all $j \in \mathbb{N}$, $E_\Omega[x^j] \leq B_\Omega(j)^j$*

Definition 2.99 (h_j^+)**.** *For all j, we define h_j^+ to be the polynomial h_j where we make all of the coefficients have positive sign.*

Lemma 2.100. *If $\Omega = N(0,1)$ then we can take $B_\Omega(j) = \sqrt{j}$ and we have that*

$$h_j^+(x) \leq \frac{1}{\sqrt{j!}}(x^2+j)^{\frac{j}{2}} \leq \left(\frac{e}{j}(x^2+j)\right)^{\frac{j}{2}}$$

Proof. To be added. □

Theorem 2.101. *Given the moment matrix $\Lambda = \sum_{U \in \mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{orth}_{\tau}(H_\tau)$, for all $\epsilon > 0$, if we take*

1. $q = \lceil 3D_V ln(n) + ln(\frac{1}{\epsilon}) + (3D_V)^k ln(D_E+1) + 3D_V ln(5) \rceil$
2. $B_{vertex} = 6qD_V$
3. $B_{edge}(e) = 2h^+_{l_e}(B_\Omega(6D_V D_E)) \max_{j \in [0, 3D_V D_E]} \left\{ \left(h_j^+(B_\Omega(2qj))\right)^{\frac{l_e}{\max\{j, l_e\}}} \right\}$

   As a special case, if $\Omega = N(0,1)$ then we can take $B_{edge}(e) = \left(400 D_V^2 D_E^2 q\right)^{l_e}$
4. $B_{norm}(\alpha) = 2eB_{vertex}^{|V(\alpha)\setminus U_\alpha| + |V(\alpha)\setminus V_\alpha|} \left(\prod_{e \in E(\alpha)} B_{edge}(e)\right) n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_{\alpha,min})}{2}}$
5. $B(\gamma) = B_{vertex}^{|V(\gamma)\setminus U_\gamma| + |V(\gamma)\setminus V_\gamma|} \left(\prod_{e \in E(\gamma)} B_{edge}(e)\right) n^{\frac{w(V(\gamma)\setminus U_\gamma)}{2}}$
6. $N(\gamma) = (3D_V)^{2|V(\gamma)\setminus V_\gamma| + |V(\gamma)\setminus U_\gamma|}$
7. $c(\alpha) = \frac{1}{100(3t_{max}D_V)^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + k|E(\alpha)|}(2t_{max})^{|V(\alpha)\setminus(U_\alpha \cup V_\alpha)|}}$

and we have SOS-symmetric coefficient matrices $\{H'_\gamma : \gamma \in \Gamma\}$ such that the following conditions hold:

1. *For all $U \in \mathcal{I}_{mid}$, $H_{Id_U} \succeq 0$*

2. *For all $U \in \mathcal{I}_{mid}$ and $\tau \in \mathcal{M}_U$,*

$$\begin{bmatrix} \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} & B_{norm}(\tau) H_\tau \\ B_{norm}(\tau) H_\tau^T & \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} \end{bmatrix} \succeq 0$$

3. *For all $U, V \in \mathcal{I}_{mid}$ where $w(U) > w(V)$ and all $\gamma \in \Gamma_{U,V}$,*

$$c(\gamma)^2 N(\gamma)^2 B(\gamma)^2 H_{Id_V}^{-\gamma,\gamma} \preceq H'_\gamma$$

then with probability at least $1 - \epsilon$,

$$\Lambda \succeq \frac{1}{2} \left(\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) \right) - 3 \left(\sum_{U \in \mathcal{I}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)} \right) Id_{sym}$$

If it is also true that whenever $||M_\alpha|| \leq B_{norm}(\alpha)$ for all $\alpha \in \mathcal{M}'$,

$$\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) \succeq 6 \left(\sum_{U \in \mathcal{I}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)} \right) Id_{sym}$$

then with probability at least $1 - \epsilon$, $\Lambda \succeq 0$.

2.11 Choosing H'_γ and Truncation Error

A canonical choice for H'_γ is to take

1. $H'_\gamma(\sigma, \sigma') = H_{V_\gamma}(\sigma \circ \gamma, \sigma' \circ \gamma)$ whenever $|V(\sigma \circ \gamma)| \leq D_V$ and $|V(\sigma' \circ \gamma)| \leq D_V$.
2. $H'_\gamma(\sigma, \sigma') = 0$ whenever $|V(\sigma \circ \gamma)| > D_V$ or $|V(\sigma' \circ \gamma)| > D_V$.

With this choice, the truncation error is

$$d_{Id_{U_\gamma}}(H_{Id_{U_\gamma}}, H'_\gamma) = \sum_{\substack{\sigma, \sigma' \in \mathcal{L}_{U_\gamma}: V(\sigma) \leq D_V, V(\sigma') \leq D_V, \\ |V(\sigma \circ \gamma)| > D_V \text{ or } |V(\sigma' \circ \gamma)| > D_V}} B_{norm}(\sigma) B_{norm}(\sigma') H_{Id_{U_\gamma}}(\sigma, \sigma')$$

3 Proof of the Main Theorem

In this section, we prove the main theorem under the assumption that the functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$ have certain properties. More precisely, we prove the following theorem.

Theorem 3.1. *For all $\epsilon > 0$ and all $\epsilon' \in (0, \frac{1}{20}]$, for any moment matrix*

$$\Lambda = \sum_{U \in \mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{orth}_\tau(H_\tau),$$

if $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$ are functions such that

1. *With probability at least* $(1-\epsilon)$, *for all shapes* $\alpha \in \mathcal{M}'$, $||M_\alpha|| \le B_{norm}(\alpha)$.

2. *For all* $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau}$, $\gamma' \in \Gamma_{*,V_\tau}$, *and all intersection patterns* $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,

$$B_{norm}(\tau_P) \le B(\gamma)B(\gamma')B_{norm}(\tau)$$

   Note: Intersection patterns and $\mathcal{P}_{\gamma,\tau,\gamma'}$ *will be defined later, see Definitions 3.8 and 3.9.*

3. *For all composable* γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.

4. $\forall U \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{U,*}} \frac{1}{|Aut(U)|c(\gamma)} < \epsilon'$

5. $\forall V \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)|c(\gamma)} < \epsilon'$

6. $\forall U \in \mathcal{I}_{mid}, \sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)|c(\tau)} < \epsilon'$

7. *For all* $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, *and* $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$,

$$\sum_{j>0} \sum_{\gamma_1,\gamma'_1,\cdots,\gamma_j,\gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1,\cdots,P_j: P_i \in \mathcal{P}_{\gamma_i,\tau_{P_{i-1}},\gamma'^T_i}} \left(\prod_{i=1}^{j} N(P_i)\right)$$
$$\le \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

   Note: $\Gamma_{\gamma,\gamma',j}$ *will be defined later, see Definition 3.18.*

and we have SOS-symmetric coefficient matrices $\{H'_\gamma : \gamma \in \Gamma\}$ *such that the following conditions hold:*

1. *For all* $U \in \mathcal{I}_{mid}$, $H_{Id_U} \succeq 0$

2. *For all* $U \in \mathcal{I}_{mid}$ *and* $\tau \in \mathcal{M}_U$,

$$\begin{bmatrix} \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H_\tau^T & \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} \end{bmatrix} \succeq 0$$

3. *For all* $U, V \in \mathcal{I}_{mid}$ *where* $w(U) > w(V)$ *and all* $\gamma \in \Gamma_{U,V}$,

$$c(\gamma)^2 N(\gamma)^2 B(\gamma)^2 H_{Id_V}^{-\gamma,\gamma} \preceq H'_\gamma$$

then with probability at least $1-\epsilon$,

$$\Lambda \succeq \frac{1}{2}\left(\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U})\right) - 3\left(\sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)}\right) Id_{sym}$$

If it is also true that whenever $||M_\alpha|| \le B_{norm}(\alpha)$ *for all* $\alpha \in \mathcal{M}'$,

$$\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) \succeq 6\left(\sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)}\right) Id_{sym}$$

then with probability at least $1-\epsilon$, $\Lambda \succeq 0$.

Throughout this section, we assume that we have functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$. If $\forall \alpha \in \mathcal{M}', ||M_\alpha|| \leq B_{norm}(\alpha)$ then we say that the norm bounds hold. For the other properties of these functions, we will either restate these properties in our intermediate results to highlight where these properties are needed or just state that the conditions on these functions are satisfied for brevity.

3.1 Warm-up: Analysis with no intersection terms

In this subsection, we show how the analysis works if we ignore the difference between M^{fact} and M^{orth}

Theorem 3.2. *For all $\epsilon' \in (0, \frac{1}{2}]$, if the norm bounds hold and the following conditions hold*

1. *For all $U \in \mathcal{I}_{mid}$, $H_{Id_U} \succeq 0$*
2. *For all $U \in \mathcal{I}_{mid}$ and all $\tau \in \mathcal{M}_U$*

$$\begin{bmatrix} \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} & B_{norm}(\tau) H_\tau \\ B_{norm}(\tau) H_\tau^T & \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} \end{bmatrix} \succeq 0$$

3. *$\forall U \in \mathcal{I}_{mid}, \sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)|c(\tau)} \leq \epsilon'$.*

then

$$\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{fact}_\tau (H_\tau) \succeq (1 - 2\epsilon') \sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) \succeq 0$$

Proof. We first show how a single term $M_\sigma M_\tau M_{\sigma'^T}$ plus its transpose $M_{\sigma'} M_{\tau^T} M_{\sigma^T}$ can be bounded.

Lemma 3.3. *If the norm bounds hold then for all $\tau \in \mathcal{M}'$ and shapes σ, σ' such that σ, τ, σ'^T are composable, for all a, b such that $a > 0$, $b > 0$, and $ab = B_{norm}(\tau)^2$,*

$$M_\sigma M_\tau M_{\sigma'^T} + M_{\sigma'} M_{\tau^T} M_{\sigma^T} \preceq a M_\sigma M_{\sigma^T} + b M_{\sigma'} M_{\sigma'^T}$$

Proof. Observe that

$$0 \preceq \left(\sqrt{a} M_\sigma - \frac{\sqrt{b}}{B_{norm}(\tau)} M_{\sigma'} M_{\tau^T}\right) \left(\sqrt{a} M_\sigma - \frac{\sqrt{b}}{B_{norm}(\tau)} M_{\sigma'} M_{\tau^T}\right)^T =$$
$$\left(\sqrt{a} M_\sigma - \frac{\sqrt{b}}{B_{norm}(\tau)} M_{\sigma'} M_\tau\right) \left(\sqrt{a} M_{\sigma^T} - \frac{\sqrt{b}}{B_{norm}(\tau)} M_\tau M_{\sigma'^T}\right) =$$
$$a M_\sigma M_{\sigma^T} - M_\sigma M_\tau M_{\sigma'^T} - M_{\sigma'} M_{\tau^T} M_{\sigma^T} + \frac{b}{B_{norm}(\tau)^2} M_{\sigma'} M_{\tau^T} M_\tau M_{\sigma'^T} \preceq$$
$$a M_\sigma M_{\sigma^T} - M_\sigma M_\tau M_{\sigma'^T} - M_{\sigma'} M_{\tau^T} M_{\sigma^T} + \frac{b}{B_{norm}(\tau)^2} M_{\sigma'} (B_{norm}(\tau)^2 Id) M_{\sigma'^T}$$

Thus, $M_\sigma M_\tau M_{\sigma'^T} + M_{\sigma'} M_{\tau^T} M_{\sigma^T} \preceq a M_\sigma M_{\sigma^T} + b M_{\sigma'} M_{\sigma'^T}$, as needed. $\square$

Unfortunately, if we try to bound everything term by term, there may be too many terms to bound. Instead, we generalize this argument for vectors and coefficient matrices.

Definition 3.4. *Let τ be a shape. We say that a vector v is a left τ-vector if the coordinates of v are indexed by left shapes $\sigma \in \mathcal{L}_{U_\tau}$. We say that a vector w is a right τ-vector if the coordinates of w are indexed by left shapes $\sigma' \in \mathcal{L}_{V_\tau}$.*

Lemma 3.5. *For all $\tau \in \mathcal{M}'$, if the norm bounds hold, v is a left τ-vector, and w is a right τ-vector then*

$$M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T) \preceq B_{norm}(\tau)\left(M^{fact}_{Id_{U_\tau}}(vv^T) + M^{fact}_{Id_{V_\tau}}(ww^T)\right)$$

and

$$-M^{fact}_{\tau}(vw^T) - M^{fact}_{\tau^T}(wv^T) \preceq B_{norm}(\tau)\left(M^{fact}_{Id_{U_\tau}}(vv^T) + M^{fact}_{Id_{V_\tau}}(ww^T)\right)$$

Proof. observe that

$$\begin{aligned}
0 \preceq &\left(\sum_{\sigma} v_\sigma M_\sigma \mp \frac{w_\sigma M_\sigma M_{\tau^T}}{B_{norm}(\tau)}\right)\left(\sum_{\sigma'} v_{\sigma'} M_{\sigma'} \mp \frac{w_{\sigma'} M_{\sigma'} M_{\tau^T}}{B_{norm}(\tau)}\right)^T = \\
&\left(\sum_{\sigma} v_\sigma M_\sigma \mp \frac{w_\sigma M_\sigma M_{\tau^T}}{B_{norm}(\tau)}\right)\left(\sum_{\sigma'} v_{\sigma'} M_{\sigma'^T} \mp \frac{w_{\sigma'} M_{\tau} M_{\sigma'^T}}{B_{norm}(\tau)}\right) = \\
&\sum_{\sigma,\sigma'} (v_\sigma v_{\sigma'}) M_\sigma M_{\sigma'^T} \mp \sum_{\sigma,\sigma'} \frac{(v_\sigma w_{\sigma'})}{B_{norm}(\tau)} M_\sigma M_\tau M_{\sigma'} \\
&\mp \sum_{\sigma,\sigma'} \frac{(w_\sigma v_{\sigma'})}{B_{norm}(\tau)} M_\sigma M_{\tau^T} M_{\sigma'} + \frac{1}{B_{norm}(\tau)^2} \sum_{\sigma,\sigma'} (v_\sigma v_{\sigma'}) M_\sigma M_\tau M_{\tau^T} M_{\sigma'^T}
\end{aligned}$$

Further observe that

1. $\sum_{\sigma,\sigma'} (v_\sigma v_{\sigma'}) M_\sigma M_{\sigma'^T} = M^{fact}_{Id_{U_\tau}}(vv^T)$
2. $\sum_{\sigma,\sigma'} (v_\sigma w_{\sigma'}) M_\sigma M_\tau M_{\sigma'^T} = M^{fact}_{\tau}(vw^T)$
3. $\sum_{\sigma,\sigma'} (w_\sigma v_{\sigma'}) M_\sigma M_{\tau^T} M_{\sigma'^T} = M^{fact}_{\tau^T}(wv^T)$
4.

$$\begin{aligned}
\sum_{\sigma,\sigma'} (w_\sigma w_{\sigma'}) M_\sigma M_\tau M_{\tau^T} M_{\sigma'^T} &= \left(\sum_{\sigma} w_\sigma M_\sigma\right) M_\tau M_{\tau^T} \left(\sum_{\sigma} w_\sigma M_\sigma\right)^T \\
&\preceq \left(\sum_{\sigma} w_\sigma M_\sigma\right) B_{norm}(\tau)^2 Id \left(\sum_{\sigma} w_\sigma M_\sigma\right)^T \\
&= B_{norm}(\tau)^2 \sum_{\sigma,\sigma'} (w_\sigma w_{\sigma'}) M_\sigma M_{\sigma'^T} \\
&= B_{norm}(\tau)^2 M^{fact}_{Id_{V_\tau}}(ww^T)
\end{aligned}$$

Putting everything together,

$$\frac{M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T)}{B_{norm}(\tau)} \preceq M^{fact}_{Id_{U_\tau}}(vv^T) + M^{fact}_{Id_{V_\tau}}(ww^T)$$

and

$$-\frac{M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T)}{B_{norm}(\tau)} \preceq M^{fact}_{Id_{U_\tau}}(vv^T) + M^{fact}_{Id_{V_\tau}}(ww^T)$$

as needed. □

Corollary 3.6. *For all $\tau \in \mathcal{M}'$, if the norm bounds hold and H_U and H_V are matrices such that*

$$\begin{bmatrix} H_U & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H^T_\tau & H_V \end{bmatrix} \succeq 0$$

then $M^{fact}_{\tau}(H_\tau) + M^{fact}_{\tau^T}(H_{\tau^T}) \preceq M^{fact}_{Id_{U_\tau}}(H_U) + M^{fact}_{Id_{V_\tau}}(H_V)$

Proof. If $\begin{bmatrix} H_U & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H^T_\tau & H_V \end{bmatrix} \succeq 0$ then we can write

$$\begin{bmatrix} H_U & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H^T_\tau & H_V \end{bmatrix} = \sum_i (v_i, w_i)(v_i, w_i)^T$$

Since the M^{fact} operations are linear, the result now follows by summing the equation

$$M^{fact}_{\tau}(v_i w_i^T) + M^{fact}_{\tau^T}(w_i v_i^T) \preceq B_{norm}(\tau)\left(M^{fact}_{Id_{U_\tau}}(v_i v_i^T) + M^{fact}_{Id_{V_\tau}}(w_i w_i^T)\right)$$

over all i. □

Theorem 3.2 now follows directly. For all $U \in \mathcal{I}_{mid}$ and all $\tau \in \mathcal{M}_U$, using Corollary 3.6 with $H_U = H_V = \frac{1}{|Aut(U)|c(\tau)} H_{Id_U}$,

$$M^{fact}_{\tau}(H_\tau) + M^{fact}_{\tau^T}(H_{\tau^T}) \preceq \frac{1}{|Aut(U)|c(\tau)} M^{fact}_{Id_U}(H_{Id_U}) + \frac{1}{|Aut(U)|c(\tau)} M^{fact}_{Id_U}(H_{Id_U})$$

Summing this equation over all $U \in \mathcal{I}_{mid}$ and all $\tau \in \mathcal{M}_U$, we obtain that

$$\sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{fact}_{\tau}(H_\tau) \preceq 2\epsilon' \sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U})$$

as needed. □

3.2 Intersection Term Analysis Strategy

As we saw in the previous subsection, the analysis works out nicely if we work with M^{fact}. Unfortunately, our matrices are expressed in terms of M^{orth}. In this subsection, we describe our strategy for analyzing the difference between M^{fact} and M^{orth}.

Recall the following expressions for $\left(M_{\tau}^{fact}(H)\right)(A,B)$ and $\left(M_{\tau}^{orth}(H)\right)(A,B)$ where A has shape U_τ and B has shape V_τ:

$$\left(M_{\tau}^{fact}(H)\right)(A,B) = \sum_{\sigma \in \mathcal{L}_{U_\tau}, \sigma' \in \mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1 \in \mathcal{R}(\sigma,A,A'), R_2 \in \mathcal{R}(\tau,A',B'), \\ R_3 \in \mathcal{R}({\sigma'}^T,B',B)}} M_{R_1}(A,A')M_{R_2}(A',B')M_{R_3}(B',B)$$

$$\begin{aligned}&\left(M_{\tau}^{orth}(H)\right)(A,B) \\ &= \sum_{\sigma \in \mathcal{L}_{U_\tau}, \sigma' \in \mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1 \in \mathcal{R}(\sigma,A,A'), R_2 \in \mathcal{R}(\tau,A',B'), \\ R_3 \in \mathcal{R}({\sigma'}^T,B',B), R_1,R_2,R_3 \text{ are properly composable}}} M_{R_1}(A,A')M_{R_2}(A',B')M_{R_3}(B',B)\end{aligned}$$

This implies that $\left(M_{\tau}^{fact}(H)\right)(A,B) - \left(M_{\tau}^{orth}(H)\right)(A,B)$ is equal to

$$\sum_{\sigma \in \mathcal{L}_{U_\tau}, \sigma' \in \mathcal{L}_{V_\tau}} H(\sigma,\sigma') \sum_{A',B'} \sum_{\substack{R_1 \in \mathcal{R}(\sigma,A,A'), R_2 \in \mathcal{R}(\tau,A',B'), \text{ and } R_3 \in \mathcal{R}({\sigma'}^T,B',B) \\ R_1,R_2,R_3 \text{ are not properly composable}}} M_{R_1}(A,A')M_{R_2}(A',B')M_{R_3}(B',B)$$

Thus, to understand the difference between M^{fact} and M^{orth}, we need to analyze the terms $\chi_{R_1}\chi_{R_2}\chi_{R_3} = \chi_{R_1 \circ R_2 \circ R_3}$ for ribbons R_1, R_2, R_3 which are composable but not properly composable. These terms, which we call intersection terms, are not negligible and must be analyzed carefully. In particular, we decompose each resulting ribbon $R = R_1 \circ R_2 \circ R_3$ into new left, middle, and right parts. We do this as follows:

1. Let V_* be the set of vertices which appear more than once in $V(R_1 \circ R_2 \circ R_3)$. In other words, V_* is the set of vertices involved in the intersections between R_1, R_2, and R_3 (not counting the facts that $B_{R_1} = A_{R_2}$ and $B_{R_2} = A_{R_3}$ because we expect these intersections).

2. Let A' be the leftmost minimum vertex separator of A_{R_1} and $B_{R_1} \cup V_*$ in R_1. We turn A' into a matrix index by specifying an ordering $O_{A'}$ for the vertices in A'.

3. Let B' be the leftmost minimum vertex separator of $A_{R_3} \cup V_*$ and B_{R_3} in R_2. We turn B' into a matrix index by specifying an ordering $O_{B'}$ for the vertices in B'.

4. Decompose R_1 as $R_1 = {R'}_1 \cup R_4$ where ${R'}_1$ is the part of R_1 between A_{R_1} and A' and R_4 is the part of R_1 between B' and $B_{R_1} = A_{R_2}$. Similarly, decompose R_3 as $R_3 = R_5 \cup {R'}_3$ where R_5 is the part of R_3 between $B_{R_1} = A_{R_2}$ and B' and ${R'}_3$ is the part of R_3 between B' and B_{R_3}.

5. Take $R'_2 = R_4 \circ R_2 \circ R_5$ and note that $R'_1 \circ R'_2 \circ R'_3 = R_1 \circ R_2 \circ R_3$. We view R'_1, R'_2, R'_3 as the left, middle, and right parts of $R = R_1 \circ R_2 \circ R_3$

While we will verify our analysis by checking the coefficients of the ribbons, we want to express everything in terms of shapes. We use the following conventions for the names of the shapes:

1. As usual, we let σ, τ, and σ'^T be the shapes of R_1, R_2, and R_3.

2. We let γ and γ'^T be the shapes of R_4 and R_5.

3. We let σ_2, τ_P, and $\sigma_2'^T$ be the shapes of R_1', R_2', and R_3'. Here P is the intersection pattern induced by R_4, R_2, and R_5 which we define in the next subsection.

Remark 3.7. *A key feature of our analysis is that it will work the same way regardless of the shapes $\sigma_2, \sigma_2'^T$ of R_1' and R_3'. In other words, if we replace σ_2 by σ_{2a} and σ_2' by σ_{2a}' for a given intersection term, this just replaces $\sigma = \sigma_2 \cup \gamma$ with $\sigma_a = \sigma_{2a} \cup \gamma$ and $\sigma' = \sigma_2' \cup \gamma'$ with $\sigma_a' = \sigma_{2a}' \cup \gamma'$. This allows us to focus on the shapes γ, τ, and γ'^T and is the reason why the $-\gamma, \gamma$ operation appears in our results.*

3.3 Intersection Term Analysis

In this section, we implement our strategy for analyzing intersection terms. For simplicity, we only give rough definitions and proof sketches here. For a more rigorous treatment, see Apendix B.

We begin by defining intersection patterns which describe how the ribbons R_1, R_2, and R_3 intersect.

Definition 3.8 (Rough Definition of Intersection Patterns)**.** *Given $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$, and ribbons R_1, R_2, and R_3 of shapes γ, τ, and γ'^T which are composable but not properly composable, we define the intersection pattern P induced by R_1, R_2, and R_3 and the resulting shape τ_P as follows:*

1. *We take $V(P) = V(\gamma \circ \tau \circ \gamma'^T)$.*

2. *We take $E(P)$ to be the set of edges (u, v) such that u, v are distinct vertices in $V(\sigma \circ \tau \circ \sigma'^T)$ but u and v correspond to the same vertex in $R_1 \circ R_2 \circ R_3$*

3. *We define τ_P to be the shape of the ribbon $R = R_1 \circ R_2 \circ R_3$*

Definition 3.9. *Given $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, and $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$, we define $\mathcal{P}_{\gamma,\tau,\gamma'^T}$ to be the set of all possible intersection patterns P which can be induced by ribbons R_1, R_2, and R_3 of shapes γ, τ, and γ'^T.*

Remark 3.10. *Note that if $\gamma = Id_{U_\tau}$ and $\gamma' = Id_{V_\tau}$ then $\mathcal{P}_{\gamma,\tau,\gamma'^T} = \emptyset$ as every intersection pattern must have an unexpected intersection so either γ or γ' must be non-trivial.*

It would be nice if the intersection pattern P together with the ribbon R allowed us to recover the original ribbons R_1, R_2, and R_3. Unfortunately, it is possible for different triples of ribbons to result in the same intersection pattern P and ribbon R. That said, the number of such triples cannot be too large, and this is sufficient for our purposes.

Definition 3.11. *Given an intersection pattern $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$, let R be a ribbon of shape τ_P. We define $N(P)$ to be the number of different triples of ribbons R_1, R_2, R_3 such that $R_1 \circ R_2 \circ R_3 = R$ and R_1, R_2, R_3 induce the intersection pattern P.*

Lemma 3.12. *For all intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$, $N(P) \leq |V(\tau_P)|^{|V(\gamma) \setminus U_\gamma| + |V(\gamma') \setminus U_{\gamma'}|}$*

Proof sketch. This can be proved by making the following observations:

1. $A_{R_1} = A_R$ and $B_{R_3} = B_R$.
2. All of the remining vertices in $V(R_1)$ and $V(R_3)$ must be equal to some vertex in $V(R)$.
3. Once R_1 and R_3 are determined, there is at most one ribbon R_2 such that R_1, R_2, R_3 are composable, $R = R_1 \circ R_2 \circ R_3$, and R_1, R_2, R_3 induce the intersection pattern P.

□

With these definitions, we can now analyze the intersection terms.

Definition 3.13. *Given a left shape σ, define e_σ to be the vector which has a 1 in coordinate σ and has a 0 in all other coordinates.*

Lemma 3.14. *For all $\tau \in \mathcal{M}'$, $\sigma \in \mathcal{L}_{U_\tau}$, and $\sigma' \in \mathcal{L}_{V_\tau}$,*

$$M^{fact}_\tau(e_\sigma e^T_{\sigma'}) - M^{orth}_\tau(e_\sigma e^T_{\sigma'}) = \sum_{\sigma_2 \in \mathcal{L}, \gamma \in \Gamma : \sigma_2 \circ \gamma = \sigma} \frac{1}{|Aut(U_\gamma)|} \sum_{P \in \mathcal{P}_{\gamma, \tau, Id_{V_\tau}}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'})$$

$$+ \sum_{\sigma'_2 \in \mathcal{L}, \gamma' \in \Gamma : \sigma'_2 \circ \gamma' = \sigma'} \frac{1}{|Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{Id_{U_\tau}, \tau, \gamma'^T}} N(P) M^{orth}_{\tau_P}(e_\sigma e^T_{\sigma'_2})$$

$$+ \sum_{\sigma_2 \in \mathcal{L}, \gamma \in \Gamma : \sigma_2 \circ \gamma = \sigma} \sum_{\sigma'_2 \in \mathcal{L}, \gamma' \in \Gamma : \sigma'_2 \circ \gamma' = \sigma'} \frac{1}{|Aut(U_\gamma)| \cdot |Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{\gamma, \tau, \gamma'^T}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'_2})$$

Proof sketch. This lemma follows from the following bijection. Consider the third term

$$\sum_{\sigma_2 \in \mathcal{L}, \gamma \in \Gamma : \sigma_2 \circ \gamma = \sigma} \sum_{\sigma'_2 \in \mathcal{L}, \gamma' \in \Gamma : \sigma'_2 \circ \gamma' = \sigma'} \frac{1}{|Aut(U_\gamma)| \cdot |Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{\gamma, \tau, \gamma'^T}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'_2})$$

On one side, we have the following data:

1. Ribbons R_1, R_2, and R_3 of shapes γ, τ, γ'^T such that R_1, R_2, R_3 are composable but R_1 and $R_2 \circ R_3$ are not properly composable (i.e. R_1 has an unexpected intersection with R_2 and/or R_3) and $R_1 \circ R_2$ and R_3 are not properly composable (i.e. R_3 has an unexpected intersection with R_1 and/or R_2).
2. An ordering $O_{A'}$ on the leftmost minimum vertex separator A' of A_{R_1} and $V_* \cup B_{R_1}$ (recall that V_* is the set of vertices which appear more than once in $V(R_1 \circ R_2 \circ R_3)$).
3. An ordering $O_{B'}$ on the rightmost minimum vertex separator B' of $V_* \cup A_{R_3}$ and B_{R_3}.

On the other side, we have the following data

1. An intersection pattern $P \in \mathcal{P}_{\gamma, \tau, \gamma'^T}$ where γ and γ'^T are non-trivial.
2. Ribbons R'_1, R'_2, R'_3 of shapes σ_2, τ_P, ${\sigma'_2}^T$ which are properly composable

3. A number in $[N(P)]$ describing which possible triple of ribbons resulted in the intersection pattern P and the ribbon R'_2.

To see this bijection, note that given the data on the first side, we can recover the ribbons R'_1, R'_2, and R'_3 as follows:

1. We decompose R_1 as $R_1 = R'_1 \circ R_4$ where $B_{R'_1} = A_{R_4} = A'$ with the ordering $O_{A'}$.

2. We decompose R_3 as $R_3 = R_5 \circ R'_3$ where where $B_{R_5} = A_{R'_3} = B'$ with the ordering $O_{B'}$.

3. We take $R'_2 = R_4 \circ R_2 \circ R_5$.

The intersection pattern P and the number in $[N(P)]$ can be obtained from R_1, R_2, and R_3.

Conversely, with the data on the other side, we can recover the data on the first side as follows:

1. R'_2 gives an ordering $O_{A'}$ for $A' = A_{R'_2}$ and an ordering $O_{B'}$ for $B' = B_{R'_2}$.

2. The ribbon R'_2, intersection pattern P, and number in $[N(P)]$ allow us to recover R_4, R_2, and R_5.

3. We take $R_1 = R'_1 \circ R_4$ and $R_3 = R_5 \circ R'_3$.

Thus, both sides have the same coefficient for each ribbon.

The analysis for the the first term is the same except that when γ' is trivial, we always take $\gamma' = Id_{V_\tau}$. Thus, we always have that $B' = B_{R'_2} = B_{R_2}$ (with the same ordering) and $R'_3 = R_3 = Id_{B'}$. Because of this, there is no need to specify R_3, R'_3, R_5, or an ordering on B'.

Similarly, the analysis for the the second term is the same except that when γ is trivial, we always take $\gamma = Id_{U_\tau}$. Thus, we always have that $A' = A_{R'_2} = A_{R_2}$ (with the same ordering) and $R'_1 = R_1 = Id_{A'}$. Because of this, there is no need to specify R_1, R'_1, R_4, or an ordering on A'. □

Applying Lemma 3.14 for all σ and σ' simultaneously, we obtain the following corollary.

Definition 3.15. *For all $U, V \in \mathcal{I}_{mid}$, given a $\gamma \in \Gamma_{U,V}$ and a vector v indexed by left shapes $\sigma \in \mathcal{L}_V$, define $v^{-\gamma}$ to be the vector indexed by left shapes $\sigma_2 \in \mathcal{L}_U$ such that $v^{-\gamma}(\sigma_2) = v(\sigma_2 \circ \gamma)$ if $\sigma_2 \circ \gamma \in \mathcal{L}_V$ and $v^{-\gamma}(\sigma_2) = 0$ otherwise.*

Proposition 3.16. *For all composable $\gamma_2, \gamma_1 \in \Gamma$ and all vectors v indexed by left shapes in $\mathcal{L}_{V_{\gamma_1}}$, $(v^{-\gamma_1})^{-\gamma_2} = v^{-\gamma_2 \circ \gamma_1}$*

Corollary 3.17. *For all $\tau \in \mathcal{M}'$, for all left τ-vectors v and all right τ-vectors w,*

$$M_\tau^{orth}(vw^T) = M_\tau^{fact}(vw^T) - \sum_{\gamma \in \Gamma_{*,U_\tau}} \frac{1}{|Aut(U_\gamma)|} \sum_{P \in \mathcal{P}_{\gamma,\tau,Id_{V_\tau}}} N(P) M_{\tau_P}^{orth}(v^{-\gamma} w^T)$$
$$- \sum_{\gamma' \in \Gamma_{*,V_\tau}} \frac{1}{|Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{Id_{U_\tau},\tau,\gamma'^T}} N(P) M_{\tau_P}^{orth}(v(w^{-\gamma})^T)$$
$$- \sum_{\gamma \in \Gamma_{*,U_\tau}} \sum_{\gamma' \in \Gamma_{*,V_\tau}} \frac{1}{|Aut(U_\gamma)| \cdot |Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}} N(P) M_{\tau_P}^{orth}(v^{-\gamma}(w^{-\gamma'})^T)$$

Applying Corollary 3.17 iteratively, we obtain the following theorem:

Definition 3.18. *Given $\gamma, \gamma' \in \Gamma \cup \{Id_U : U \in \mathcal{I}_{mid}\}$ and $j > 0$, let $\Gamma_{\gamma,\gamma',j}$ be the set of all $\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma \cup \{Id_U : U \in \mathcal{I}_{mid}\}$ such that:*

1. *$\gamma_j, \ldots, \gamma_1$ are composable and $\gamma_j \circ \ldots \circ \gamma_1 = \gamma$*
2. *$\gamma'_j, \ldots, \gamma'_1$ are composable and $\gamma'_j \circ \ldots \circ \gamma'_1 = \gamma'$*
3. *For all $i \in [1, j]$, γ_i or γ'_i is non-trivial (i.e. $\gamma_i \neq Id_{U_{\gamma_i}}$ or $\gamma'_i \neq Id_{U_{\gamma'_i}}$).*

Remark 3.19. *Note that if $\gamma = Id_U$ and $\gamma' = Id_V$ then for all $j > 0$, $\Gamma_{\gamma,\gamma',j} = \emptyset$.*

Theorem 3.20. *For all $\tau \in \mathcal{M}'$, left τ-vectors v, and right τ-vetors w,*

$$M^{orth}_{\tau}(vw^T) = M^{fact}_{\tau}(vw^T) + \sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \sum_{j>0} (-1)^j \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i: \gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i: \gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}$$
$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}} \left(\prod_{i=1}^{j} N(P_i)\right) M^{fact}_{\tau_{P_j}}(v^{-\gamma}(w^{-\gamma'})^T)$$

where we take $\tau_{P_0} = \tau$.

3.4 Bounding the difference between M^{fact} and M^{orth}

In this subsection, we bound the difference between $M^{fact}_{\tau}(H_\tau)$ and $M^{orth}_{\tau}(H_\tau)$. We recall the following conditions on $B(\gamma)$, $N(\gamma)$, and $c(\gamma)$:

1. For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, and $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$,

$$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \left(\prod_{i: \gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|}\right) \left(\prod_{i: \gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}\right)$$
$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}} \left(\prod_{i=1}^{j} N(P_i)\right) \leq \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

2. For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau}$, and $\gamma' \in \Gamma_{*,V_\tau}$, for all $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$, $B_{norm}(\tau_P) \leq B(\gamma)B(\gamma')B_{norm}(\tau)$
3. $\forall V \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)| c(\gamma)} \leq \epsilon' \leq \frac{1}{20}$

With these conditions, we can now bound the difference between M^{fact} and M^{orth}.

Lemma 3.21. *If the norm bounds and the conditions on $B(\gamma)$, $N(\gamma)$, and $c(\gamma)$ hold then for all $\tau \in \mathcal{M}'$, left τ-vectors v, and right τ-vetors w,*

$$\left(M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T)\right) - \left(M^{orth}_{\tau}(vw^T) + M^{orth}_{\tau^T}(wv^T)\right) \preceq$$

$$\epsilon' B_{norm}(\tau) M^{fact}_{Id_{U_\tau}}(vv^T) + 2\sum_{\gamma \in \Gamma_{*,U_\tau}} \frac{B(\gamma)^2 N(\gamma)^2 B_{norm}(\tau) c(\gamma)}{|Aut(U_\gamma)|} M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) +$$
$$\epsilon' B_{norm}(\tau) M^{fact}_{Id_{V_\tau}}(ww^T) + 2\sum_{\gamma' \in \Gamma_{*,V_\tau}} \frac{B(\gamma')^2 N(\gamma')^2 B_{norm}(\tau) c(\gamma')}{|Aut(U_{\gamma'})|} M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)$$

Proof. By Theorem 3.20, taking $\tau_{P_0} = \tau$,

$$M^{orth}_{\tau}(vw^T) = M^{fact}_{\tau}(vw^T) +$$
$$\sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \sum_{j>0} (-1)^j \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}$$
$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right) M^{fact}_{\tau_{P_j}}(v^{-\gamma}(w^{-\gamma'})^T)$$

Taking the transpose of this equation gives

$$M^{orth}_{\tau^T}(wv^T) = M^{fact}_{\tau^T}(wv^T) +$$
$$\sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \sum_{j>0} (-1)^j \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}$$
$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right) M^{fact}_{\tau^T_{P_j}}(w^{-\gamma'}(v^{-\gamma})^T)$$

Now observe that by Lemma 3.5, if the norm bounds hold,

$$\pm\left(M^{fact}_{\tau_{P_j}}(v^{-\gamma}(w^{-\gamma'})^T) + M^{fact}_{\tau^T_{P_j}}(w^{-\gamma'}(v^{-\gamma})^T)\right) =$$
$$\pm M^{fact}_{\tau_{P_j}}\left(\left(\sqrt{\frac{N(\gamma)B(\gamma)c(\gamma)}{N(\gamma')B(\gamma')c(\gamma')}} v^{-\gamma}\right)\left(\sqrt{\frac{N(\gamma')B(\gamma')c(\gamma')}{N(\gamma)B(\gamma)c(\gamma)}} (w^{-\gamma'})^T\right)\right) \pm$$
$$M^{fact}_{\tau^T_{P_j}}\left(\left(\sqrt{\frac{N(\gamma')B(\gamma')c(\gamma')}{N(\gamma)B(\gamma)c(\gamma)}} w^{-\gamma'}\right)\left(\sqrt{\frac{N(\gamma)B(\gamma)c(\gamma)}{N(\gamma')B(\gamma')c(\gamma')}} (v^{-\gamma})^T\right)\right) \preceq$$
$$B_{norm}(\tau_{P_j})\left(\frac{N(\gamma)B(\gamma)c(\gamma)}{N(\gamma')B(\gamma')c(\gamma')} M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) + \frac{N(\gamma')B(\gamma')c(\gamma')}{N(\gamma)B(\gamma)c(\gamma)} M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)\right)$$

Combining these equations,

$$\left(M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T)\right) - \left(M^{orth}_{\tau}(vw^T) + M^{orth}_{\tau^T}(wv^T)\right) \preceq$$

$$\sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i: \gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i: \gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}$$

$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right) B_{norm}(\tau_{P_j})$$

$$\left(\frac{N(\gamma)B(\gamma)c(\gamma)}{N(\gamma')B(\gamma')c(\gamma')} M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) + \frac{N(\gamma')B(\gamma')c(\gamma')}{N(\gamma)B(\gamma)c(\gamma)} M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)\right)$$

From the conditions on $B(\gamma)$ and $N(\gamma)$,

1. $B_{norm}(\tau_{P_j}) \le B(\gamma)B(\gamma')B_{norm}(\tau)$
2. $$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \left(\prod_{i: \gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|}\right) \left(\prod_{i: \gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|}\right)$$
$$\sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right) \le \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

Putting these equations together,

$$\left(M^{fact}_{\tau}(vw^T) + M^{fact}_{\tau^T}(wv^T)\right) - \left(M^{orth}_{\tau}(vw^T) + M^{orth}_{\tau^T}(wv^T)\right) \preceq$$

$$\sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \frac{B(\gamma)^2 N(\gamma)^2 B_{norm}(\tau) c(\gamma)}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}} c(\gamma')} M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) +$$

$$\sum_{\substack{\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}, \gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}: \\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \frac{B(\gamma')^2 N(\gamma')^2 B_{norm}(\tau) c(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}} c(\gamma)} M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)$$

Now observe that

$$\sum_{\substack{\gamma\in\Gamma_{*,U_\tau}\cup\{Id_{U_\tau}\},\gamma'\in\Gamma_{*,V_\tau}\cup\{Id_{V_\tau}\}:\\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \frac{B(\gamma)^2N(\gamma)^2B_{norm}(\tau)c(\gamma)}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}c(\gamma')}M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) \preceq$$

$$\left(\sum_{\gamma'\in\Gamma_{*,V_\tau}}\frac{1}{|Aut(U_{\gamma'})|c(\gamma')}\right)B_{norm}(\tau)M^{fact}_{Id_{U_\tau}}(vv^T)+$$

$$\sum_{\gamma\in\Gamma_{*,U_\tau}}\left(\sum_{\gamma'\in\Gamma_{*,V_\tau}\cup\{Id_{V_\tau}\}}\frac{1}{(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}c(\gamma')}\right)\frac{B(\gamma)^2N(\gamma)^2B_{norm}(\tau)c(\gamma)}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}}M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T) \preceq$$

$$\epsilon' B_{norm}(\tau)M^{fact}_{Id_{U_\tau}}(vv^T) + 2\sum_{\gamma\in\Gamma_{*,U_\tau}}\frac{B(\gamma)^2N(\gamma)^2B_{norm}(\tau)c(\gamma)}{|Aut(U_\gamma)|}M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T)$$

Following similar logic,

$$\sum_{\substack{\gamma\in\Gamma_{*,U_\tau}\cup\{Id_{U_\tau}\},\gamma'\in\Gamma_{*,V_\tau}\cup\{Id_{V_\tau}\}:\\ \gamma \text{ or } \gamma' \text{ is non-trivial}}} \frac{B(\gamma')^2N(\gamma')^2B_{norm}(\tau)c(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}c(\gamma)}M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T) \preceq$$

$$\epsilon' B_{norm}(\tau)M^{fact}_{Id_{V_\tau}}(ww^T) + 2\sum_{\gamma'\in\Gamma_{*,V_\tau}}\frac{B(\gamma')^2N(\gamma')^2B_{norm}(\tau)c(\gamma')}{|Aut(U_{\gamma'})|}M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)$$

Putting everything together,

$$\left(M^{fact}_\tau(vw^T)+M^{fact}_{\tau^T}(wv^T)\right)-\left(M^{orth}_\tau(vw^T)+M^{orth}_{\tau^T}(wv^T)\right)\preceq$$

$$\epsilon' B_{norm}(\tau)M^{fact}_{Id_{U_\tau}}(vv^T) + 2\sum_{\gamma\in\Gamma_{*,U_\tau}}\frac{B(\gamma)^2N(\gamma)^2B_{norm}(\tau)c(\gamma)}{|Aut(U_\gamma)|}M^{fact}_{Id_{U_\gamma}}(v^{-\gamma}(v^{-\gamma})^T)+$$
$$\epsilon' B_{norm}(\tau)M^{fact}_{Id_{V_\tau}}(ww^T) + 2\sum_{\gamma'\in\Gamma_{*,V_\tau}}\frac{B(\gamma')^2N(\gamma')^2B_{norm}(\tau)c(\gamma')}{|Aut(U_{\gamma'})|}M^{fact}_{Id_{U_{\gamma'}}}(w^{-\gamma'}(w^{-\gamma'})^T)$$

as needed. □

Using Lemma 3.21 we have the following corollaries:

Corollary 3.22. *For all $U \in \mathcal{I}_{mid}$, if the norm bounds and the conditions on $B(\gamma)$, $N(\gamma)$, and $c(\gamma)$ hold and $H_{Id_U} \succeq 0$ then*

$$M^{fact}_{Id_U}(H_{Id_U}) - M^{orth}_{Id_U}(H_{Id_U}) \preceq \epsilon' M^{fact}_{Id_U}(H_{Id_U}) + 2\sum_{\gamma\in\Gamma_{*,U}}\frac{B(\gamma)^2N(\gamma)^2c(\gamma)}{|Aut(U_\gamma)|}M^{fact}_{Id_{U_\gamma}}(H^{-\gamma,\gamma}_{Id_U})$$

Corollary 3.23. *For all $U \in \mathcal{I}_{mid}$ and all $\tau \in \mathcal{M}_U$, if the norm bounds and the conditions on $B(\gamma)$, $N(\gamma)$, and $c(\gamma)$ hold and*

$$\begin{bmatrix} \frac{1}{||Aut(U)|c(\tau)}H_{Id_U} & B_{norm}(\tau)H_\tau \\ B_{norm}(\tau)H^T_\tau & \frac{1}{|Aut(U)|c(\tau)}H_{Id_U} \end{bmatrix} \succeq 0$$

then

$$\left(M^{fact}_{\tau}(H_\tau) + M^{fact}_{\tau^T}(H^T_\tau)\right) - \left(M^{orth}_{\tau}(H_\tau) + M^{orth}_{\tau^T}(H^T_\tau)\right) \preceq$$

$$2\epsilon' \frac{1}{|Aut(U)|c(\tau)} M^{fact}_{Id_U}(H_{Id_U}) + 4 \sum_{\gamma \in \Gamma_{*,U}} \frac{B(\gamma)^2 N(\gamma)^2 c(\gamma)}{|Aut(U_\gamma)| \cdot |Aut(U)| c(\tau)} M^{fact}_{Id_{U_\gamma}}(H^{-\gamma,\gamma}_{Id_U})$$

3.5 Proof of the Main Theorem

We now prove the following theorem which is a slight modification of Theorem 3.1 and which implies Theorem 3.1.

Theorem 3.24. *For all $\epsilon > 0$ and all $\epsilon' \in (0, \frac{1}{20}]$, for any moment matrix*

$$\Lambda = \sum_{U \in \mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{orth}_{\tau}(H_\tau),$$

if we have that for all $\alpha \in \mathcal{M}'$, $||M_\alpha|| \le B_{norm}(\alpha)$ and $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$ are functions such that

1. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau}$, $\gamma' \in \Gamma_{*,V_\tau}$, and all intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,*

$$B_{norm}(\tau_P) \le B(\gamma)B(\gamma')B_{norm}(\tau)$$

2. *For all composable γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.*
3. $\forall U \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{U,*}} \frac{1}{|Aut(U)|c(\gamma)} < \epsilon'$
4. $\forall V \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)|c(\gamma)} < \epsilon'$
5. $\forall U \in \mathcal{I}_{mid}, \sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)|c(\tau)} < \epsilon'$
6. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, and $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$,*

$$\sum_{j>0} \sum_{\gamma_1,\gamma'_1,\cdots,\gamma_j,\gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1,\cdots,P_j: P_i \in \mathcal{P}_{\gamma_i,\tau_{P_{i-1}},{\gamma'_i}^T}} \left(\prod_{i=1}^{j} N(P_i)\right)$$
$$\le \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

and we have SOS-symmetric coefficient matrices $\{H'_\gamma : \gamma \in \Gamma\}$ such that the following conditions hold:

1. *For all $U \in \mathcal{I}_{mid}$, $H_{Id_U} \succeq 0$*

2. *For all* $U \in \mathcal{I}_{mid}$ *and* $\tau \in \mathcal{M}_U$,

$$\begin{bmatrix} \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} & B_{norm}(\tau) H_\tau \\ B_{norm}(\tau) H_\tau^T & \frac{1}{|Aut(U)|c(\tau)} H_{Id_U} \end{bmatrix} \succeq 0$$

3. *For all* $U, V \in \mathcal{I}_{mid}$ *where* $w(U) > w(V)$ *and all* $\gamma \in \Gamma_{U,V}$,

$$c(\gamma)^2 N(\gamma)^2 B(\gamma)^2 H_{Id_V}^{-\gamma,\gamma} \preceq H'_\gamma$$

then

$$\Lambda \succeq \frac{1}{2}\left(\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U})\right) - 3\left(\sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)}\right) Id_{sym}$$

If it is also true that

$$\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) \succeq 6\left(\sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{U,*}} \frac{d_{Id_U}(H'_\gamma, H_{Id_U})}{|Aut(U)|c(\gamma)}\right) Id_{sym}$$

then $\Lambda \succeq 0$.

Proof. We make the following observations:

1. By Theorem 3.2,

$$\sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) + \sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} M^{fact}_{\tau}(H_\tau) \succeq (1 - 2\epsilon') \sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U})$$

2. By Corollary 3.22,

$$\sum_{U \in \mathcal{I}_{mid}} \left(M^{fact}_{Id_U}(H_{Id_U}) - M^{orth}_{Id_U}(H_{Id_U})\right) \preceq \epsilon' \sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) + 2 \sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{*,U}} \frac{M^{fact}_{Id_{U_\gamma}}(H'_\gamma)}{c(\gamma)|Aut(U_\gamma)|}$$

3. By Corollary 3.23,

$$\sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} \left(M^{fact}_{\tau}(H_\tau) - M^{orth}_{\tau}(H_\tau)\right) \preceq$$

$$\sum_{U \in \mathcal{I}_{mid}} \sum_{\tau \in \mathcal{M}_U} \left(\frac{2\epsilon'}{|Aut(U)|c(\tau)} M^{fact}_{Id_U}(H_{Id_U}) + 4 \sum_{\gamma \in \Gamma_{*,U}} \frac{B(\gamma)^2 N(\gamma)^2 c(\gamma)}{|Aut(U_\gamma)| \cdot |Aut(U)|c(\tau)} M^{fact}_{Id_{U_\gamma}}(H^{-\gamma,\gamma}_{Id_U})\right) \preceq$$

$$2\epsilon'^2 \sum_{U \in \mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) + 4\epsilon' \sum_{U \in \mathcal{I}_{mid}} \sum_{\gamma \in \Gamma_{*,U}} \frac{M^{fact}_{Id_{U_\gamma}}(H'_\gamma)}{c(\gamma)|Aut(U_\gamma)|}$$

4.

$$\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{*,U}}\frac{M^{fact}_{Id_{U_\gamma}}(H'_\gamma)}{c(\gamma)|Aut(U_\gamma)|} = \sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{*,U}}\frac{M^{fact}_{Id_{U_\gamma}}(H_{Id_{U_\gamma}}) + \left(M^{fact}_{Id_{U_\gamma}}(H'_\gamma) - M^{fact}_{Id_{U_\gamma}}(H_{Id_{U_\gamma}})\right)}{c(\gamma)|Aut(U_\gamma)|} \preceq$$

$$\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{*,U}}\frac{M^{fact}_{Id_{U_\gamma}}(H_{Id_{U_\gamma}})}{c(\gamma)|Aut(U_\gamma)|} + \left(\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_{U_\gamma}}(H'_\gamma, H_{Id_{U_\gamma}})}{|Aut(U_\gamma)|c(\gamma)}\right) Id_{sym} \preceq$$

$$\epsilon' \sum_{U\in\mathcal{I}_{mid}} M^{fact}_U(H_{Id_U}) + \left(\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_{U_\gamma}}(H'_\gamma, H_{Id_{U_\gamma}})}{|Aut(U_\gamma)|c(\gamma)}\right) Id_{sym}$$

Putting everything together,

$$\Lambda = \sum_{U\in\mathcal{I}_{mid}} M^{orth}_{Id_U}(H_{Id_U}) + \sum_{U\in\mathcal{I}_{mid}}\sum_{\tau\in\mathcal{M}_U} M^{orth}_{\tau}(H_\tau) =$$

$$\sum_{U\in\mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) + \sum_{U\in\mathcal{I}_{mid}}\sum_{\tau\in\mathcal{M}_U} M^{fact}_{\tau}(H_\tau) + \sum_{U\in\mathcal{I}_{mid}}\left(M^{fact}_{Id_U}(H_{Id_U}) - M^{orth}_{Id_U}(H_{Id_U})\right) +$$

$$\sum_{U\in\mathcal{I}_{mid}}\sum_{\tau\in\mathcal{M}_U}\left(M^{fact}_{\tau}(H_\tau) - M^{orth}_{\tau}(H_\tau)\right) \succeq$$

$$(1 - 3\epsilon' - 2\epsilon'^2)\sum_{U\in\mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) - (2 + 4\epsilon')\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{*,U}}\frac{M^{fact}_{Id_{U_\gamma}}(H'_\gamma)}{c(\gamma)|Aut(U_\gamma)|} \succeq$$

$$(1 - 5\epsilon' - 6\epsilon'^2)\sum_{U\in\mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) - (2 + 4\epsilon')\left(\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_{U_\gamma}}(H'_\gamma, H_{Id_{U_\gamma}})}{|Aut(U_\gamma)|c(\gamma)}\right) Id_{sym} \succeq$$

$$\frac{1}{2}\sum_{U\in\mathcal{I}_{mid}} M^{fact}_{Id_U}(H_{Id_U}) - 3\left(\sum_{U\in\mathcal{I}_{mid}}\sum_{\gamma\in\Gamma_{U,*}}\frac{d_{Id_{U_\gamma}}(H'_\gamma, H_{Id_{U_\gamma}})}{|Aut(U_\gamma)|c(\gamma)}\right) Id_{sym}$$

□

4 Choosing the functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$

In this subsection, we give functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$ which satisfy the conditions needed for our machinery.

4.1 Theorem Statements

Recall the following definitions from Section 2.10.

Definition 4.1. *We define S_α to be the leftmost minimum vertex separator of α*

Definition 4.2 (Simplified Isolated Vertices)**.** *Under our simplifying assumptions, we define*

$$I_\alpha = \{v \in W_\alpha : v \text{ is not incident to any edges in } E(\alpha)\}$$

Theorem 4.3 (Simplified $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$)**.** *Under our simplifying assumptions, for all $\epsilon, \epsilon' > 0$ and all $D_V \in \mathbb{N}$, if we take*

1. $q = 3\left\lceil D_V ln(n) + \frac{ln(\frac{1}{\epsilon})}{3} + D_V ln(5) + 3D_V^2 ln(2)\right\rceil$
2. $B_{vertex} = 6D_V\sqrt[4]{2eq}$
3. $B_{norm}(\alpha) = B_{vertex}^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha))+w(I_\alpha)-w(S_\alpha)}{2}}$
4. $B(\gamma) = B_{vertex}^{|V(\gamma)\setminus U_\gamma|+|V(\gamma)\setminus V_\gamma|} n^{\frac{w(V(\gamma)\setminus U_\gamma)}{2}}$
5. $N(\gamma) = (3D_V)^{2|V(\gamma)\setminus V_\gamma|+|V(\gamma)\setminus U_\gamma|}$
6. $c(\alpha) = \frac{\epsilon'}{5(3D_V)^{|U_\alpha\setminus V_\alpha|+|V_\alpha\setminus U_\alpha|+2|E(\alpha)|}2^{|V(\alpha)\setminus(U_\alpha\cup V_\alpha)|}}$

then the following conditions hold:

1. *With probability at least $(1-\epsilon)$, $\forall \alpha \in \mathcal{M}'$, $||M_\alpha|| \leq B_{norm}(\alpha)$*
2. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$, and intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,*
$$B_{norm}(\tau_P) \leq B(\gamma)B(\gamma')B_{norm}(\tau)$$
3. *For all composable γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.*
4. $\forall U \in \mathcal{I}_{mid}$, $\sum_{\gamma \in \Gamma_{U,*}} \frac{1}{|Aut(U)|c(\gamma)} < \epsilon'$
5. $\forall V \in \mathcal{I}_{mid}$, $\sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)|c(\gamma)} < \epsilon'$
6. $\forall U \in \mathcal{I}_{mid}$, $\sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)|c(\tau)} < \epsilon'$
7. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, and $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$,*

$$\sum_{j>0} \sum_{\gamma_1,\gamma_1',\cdots,\gamma_j,\gamma_j' \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma_i' \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i'})|} \sum_{P_1,\cdots,P_j:P_i \in \mathcal{P}_{\gamma_i,\tau_{P_{i-1}},\gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right)$$
$$\leq \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

4.1.1 General functions $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$*

Recall the following definitions from Section 2.10.1.

Definition 4.4 ($S_{\alpha,min}$ and $S_{\alpha,max}$)**.** *Given a shape $\alpha \in \mathcal{M}'$, define $S_{\alpha,min}$ to be the leftmost minimum vertex separator of α if all edges with multiplicity at least 2 are deleted and define $S_{\alpha,max}$ to be the leftmost minimum vertex separator of α if all edges with multiplicity at least 2 are present.*

Definition 4.5 (General I_α)**.** *Given a shape α, define I_α to be the set of vertices in $V(\alpha)\setminus(U_\alpha \cup V_\alpha)$ such that all edges incident with that vertex have multplicity at least 2.*

Definition 4.6 (B_Ω)**.** *We take $B_\Omega(j)$ to be a non-decreasing function such that for all $j \in \mathbb{N}$, $E_\Omega[x^j] \leq B_\Omega(j)^j$*

Definition 4.7. *For all i, we define h_i^+ to be the polynomial h_i where we make all of the coefficients have positive sign.*

Lemma 4.8. *If $\Omega = N(0,1)$ then we can take $B_\Omega(j) = \sqrt{j}$ and we have that*

Theorem 4.9 (General $B_{norm}(\alpha)$, $B(\gamma)$, $N(\gamma)$, and $c(\alpha)$)**.** *For all $\epsilon, \epsilon' > 0$ and all $D_V, D_E \in \mathbb{N}$, if we take*

1. $q = \left\lceil 3D_V ln(n) + ln(\frac{1}{\epsilon}) + (3D_V)^k ln(D_E + 1) + 3D_V ln(5) \right\rceil$
2. $B_{vertex} = 6qD_V$
3. $B_{edge}(e) = 2h^+_{l_e}(B_\Omega(6D_V D_E)) \max_{j \in [0, 3D_V D_E]} \left\{ \left(h^+_j(B_\Omega(2qj))\right)^{\frac{l_e}{\max\{j, l_e\}}} \right\}$
4. $B_{norm}(\alpha) = 2eB_{vertex}^{|V(\alpha)\setminus U_\alpha| + |V(\alpha)\setminus V_\alpha|} \left(\prod_{e \in E(\alpha)} B_{edge}(e)\right) n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_\alpha)}{2}}$
5. $B(\gamma) = B_{vertex}^{|V(\gamma)\setminus U_\gamma| + |V(\gamma)\setminus V_\gamma|} \left(\prod_{e \in E(\gamma)} B_{edge}(e)\right) n^{\frac{w(V(\gamma)\setminus U_\gamma)}{2}}$
6. $N(\gamma) = (3D_V)^{2|V(\gamma)\setminus V_\gamma| + |V(\gamma)\setminus U_\gamma|}$
7. $c(\alpha) = \frac{\epsilon'}{5(3t_{max}D_V)^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + k|E(\alpha)|}(2t_{max})^{|V(\alpha)\setminus(U_\alpha \cup V_\alpha)|}}$

then the following conditions hold:

1. *With probability at least $(1-\epsilon)$, $\forall \alpha \in \mathcal{M}'$, $||M_\alpha|| \leq B_{norm}(\alpha)$*
2. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$, and intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,*
$$B_{norm}(\tau_P) \leq B(\gamma)B(\gamma')B_{norm}(\tau)$$
3. *For all composable γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.*
4. $\forall U \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{U,*}} \frac{1}{|Aut(U)|c(\gamma)} < \epsilon'$
5. $\forall V \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)|c(\gamma)} < \epsilon'$
6. $\forall U \in \mathcal{I}_{mid}, \sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)|c(\tau)} < \epsilon'$
7. *For all $\tau \in \mathcal{M}'$, $\gamma \in \Gamma_{*,U_\tau} \cup \{Id_{U_\tau}\}$, and $\gamma' \in \Gamma_{*,V_\tau} \cup \{Id_{V_\tau}\}$,*
$$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i)\right)$$
$$\leq \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}}(|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

Remark 4.10. *Recall that if* $\Omega = N(0,1)$ *then we may take* $B_\Omega(j) = \sqrt{j}$ *and we have that*

$$h_j^+(x) \leq \frac{1}{\sqrt{j!}}(x^2+j)^{\frac{j}{2}} \leq \left(\frac{e}{j}(x^2+j)\right)^{\frac{j}{2}}$$

Thus, when $\Omega = N(0,1)$ *we can take*

$$B_{edge}(e) = 2\left(\frac{e}{l_e}(6D_V D_E + l_e)\right)^{l_e}(e(6D_V D_E q + 1))^{l_e} \leq \left(400 D_V^2 D_E^2 q\right)^{l_e}$$

4.2 Choosing $B_{norm}(\alpha)$

We need matrix norm bounds which hold for all $\alpha \in \mathcal{M}'$. For convenience, we recall the definition of $\mathcal{M}'$ below.

Definition 4.11 ($\mathcal{M}'$)**.** *We define* $\mathcal{M}'$ *to be the set of all shapes* α *such that*

1. $|V(\alpha)| \leq 3D_V$

*2.** $\forall e \in E(\alpha), l_e \leq D_E$

*3.** *All edges* $e \in E(\alpha)$ *have multiplicity at most* $3D_V$*.*

To obtain such norm bounds, we start with the norm bounds in the graph matrix norm bound paper. We then modify these bounds as follows:

1. We make the bounds more compatible with the conditions of our machinery. To do this, we upper bound many of the terms in the norm bound by $B_{vertex}^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|}$ where B_{vertex} is a function of our parameters. In general, we will also need to upper bound some of the terms by $\prod_{e\in E(\alpha)}(B_{edge}(e))$ where $B_{edge}(e)$ is a function of l_e, Ω, and our parameters.

2. We generalize the bounds so that they apply to improper shapes as well as proper shapes. Under our simplifying assumptions, all we need to do here is to take isolated vertices into account. In general, we also need to handle multi-edges.

4.2.1 Simplified $B_{norm}(\alpha)$

Under our simplifying assumptions, we start with the following norm bound from the updated graph matrix norm bound paper [2]:

Theorem 4.12 (Simplified Graph Matrix Norm Bounds)**.** *Under our simplifying assumptions, for all* $\epsilon > 0$ *and all proper shapes* α*, taking* $c_\alpha = |V(\alpha) \setminus (U_\alpha \cup V_\alpha)| + |S_\alpha \setminus (U_\alpha \cap V_\alpha)|$*,*

$$Pr\left(||M_\alpha|| > (2|V_\alpha \setminus (U_\alpha \cap V_\alpha)|)^{|V(\alpha)\setminus(U_\alpha\cap V_\alpha)|}(2eq)^{\frac{c_\alpha}{2}} n^{\frac{w(V(\alpha))-w(S_\alpha)}{2}}\right) < \epsilon$$

where $q = 3\left\lceil \frac{ln(\frac{n^{w(S_\alpha)}}{\epsilon})}{3c_\alpha}\right\rceil$

Corollary 4.13. *For all shapes α and all $\epsilon > 0$,*

$$Pr\left(||M_\alpha|| > \left(2|V_\alpha|\sqrt[4]{2eq}\right)^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha))+w(I_\alpha)-w(S_\alpha)}{2}}\right) < \epsilon$$

where $q = 3\left\lceil\frac{ln(\frac{n^{w(S_\alpha)}}{\epsilon})}{3c_\alpha}\right\rceil$.

Proof. Observe that adding an isolated vertex to α is equivalent to multiplying M_α by $n - |V(\alpha)|$. Thus, if the bound holds for all proper α then it will hold for improper α as well.

We now make the following observations:

1. $|S_\alpha \setminus (U_\alpha \cap V_\alpha)| \leq |U_\alpha \setminus V_\alpha|$, so $c_\alpha = |W_\alpha| + |S_\alpha \setminus (U_\alpha \cap V_\alpha)| \leq |V(\alpha) \setminus V_\alpha|$. Similarly, $|S_\alpha \setminus (U_\alpha \cap V_\alpha)| \leq |V_\alpha \setminus U_\alpha|$, so $c_\alpha \leq |V(\alpha) \setminus U_\alpha|$. Thus, $c_\alpha \leq \frac{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|}{2}$.

2. $|V(\alpha) \setminus (U_\alpha \cap V_\alpha)| \leq |V(\alpha) \setminus U_\alpha| + |V(\alpha) \setminus V_\alpha|$

Thus, by Theorem 4.12, for all proper shapes α and all $\epsilon > 0$,

$$Pr\left(||M_\alpha|| > \left(2|V_\alpha|\sqrt[4]{2eq}\right)^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha))+w(I_\alpha)-w(S_\alpha)}{2}}\right) < \epsilon''$$

where $q = 3\left\lceil\frac{ln(\frac{n^{w(S_\alpha)}}{\epsilon})}{3c_\alpha}\right\rceil$. □

Corollary 4.14. *For all $z \in \mathbb{N}$ and all $\epsilon > 0$, taking $\epsilon'' = \frac{\epsilon}{5^z 2^{z^2}}$, with probability at least $1 - \epsilon$ we have that for all shapes α such that $|V(\alpha)| \leq z$,*

$$||M_\alpha|| \leq \left(2|V_\alpha|\sqrt[4]{2eq}\right)^{|V(\alpha)\setminus U_\alpha|+|V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha))+w(I_\alpha)-w(S_\alpha)}{2}}$$

where $q = 3\left\lceil\frac{ln(\frac{n^{w(S_\alpha)}}{\epsilon''})}{3c_\alpha}\right\rceil$.

Proof. This result can be proved from Corollary 4.13 using a union bound and the following proposition:

Proposition 4.15. *Under our simplifying assumptions, for all $z \in \mathbb{N}$, there are at most $5^z 2^{z^2}$ proper shapes α such that $V(\alpha) \leq z$.*

Proof. Observe that we can construct any proper shape α with at most m vertices as follows:

1. Start with z vertices $v_1, \ldots, v_z$.

2. For each vertex v_i, choose whether $v_i \in V(\alpha) \setminus U_\alpha \setminus V_\alpha$, $v_i \in U_\alpha \setminus V_\alpha$, $v_i \in V_\alpha \setminus U_\alpha$, $v_i \in U_\alpha \cap V_\alpha$, or $v_i \notin V(\alpha)$.

3. For each pair of vertices $v_i, v_j \in V(\alpha)$, choose whether or not $(v_i, v_j) \in E(\alpha)$

□

□

Corollary 4.16. *For all $D_V \in \mathbb{N}$ and all $\epsilon > 0$, taking*

$$q = 3\left\lceil \frac{ln(\frac{5^{3D_V}2^{9D_V^2}n^{3D_V}}{\epsilon})}{3} \right\rceil = 3\left\lceil D_V ln(n) + \frac{ln(\frac{1}{\epsilon})}{3} + D_V ln(5) + 3D_V^2 ln(2) \right\rceil,$$

$B_{vertex} = 6D_V\sqrt[4]{2eq}$, and

$$B_{norm}(\alpha) = B_{vertex}^{|V(\alpha)\setminus U_\alpha| + |V(\alpha)\setminus V_\alpha|} n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_\alpha)}{2}},$$

with probability at least $(1-\epsilon)$ we have that for all shapes $\alpha \in \mathcal{M}'$, $||M_\alpha|| \leq B_{norm}(\alpha)$

Proof. This follows from Corollary 4.14 and the fact that for all $\alpha \in \mathcal{M}'$, $w(S_\alpha) \leq |V(\alpha)| \leq 3D_V$ □

4.2.2 General $B_{norm}(\alpha)$

In general, we start with the following norm bound from the updated graph matrix norm bound paper [2]:

Theorem 4.17 (General Graph Matrix Norm Bounds)**.** *For all $\epsilon > 0$ and all proper shapes α, taking $q = \lceil ln(\frac{n^{w(S_\alpha)}}{\epsilon}) \rceil$*

$$P\left(||M_\alpha|| > 2e(2q|V(\alpha)|)^{|V(\alpha)\setminus(U_\alpha\cap V_\alpha)|}\left(\prod_{e\in E(\alpha)} h^+_{l_e}(B_\Omega(2ql_e))\right) n^{\frac{(w(V(\alpha)) - w(S_\alpha))}{2}}\right) < \epsilon$$

Corollary 4.18. *For all $\epsilon > 0$, for all $z, l_{max}, m \in \mathbb{N}$, taking $\epsilon'' = \frac{\epsilon}{5^z(l_{max}+1)^{z^k}}$, with probability at least $1-\epsilon$, for all shapes α such that*

1. *$|V(\alpha)| \leq z$.*
2. *All edges in $E(\alpha)$ have label at most l_{max}.*
3. *All edges in $E(\alpha)$ have multiplicity at most m.*

,

$$||M_\alpha|| \leq 2e(2q|V(\alpha)|)^{|V(\alpha)\setminus U_\alpha| + |V(\alpha)\setminus V_\alpha|}\left(\prod_{e\in E(\alpha)} 2h^+_{l_e}(B_\Omega(2ml_{max})) \max_{j\in[0,ml_{max}]}\left\{\left(h^+_j(B_\Omega(2qj))\right)^{\frac{l_e}{\max\{j,l_e\}}}\right\}\right) n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_{\alpha,min})}{2}}$$

where $q = \left\lceil ln\left(\frac{n^{w(S_{\alpha,max})}}{\epsilon''}\right)\right\rceil$

Proof. Observe that for each α which has multi-edges, we can write $M_\alpha = \sum_i c_i M_{\alpha_i}$ where each α_i has no multiple edges. We first upper bound $\sum_i |c_i|$.

Lemma 4.19. *For any $a_1, \ldots, a_m \in \mathbb{N} \cup \{0\}$, taking $p_{max} = \sum_{i=1}^{m} a_i$ and writing $\prod_{i=1}^{m} h_{a_i} = \sum_{k=0}^{p_{max}} c_k h_k$,*

$$\sum_{k=0}^{p_{max}} |c_k| \leq (p_{max} + 1) \prod_{i=1}^{m} h^{+}_{a_i}(B_{\Omega}(2p_{max})) \leq \prod_{i=1}^{m} 2h^{+}_{a_i}(B_{\Omega}(2p_{max}))$$

Proof. To be added. □

Corollary 4.20. *For any shape α such that every edge of α has multiplicity at most m and label at most l_e, if we write $M_\alpha = \sum_i c_i M_{\alpha_i}$ where each α_i has no multi-edges then $\sum_i |c_i| \leq \prod_{e \in E(\alpha)} 2h^{+}_{l_e}(B_{\Omega}(2ml_{max}))$*

The result now follows from Theorem 4.17 and the following observations:

1. $|V(\alpha) \setminus (U_\alpha \cap V_\alpha)| \leq |V(\alpha) \setminus U_\alpha| + |V(\alpha) \setminus V_\alpha|$.
2. For any α, writing $M_\alpha = \sum_i c_i M_{\alpha_i}$ where each α_i has no multi-edges, for all α_i,

$$w(V(\alpha_i)) + w(I_{\alpha_i}) - w(S_{\alpha_i}) \leq w(V(\alpha)) + w(I_\alpha) - w(S_{\alpha,min})$$

3. For any $a_1, \ldots, a_m \in \mathbb{N} \cup \{0\}$ such that $\forall i' \in [m], a_{i'} \leq l_{max}$, for all $j \in [0, ml_{max}]$

$$h^{+}_{j}(B_{\Omega}(2qj)) \leq \prod_{i'=1}^{m} \left(h^{+}_{j}(B_{\Omega}(2qj))\right)^{\frac{a_{i'}}{\max\{j, a_{i'}\}}} \leq \prod_{i'=1}^{m} \max_{j' \in [0, ml_{max}]} \left\{ \left(h^{+}_{j'}(B_{\Omega}(2qj'))\right)^{\frac{a_{i'}}{\max\{j', a_{i'}\}}} \right\}$$

Proposition 4.21. *For all $z, l_{max} \in \mathbb{N}$, there are at most $5^z (l_{max} + 1)^{z^k}$ proper shapes α such that $|V(\alpha)| \leq z$ and every edge in $E(\alpha)$.*

Proof. This can be proved in the same way as before. Observe that we can construct any proper shape α with at most z vertices as follows:

1. Start with z vertices $v_1, \ldots, v_z$.
2. For each vertex v_i, choose whether $v_i \in V(\alpha) \setminus U_\alpha \setminus V_\alpha$, $v_i \in U_\alpha \setminus V_\alpha$, $v_i \in V_\alpha \setminus U_\alpha$, $v_i \in U_\alpha \cap V_\alpha$, or $v_i \notin V(\alpha)$.
3. For each k tuple of vertices in $V(\alpha)$, choose the label of the hyperedge between these vertices (or 0 if the hyperedge is not in $E(\alpha)$).

□

□

Corollary 4.22. *For all $D_V, D_E \in \mathbb{N}$ and all $\epsilon > 0$, taking*

$$B_{norm}(\alpha) = 2eB_{vertex}^{|V(\alpha) \setminus U_\alpha| + |V(\alpha) \setminus V_\alpha|} \left(\prod_{e \in E(\alpha)} B_{edge}(e) \right) n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_\alpha)}{2}}$$

where

1. $q = \left\lceil ln\left(\frac{n^{3D_V}}{\epsilon''}\right)\right\rceil = \lceil 3D_V ln(n) + ln(\frac{1}{\epsilon}) + (3D_V)^k ln(D_E+1) + 3D_V ln(5)\rceil$

2. $B_{vertex} = 6qD_V$

3. $B_{edge}(e) = 2h^+_{l_e}(B_\Omega(6D_V D_E)) \max_{j\in[0,3D_V D_E]}\left\{\left(h^+_j(B_\Omega(2qj))\right)^{\frac{l_e}{\max\{j,l_e\}}}\right\}$

with probability at least $(1-\epsilon)$*, for all shapes* $\alpha \in \mathcal{M}'$, $||M_\alpha|| \le B_{norm}(\alpha)$.

4.3 Choosing $B(\gamma)$

We now describe how to choose the function $B(\gamma)$. Recall that we want the following conditions to hold:

1. For all γ, τ, γ' and all intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,

$$B_{norm}(\tau_P) \le B(\gamma)B(\gamma')B_{norm}(\tau)$$

2. For all composable γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.

The most important part of choosing $B(\gamma)$ is to make sure that the factors of n are controlled. For this, we use the following intersection tradeoff lemma. Under our simplifying assumptions, this lemma follows from Lemma 7.12 of [1]. We defer the general proof of this lemma to the end of this section.

Lemma 4.23 (Intersection Tradeoff Lemma)**.** *For all* γ, τ, γ' *and all intersection patterns* $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,

$$w(V(\tau_P)) + w(I_{\tau_P}) - w(S_{\tau_P,min}) \le w(V(\tau)) + w(I_\tau) - w(S_{\tau,min}) + w(V(\gamma)\setminus U_\gamma) + w(V(\gamma')\setminus U_{\gamma'})$$

Based on this intersection tradeoff lemma, we can choose the function $B(\gamma)$ as follows.

Corollary 4.24. *If we take*

$$B_{norm}(\alpha) = C \cdot B_{vertex}^{|V(\alpha)\setminus U_\alpha| + |V(\alpha)\setminus V_\alpha|}\left(\prod_{e\in E(\alpha)} B_{edge}(e)\right) n^{\frac{w(V(\alpha)) + w(I_\alpha) - w(S_\alpha)}{2}}$$

for some constant $C > 0$ *and take*

$$B(\gamma) = B_{vertex}^{|V(\gamma)\setminus U_\gamma| + |V(\gamma)\setminus V_\gamma|}\left(\prod_{e\in E(\gamma)} B_{edge}(e)\right) n^{\frac{w(V(\gamma)\setminus U_\gamma)}{2}}$$

then the following conditions hold:

1. For all γ, τ, γ' *and all intersection patterns* $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,

$$B_{norm}(\tau_P) \le B(\gamma)B(\gamma')B_{norm}(\tau)$$

2. *For all composable* γ_1, γ_2, $B(\gamma_1)B(\gamma_2) = B(\gamma_1 \circ \gamma_2)$.

Proof. We have that

$$B_{norm}(\tau_P) = B_{vertex}^{|V(\tau_P) \setminus U_{\tau_P}| + |V(\tau_P) \setminus V_{\tau_P}|} \left(\prod_{e \in E(\tau_P)} B_{edge}(e) \right) n^{\frac{w(V(\tau_P)) + w(I_{\tau_P}) - w(S_{\tau_P})}{2}}$$

and

$$B(\gamma)B(\gamma')B_{norm}(\tau) = B_{vertex}^{|V(\gamma) \setminus U_\gamma| + |V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus U_{\gamma'}| + |V(\gamma') \setminus V_{\gamma'}| + |V(\tau) \setminus U_\tau| + |V(\tau) \setminus V_\tau|} \left(\prod_{e \in E(\gamma) \cup E(\gamma') \cup E(\tau)} B_{edge}(e) \right) n^{\frac{w(V(\gamma) \setminus U_\gamma) + w(V(\gamma') \setminus U_{\gamma'}) + w(V(\tau)) + w(I_\tau) - w(S_\tau)}{2}}$$

The first condition now follows immediately from the following observations:

1. $$|V(\gamma) \setminus U_\gamma| + |V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus U_{\gamma'}| + |V(\gamma') \setminus V_{\gamma'}| + |V(\tau) \setminus U_\tau| + |V(\tau) \setminus V_\tau|$$
$$= |V(\gamma \circ \tau \circ \gamma'^T) \setminus U_{\gamma \circ \tau \circ \gamma'^T}| + |V(\gamma \circ \tau \circ \gamma'^T) \setminus V_{\gamma \circ \tau \circ \gamma'^T}| \geq |V(\tau_P) \setminus U_{\tau_P}| + |V(\tau_P) \setminus V_{\tau_P}|$$

2. $E(\tau_P) = E(\gamma) \cup E(\tau) \cup E(\gamma'^T)$ so $\prod_{e \in E(\tau_P)} B_{edge}(e) = \prod_{e \in E(\gamma) \cup E(\gamma') \cup E(\tau)} B_{edge}(e)$.

3. By the intersection tradeoff lemma,
$$w(V(\tau_P)) + w(I_{\tau_P}) - w(S_{\tau_P}) \leq w(V(\tau)) + w(I_\tau) - w(S_\tau) + w(V(\gamma) \setminus U_\gamma) + w(V(\gamma') \setminus U_{\gamma'})$$

The second condition follows from the form of $B(\gamma)$. □

4.4 Choosing $N(\gamma)$

To choose $N(\gamma)$, we use the following lemma:

Lemma 4.25. *For all* $D_V \in \mathbb{N}$, *for all composable* γ, τ, γ'^T *such that* $|V(\gamma)| \leq D_V$, $|V(\tau)| \leq D_V$, *and* $|V(\gamma')| \leq D_V$,

$$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma, \gamma', j}} \prod_{i: \gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i: \gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, \gamma_i'^T}} \left(\prod_{i=1}^{j} N(P_i) \right)$$
$$\leq \frac{(3D_V)^{2(|V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus V_{\gamma'}|) + (|V(\gamma) \setminus U_\gamma| + |V(\gamma') \setminus U_{\gamma'}|)}}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

Proof sketch. Observe that aside from the orderings (which are canceled out by the $|Aut(U_{\gamma_i})|$ and $|Aut(U_{\gamma'_i})|$ factors), the intersection patterns $\{P_i : i \in [j]\}$ are determined by the following data on each vertex $v \in (V(\gamma) \setminus V_\gamma) \cup (V(\gamma'^T) \setminus V_{\gamma'^T})$:

1. The first $i \in [j]$ such that $v \in (V(\gamma_i) \setminus V_{\gamma_i}) \cup (V({\gamma'_i}^T) \setminus V_{{\gamma'_i}^T})$. There are at most j possibilities for this.

2. A vertex u (if one exists) in $V(\gamma_{i-1} \circ \ldots \circ \gamma_1 \circ \tau \circ {\gamma'_1}^T \ldots \circ {\gamma'_{i-1}}^T)$ such that u and v are equal. There are at most $3D_V$ possibilities for this.

Using these observations and taking $j_{max} = |V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus V_{\gamma'}|$,

$$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}} 1$$
$$\leq \sum_{j=1}^{j_{max}} \frac{(3jD_V)^{|V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus V_{\gamma'}|}}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$
$$\leq j_{max} \left(\frac{2}{3}\right)^{j_{max}} \frac{(3D_V)^{2(|V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus V_{\gamma'}|)}}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$
$$< \frac{(3D_V)^{2(|V(\gamma) \setminus V_\gamma| + |V(\gamma') \setminus V_{\gamma'}|)}}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

Now recall that by Lemma , for any $\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T$ and any intersection pattern $P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}$,

$$N(P_i) \leq |V(\tau_{P_i})|^{|V(\gamma_i) \setminus U_{\gamma_i}| + |V(\gamma'_i) \setminus U_{\gamma'_i}|}$$

Thus, for any $P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}$, $\prod_{i=1}^{j} N(P_i) \leq (3D_V)^{|V(\gamma) \setminus U_\gamma| + |V(\gamma') \setminus U_{\gamma'}|}$. Putting everything together, the result follows. □

Corollary 4.26. *For all $D_V \in \mathbb{N}$, if we take $N(\gamma) = (3D_V)^{2|V(\gamma) \setminus V_\gamma| + |V(\gamma) \setminus U_\gamma|}$ then for all composable $\gamma, \tau, {\gamma'}^T$ such that $|V(\gamma)| \leq D_V$, $|V(\tau)| \leq D_V$, and $|V(\gamma')| \leq D_V$,*

$$\sum_{j>0} \sum_{\gamma_1, \gamma'_1, \cdots, \gamma_j, \gamma'_j \in \Gamma_{\gamma,\gamma',j}} \prod_{i:\gamma_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma_i})|} \prod_{i:\gamma'_i \text{ is non-trivial}} \frac{1}{|Aut(U_{\gamma'_i})|} \sum_{P_1, \cdots, P_j : P_i \in \mathcal{P}_{\gamma_i, \tau_{P_{i-1}}, {\gamma'_i}^T}} \left(\prod_{i=1}^{j} N(P_i)\right)$$
$$\leq \frac{N(\gamma)N(\gamma')}{(|Aut(U_\gamma)|)^{1_{\gamma \text{ is non-trivial}}} (|Aut(U_{\gamma'})|)^{1_{\gamma' \text{ is non-trivial}}}}$$

4.5 Choosing $c(\alpha)$

In this section, we describe how to choose $c(\alpha)$. For simplicity, we first describe how to choose $c(\alpha)$ under our simplifying assumptions. We then describe the minor adjustments that are needed when we have hyperedges and multiple types of vertices.

Lemma 4.27. *Under our simplifying assumptons, for all $U \in \mathcal{I}_{mid}$,*

$$\sum_{\alpha : U_\alpha \equiv U, \alpha \text{ is proper and non-trivial}} \frac{1}{|Aut(U_\alpha \cap V_\alpha)|(3D_V)^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + 2|E(\alpha)|} 2^{|V(\alpha) \setminus (U_\alpha \cup V_\alpha)|}} < 5$$

Proof. In order to choose α, it is sufficient to choose the following:

1. The number j_1 of vertices in $U_\alpha \setminus V_\alpha$, the number j_2 of vertices in $V_\alpha \setminus U_\alpha$, and the number j_3 of vertices in $V(\alpha) \setminus (U_\alpha \cup V_\alpha)$.
2. A mapping in $Aut(U_\alpha \cap V_\alpha)$ determining how the vertices in $U_\alpha \cap V_\alpha$ match up with each other.
3. The position of each vertex $u \in U_\alpha \setminus V_\alpha$ within U_α (there are at most $|U_\alpha| \leq D_V$ choices for this).
4. The position of each vertex $v \in V_\alpha \setminus U_\alpha$ within V_α (there are at most $|U_\alpha| \leq D_V$ choices for this).
5. The number j_4 of edges in $E(\alpha)$.
6. The endpoints of each edge in $E(\alpha)$.

This implies that for all $j_1, j_2, j_3, j_4 \geq 0$

$$\sum_{\substack{\alpha: U_\alpha \equiv U, |U_\alpha \setminus V_\alpha| = j_1, |V_\alpha \setminus U_\alpha| = j_2 \\ |V(\alpha) \setminus (U_\alpha \cup V_\alpha)| = j_3, |E(\alpha)| = j_4}} \frac{1}{|Aut(U_\alpha \cap V_\alpha)|(D_V)^{j_1+j_2}(D_V)^{2j_4}} \leq 1$$

Using this, we have that

$$\sum_{\alpha: U_\alpha \equiv U, \alpha \text{ is proper and non-trivial}} \frac{1}{|Aut(U_\alpha \cap V_\alpha)|(3D_V)^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + 2|E(\alpha)|} 2^{|V(\alpha) \setminus (U_\alpha \cup V_\alpha)|}}$$

$$\leq \sum_{j_1, j_2, j_3, j_4 \in \mathbb{N} \cup \{0\}: j_1 + j_2 + j_3 + j_4 \geq 1} \frac{1}{3^{j_1+j_2} 9^{j_4} 2^{j_3}} \leq 2 \left(\frac{3}{2}\right)^2 \frac{9}{8} - 1 < 5$$

□

Corollary 4.28. *For all $\epsilon' > 0$, if we take*

$$c(\alpha) = \frac{\epsilon'}{5(3D_V)^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + 2|E(\alpha)|} 2^{|V(\alpha) \setminus (U_\alpha \cup V_\alpha)|}}$$

then

1. $\forall U \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{U,*}} \frac{1}{|Aut(U)| c(\gamma)} < \epsilon'$
2. $\forall V \in \mathcal{I}_{mid}, \sum_{\gamma \in \Gamma_{*,V}} \frac{1}{|Aut(U_\gamma)| c(\gamma)} < \epsilon'$
3. $\forall U \in \mathcal{I}_{mid}, \sum_{\tau \in \mathcal{M}_U} \frac{1}{|Aut(U)| c(\tau)} < \epsilon'$

4.5.1 Choosing $c(\alpha)$ in general*

When we have multiple types of vertices and hyperedges of arity k, Lemma 4.29 can be generalized as follows:

Lemma 4.29. *Under our simplifying assumptons, for all $U \in \mathcal{I}_{mid}$,*

$$\sum_{\alpha: U_\alpha \equiv U, \alpha \text{ is proper and non-trivial}} \frac{1}{|Aut(U_\alpha \cap V_\alpha)|(3D_V t_{max})^{|U_\alpha \setminus V_\alpha| + |V_\alpha \setminus U_\alpha| + k|E(\alpha)|}(2t_{max})^{|V(\alpha) \setminus (U_\alpha \cup V_\alpha)|}} < 5$$

Proof sketch. This can be proved in the same way as Lemma 4.29 with the following modifications:

1. In addition to choosing the number of vertices in $U_\alpha \setminus V_\alpha$, $V_\alpha \setminus U_\alpha$, and $V(\alpha) \setminus (U_\alpha \cap V_\alpha)$, we also have to choose the types of these vertices.

2. For each hyperedge, we have to choose k endpoints rather than 2 endpoints.

□

4.6 Proof of the Generalized Intersection Tradeoff Lemma

We now prove the generalized intersection tradeoff lemma.

Lemma 4.30. *For all γ, τ, γ' and all intersection patterns $P \in \mathcal{P}_{\gamma,\tau,\gamma'}$,*

$$w(V(\tau_P)) + w(I_{\tau_P}) - w(S_{\tau_P,min}) \leq w(V(\tau)) + w(I_\tau) - w(S_{\tau,min}) + w(V(\gamma) \setminus U_\gamma) + w(V(\gamma') \setminus U_{\gamma'})$$

Proof.

Definition 4.31.

1. *We define I_{LM} to be the set of vertices which, after intersections, touch γ and τ but not ${\gamma'}^T$. In particular, I_{LM} consists of the vertices which result from intersecting a pair of vertices in $V(\gamma) \setminus V_\gamma$ and $V(\tau) \setminus U_\tau \setminus V_\tau$ and the vertices which are in $U_\tau \setminus V_\tau$ and are not intersected with any other vertex.*

2. *We define I_{MR} to be the set of vertices which, after intersections, touch τ and ${\gamma'}^T$ but not γ. In particular, I_{MR} consists of the vertices which result from intersecting a pair of vertices in $V(\tau) \setminus U_\tau \setminus V_\tau$ and $V({\gamma'}^T) \setminus U_{{\gamma'}^T}$ and the vertices which are in $V_\tau \setminus U_\tau$ and are not intersected with any other vertex.*

3. *We define I_{LR} to be the set of vertices which, after intersections, touch γ and ${\gamma'}^T$ but not τ. In particular, I_{LR} consists of the vertices which result from intersecting a pair of vertices in $V(\gamma) \setminus V_\gamma$ and $V({\gamma'}^T) \setminus U_{{\gamma'}^T}$.*

4. *We define I_{LMR} to be the set of vertices which, after intersections, touch γ, τ, and ${\gamma'}^T$. In particular, I_{LMR} consists of the vertices which result from intersecting a triple of vertices in $V(\gamma) \setminus V_\gamma$, $V(\tau) \setminus U_\tau \setminus V_\tau$, and $V({\gamma'}^T) \setminus U_{{\gamma'}^T}$, intersecting a pair of vertices in $V(\gamma) \setminus V_\gamma$ and $V_\tau \setminus U_\tau$, intersecting a pair of vertices in $U_\tau \setminus V_\tau$ and $V({\gamma'}^T) \setminus U_{{\gamma'}^T}$, and single vertices in $U_\tau \cap V_\tau$.*

The main idea is as follows. A priori, any of the vertices in $I_{LM} \cup I_{MR} \cup I_{LR} \cup I_{LMR}$ could become isolated. We handle this by keeping track of the following types of flows:

1. Flows from U_γ to $I_{LM} \cup I_{LR} \cup I_{LMR}$
2. Flows from $I_{LR} \cup I_{MR} \cup I_{LMR}$ to $V_{\gamma'^T}$
3. Flows from I_{LM} to I_{MR}. For technical reasons, we also view vertices in I_{LMR} as having flow to themselves.

We then observe that flows to and from these vertices prevent these vertices from being isolated and can provide flow from U_γ to $V_{\gamma'^T}$, which gives a lower bound on $w(S_{\tau_P})$.

We now implement this idea.

Definition 4.32 (Flow Graph)**.** *Given a shape α, we define the directed graph H_α as follows:*

1. *For each vertex $v \in V(\alpha)$, we create two vertices v_{in} and v_{out}. We then create a directed edge from v_{in} to v_{out} with capacity $w(v)$*
2. *For each pair of vertices (v, w) which is an edge of multiplicity 1 in $E(\alpha)$ (or part of a hyperedge of multiplicity 1 in $E(\alpha)$), we create a directed edge with infinite capacity from v_{out} to w_{in} and we create a directed edge with infinite capacity from w_{out} to v_{in}.*
3. *We define U_{H_α} to be $U_{H_\alpha} = \{u_{in} : u \in U_\alpha\}$ and we define V_{H_α} to be $V_{H_\alpha} = \{v_{out} : v \in V_\alpha\}$*

Lemma 4.33. *The maximum flow from U_{H_α} to V_{H_α} is equal to the minimum weight of a separator between U_α and V_α.*

Proof. This can be proved using the max flow min cut theorem. □

Definition 4.34 (Modified Flow Graph)**.** *Given a shape α together with a set $I_L \subseteq V(\alpha)$ of vertices in α (which will be the vertices in α which are intersected with a vertex to the left of α) and a set $I_R \subseteq V(\alpha)$ of vertices in α (which will be the vertices in α which are intersected with a vertex to the right of α), we define the modified flow graph $H_\alpha^{I_L,I_R}$ as follows:*

1. *We start with the flow graph H_α*
2. *For each vertex $u \in I_L$, we delete all of the edges into u_{in} and add u_{in} to U_{H_α}*
3. *For each vertex $v \in I_R$, we delete all of the edges out of v_{out} and add v_{out} to V_{H_α}*
4. *We call the resulting graph $H_\alpha^{I_L,I_R}$ and the resulting sets $U_{H_\alpha^{I_L,I_R}}$ and $V_{H_\alpha^{I_L,I_R}}$*

Lemma 4.35. *The maximum flow from $U_{H_\alpha^{I_L,I_R}}$ to $V_{H_\alpha^{I_L,I_R}}$ in $H_\alpha^{I_L,I_R}$ is at least as large as the maximum flow from U_{H_α} to V_{H_α} in H_α*

Proof sketch. Observe that if we have a cut C in $H_\alpha^{I_L,I_R}$ which separates $U_{H_\alpha^{I_L,I_R}}$ and $V_{H_\alpha^{I_L,I_R}}$ then C separates U_{H_α} and V_{H_α} in H_α □

Before the intersections, we have the following flows.

1. We take F_1 to be the maximum flow from U_γ to V_γ in γ. Note that F_1 has value $w(V_\gamma)$

2. We take F_2 to be the maximum flow from U_τ to V_τ in τ. Note that F_2 has value $w(S_{\tau,min})$

3. We take F_3 to be the maximum flow from $U_{\gamma'^T}$ to $V_{\gamma'^T}$ in ${\gamma'}^T$. Note that F_1 has value $w(U_{\gamma'^T})$

After the intersections, we take the following flows:

1. We take F_1' to be the maximum flow from $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ to $V_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ in $H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}$.

2. We take F_2' to be the maximum flow from $U_{H_\tau^{I_{LM}\cup I_{LMR},I_{MR}\cup I_{LMR}}}$ to $V_{H_\tau^{I_{LM}\cup I_{LMR},I_{MR}\cup I_{LMR}}}$ in $H_\tau^{I_{LM}\cup I_{LMR},I_{MR}\cup I_{LMR}}$

3. We take F_3' to be the maximum flow from $U_{H_{\gamma'^T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$ to $V_{H_{\gamma'^T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$ in $H_{\gamma'^T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}$.

Observe that because of how intersection patterns are defined, $val(F_1') = w(U_\gamma)$ and $val(F_3') = w(V_{\gamma'^T})$. By Lemma 4.35, the value of F_2' is at least as large as the value of F_2, so $val(F_2') \geq w(S_{\tau,min})$.

We now consider $F_1' + F_2' + F_3'$. As is, this is not a flow, but we can fix this.

Definition 4.36. *For each vertex $v \in V(\tau_P)$,*

1. *We define $f_{in}(v)$ to be the flow into v_{in} in $F_1' + F_2' + F_3'$.*

2. *We define $f_{out}(v)$ to be the flow out of v_{out} in $F_1' + F_2' + F_3'$.*

3. *We define $f_{through}(v)$ to be the flow from v_{in} to v_{out} in $F_1' + F_2' + F_3'$*

4. *We define $f_{imbalance}(v)$ to be $f_{imbalance}(v) = |f_{in}(v) - f_{out}(v)|$*

5. *We define $f_{excess}(v)$ to be $f_{excess}(v) = f_{through}(v) - max\{f_{in}(v), f_{out}(v)\}$*

With this information, we fix the flow $F_1' + F_2' + F_3'$ as follows. For each vertex $v \in V(\tau_P)$,

1. *If $f_{in}(v) > f_{out}(v)$ then we create a vertex $v_{supplemental,out}$ and an edge from v_{out} to $v_{supplemental,out}$ with capacity $f_{imbalance}(v)$ and we route $f_{imbalance}(v)$ of flow along this edge. We then add $v_{supplemental,out}$ to a set of vertices $V_{supplemental}$.*

2. *If $f_{in}(v) < f_{out}(v)$ then we create a vertex $v_{supplemental,in}$ and an edge from $v_{supplemental,in}$ to v_{in} with capacity $f_{imbalance}(v)$ and we route $f_{imbalance}(v)$ of flow along this edge. We then add $v_{supplemental,out}$ to a set of vertices $V_{supplemental}$.*

3. *We reduce the flow on the edge from v_{in} to v_{out} by $f_{excess}(v)$*

We call the resulting flow F'

Proposition 4.37. *F' is a flow from $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}} \cup U_{supplemental}$ to $V_{H_{\gamma'^T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}} \cup V_{supplemental}$ with value $val(F') = val(F_1') + val(F_2') + val(F_3') - \sum_{v \in V(\tau)} f_{excess}(v)$*

Corollary 4.38. *There exists a flow F'' from $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ to $V_{H_{\gamma' T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$ with value $val(F'') \geq val(F'_1) + val(F'_2) + val(F'_3) - \sum_{v\in V(\tau)} (f_{excess}(v) + f_{imbalance}(v))$*

Proof. Consider the minimum cut C between $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ and $V_{H_{\gamma' T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$. If we add all of the supplemental edges to C then this gives a cut C' between $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ and $V_{H_{\gamma' T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$ with capacity

$$capacity(C') = capacity(C) + \sum_{v\in V(\tau)} f_{imbalance}(v) \geq val(F')$$

Thus, $capacity(C) \geq val(F') - \sum_{v\in V(\tau)} f_{imbalance}(v)$ so there exists a flow F'' from $U_{H_\gamma^{\emptyset,I_{LM}\cup I_{LR}\cup I_{LMR}}}$ to $V_{H_{\gamma' T}^{I_{MR}\cup I_{LR}\cup I_{LMR},\emptyset}}$ with value

$$val(F'') = capacity(C) \geq val(F'_1) + val(F'_2) + val(F'_3) - \sum_{v\in V(\tau)} (f_{excess}(v) + f_{imbalance}(v))$$

□

We now make the following observations:

Lemma 4.39.

1. *For all vertices $v \notin I_{LM} \cup I_{MR} \cup I_{LR} \cup I_{LMR}$, $f_{excess}(v) = f_{imbalance}(v) = 0$ (and these vertices can never be isolated).*

2. *For all vertices $v \in I_{LM}$, $f_{excess}(v) + f_{imbalance}(v) \leq w(v)$. Moreover, for all vertices $v \in I_{LM}$ which are isolated, $f_{excess}(v) = f_{imbalance}(v) = 0$.*

3. *For all vertices $v \in I_{MR}$, $f_{excess}(v) + f_{imbalance}(v) \leq w(v)$. Moreover, for all vertices $v \in I_{LM}$ which are isolated, $f_{excess}(v) = f_{imbalance}(v) = 0$.*

4. *For all vertices $v \in I_{LR}$, $f_{excess}(v) + f_{imbalance}(v) \leq w(v)$. Moreover, for all vertices $v \in I_{LM}$ which are isolated, $f_{excess}(v) = f_{imbalance}(v) = 0$.*

5. *For all vertices $v \in I_{LMR}$, $f_{excess}(v) + f_{imbalance}(v) \leq 2w(v)$. Moreover, for all vertices $v \in I_{LMR}$ which are isolated, $f_{excess}(v) = w(v)$ and $f_{imbalance}(v) = 0$.*

Proof. For the first statement, observe that for vertices $v \notin I_{LM} \cup I_{MR} \cup I_{LR} \cup I_{LMR}$, neither v_{in} nor v_{out} is ever a sink or source so the flow into these vertices must equal the flow out of these vertices and thus $f_{in}(v) = f_{out}(v) = f_{through}(v)$.

For the second statement, observe that for a vertex $v \in I_{LM}$,

1. F'_1 will have a flow of $f_{in}(v)$ into v_{in} and along the edge from v_{in} to v_{out}

2. F'_2 will have a flow of $f_{out}(v)$ along the edge from v_{in} to v_{out} and out of v_{out}.

Thus, $f_{excess}(v) = f_{in}(v)+f_{out}(v)-\max\{f_{in}(v), f_{out}(v)\}$. Since $f_{imbalance}(v) = |f_{in}(v)-f_{out}(v)|$, $f_{excess}(v) + f_{imbalance}(v) = f_{in}(v) + f_{out}(v) - \min\{f_{in}(v), f_{out}(v)\} \leq w(v)$.

If v is isolated then neither F'_1 nor F'_2 can have any flow to v_{in} or out of v_{out} so $f_{in}(v) = f_{through}(v) = f_{out}(v) = 0$

The third and fourth statements can be proved in the same way as the second statement.

For the fifth statement, observe that for a vertex $v \in I_{LMR}$,

1. F'_1 will have a flow of $f_{in}(v)$ into v_{in} and along the edge from v_{in} to v_{out}.
2. F'_2 will have a flow of $w(v)$ along the edge from v_{in} to v_{out}
3. F'_3 will have a flow of $f_{out}(v)$ along the edge from v_{in} to v_{out} and out of v_{out}.

Thus, $f_{excess}(v) = w(v)+f_{in}(v)+f_{out}(v)-\max\{f_{in}(v), f_{out}(v)\}$. Since $f_{imbalance}(v) = |f_{in}(v)-f_{out}(v)|$, $f_{excess}(v) + f_{imbalance}(v) = w(v) + f_{in}(v) + f_{out}(v) - \min\{f_{in}(v), f_{out}(v)\} \leq 2w(v)$.

If v is isolated then neither F'_1 nor F'_3 can have any flow to v_{in} or out of v_{out} so $f_{in}(v) = f_{out}(v) = 0$ and $f_{through}(v) = w(v)$. □

Putting everything together, we have the following corollary:

Corollary 4.40.

$$\sum_{v \in V(\tau_P)} (f_{excess}(v) + f_{imbalance}(v)) \leq w(I_{LM})+w(I_{LR})+w(I_{MR})+2w(I_{LMR})-(w(I_{\tau_P})-w(I_\tau))$$

Combining this with Corollary 4.38,

$$\begin{aligned} w(S_{\tau_P,min}) &\geq val(F'_1) + val(F'_2) + val(F'_3) - \sum_{v \in V(\tau_P)} (f_{excess}(v) + f_{imbalance}(v)) \\ &\geq w(U_\gamma) + w(S_{\tau,min}) + w(V_{\gamma'^T}) - w(I_{LM}) - w(I_{LR}) - w(I_{MR}) - 2w(I_{LMR}) + (w(I_{\tau_P}) - w(I_\tau)) \end{aligned}$$

Since $w(V(\tau_P)) = w(V(\tau))+w(V(\gamma))+w(V(\gamma'))-w(I_{LM})-w(I_{LR})-w(I_{MR})-2w(I_{LMR})$,

$$w(S_{\tau_P,min}) \geq w(U_\gamma)+w(S_{\tau,min})+w(V_{\gamma'^T})+w(V(\tau_P))-w(V(\tau))-w(V(\gamma))-w(V(\gamma'))+(w(I_{\tau_P})-w(I_\tau))$$

Rearranging this gives

$$w(V(\tau_P))-w(S_{\tau_P,min})+w(I_{\tau_P}) \leq w(V(\tau))-w(S_{\tau,min})+w(I_\tau)+w(V(\gamma)\setminus U_\gamma)+w(V(\gamma')\setminus U_{\gamma'})$$

which is the generalized intersection tradeoff lemma. □

5 Showing Positivity

In this section, we describe how to show that $\sum_{V \in \mathcal{I}_{mid}} M^{fact}(H_{Id_V}) \succeq \delta Id_{Sym}$ for some $\delta > 0$ where δ will depend on n and other parameters. For now, we assume that the indices of Λ are multilinear monomials. We will then describe the adjustments that are needed to handle non-multilinear matrix indices.

We start with a few more definitions.

Definition 5.1. *For all $V \in \mathcal{I}_{mid}$ we define $Id_{Sym,V}$ to be the matrix such that*

1. $Id_{Sym,V}(A,B) = 1$ *if A and B both have index shape V.*
2. *Otherwise, $Id_{Sym,V}(A,B) = 0$.*

Proposition 5.2. $Id_{Sym} = \sum_{V \in \mathcal{I}_{mid}} Id_{Sym,V}$

Definition 5.3. *For all $V \in \mathcal{I}_{mid}$ we define $\lambda_V = H_{Id_V}(Id_V, Id_V)$*

We now describe our strategy for showing $\sum_{V \in \mathcal{I}_{mid}} M^{fact}(H_{Id_V}) \succeq \delta Id_{Sym}$. The idea is as follows. We will consider the index shapes $V \in \mathcal{I}_{mid}$ from largest weight to smallest weight and we will show that for each $V \in \mathcal{I}_{mid}$, there exists a $\delta_V > 0$ such that $\sum_{V \in \mathcal{I}_{mid}} M^{fact}(H_{Id_V}) \succeq \delta_V \sum_{U \in \mathcal{I}_{mid}: w(U) \geq w(V)} Id_{Sym,U}$.

For the first step, letting V_{max} be the maximum weight index shape in $\mathcal{I}_{mid}$, $M^{fact}(H_{Id_{V_{max}}}) = \lambda_{V_{max}} Id_{Sym,V_{max}}$ because there are no non-trivial left shapes σ such that $V_\sigma = V_{max}$. For other $V \in \mathcal{I}_{mid}$, $\lambda_V Id_{Sym,V}$ is a part of $M^{fact}(H_{Id_V})$ but $M^{fact}(H_{Id_V})$ will also contain terms of the form $H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T}$ where $U_\sigma \neq V$ or $U_{\sigma'} \neq V$.

We can handle the terms $H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T}$ where $U_\sigma \neq V$ and $U_{\sigma'} \neq V$ by bounding these terms in terms of Id_{Sym,U_σ} and $Id_{Sym,U_{\sigma'}}$. Since $w(U_\sigma) > w(V)$ and $w(U_{\sigma'}) > w(V)$, Id_{Sym,U_σ} and $Id_{Sym,U_{\sigma'}}$ are already available to us. To handle the terms $H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T}$ where exactly one of U_σ and $U_{\sigma'}$ are equal to V, we use the following trick.

Definition 5.4. *Given $V \in \mathcal{I}_{mid}$, define H''_{Id_V} to be the coefficient matrix such that*

1. *If $U_\sigma = U_{\sigma'} = V$ then $H''_{Id_V}(\sigma, \sigma') = \frac{1}{2} H_{Id_V}(\sigma, \sigma')$*
2. *If exactly one of U_σ and $U_{\sigma'}$ are equal to V then $H''_{Id_V}(\sigma, \sigma') = H_{Id_V}(\sigma, \sigma')$*
3. *If $U_\sigma \neq V$ and $U_{\sigma'} \neq V$ then $H''_{Id_V}(\sigma, \sigma') = 2H_{Id_V}(\sigma, \sigma')$*

Proposition 5.5. $M^{fact}(H''_{Id_V}) \succeq 0$

Proof. Since $H_{Id_V} \succeq 0$, $H''_{Id_V} \succeq 0$ and thus $M^{fact}(H''_{Id_V}) \succeq 0$. □

Corollary 5.6. *For all $V \in \mathcal{I}_{mid}$,*

$$M^{fact}(H_{Id_V}) + \sum_{\sigma, \sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T} \succeq \frac{\lambda_V}{2} Id_{Sym,V}$$

Proof. Observe that

$$M^{fact}(H_{Id_V}) - \frac{\lambda_V}{2} Id_{Sym,V} + \sum_{\sigma, \sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T} = M^{fact}(H''_{Id_V}) \succeq 0$$

□

We now analyze the terms $\sum_{\sigma, \sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} H_{Id_V}(\sigma, \sigma') M_\sigma M_{\sigma'^T}$.

Definition 5.7. *Given $U, V \in \mathcal{I}$ with $w(U) > w(V)$, we define $W(U,V)$ to be*

$$W(U,V) = \frac{1}{|Aut(U)|} \sum_{\sigma \in \mathcal{L}_V : U_\sigma = U} \sum_{\sigma' \in \mathcal{L}_V : U_{\sigma'} \neq V} B_{norm}(\sigma) B_{norm}(\sigma') H_{Id_U}(\sigma, \sigma')$$

Lemma 5.8. *For all $V \in \mathcal{I}_{mid}$,*

$$\sum_{\sigma,\sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} H_{Id_U}(\sigma,\sigma') M_\sigma M_{\sigma'^T} \preceq \sum_{U \in \mathcal{I}_{mid} : w(U) > w(V)} W(U,V) Id_{Sym,U}$$

Proof. Observe that for all $\sigma, \sigma' \in \mathcal{L}_V$ such that $U_\sigma \neq V$ and $U_{\sigma'} \neq V$, $||M_\sigma M_{\sigma'^T}|| \leq B_{norm}(\sigma) B_{norm}(\sigma')$ and thus

$$\frac{1}{2}\left(M_\sigma M_{\sigma'^T} + M_{\sigma'} M_{\sigma^T}\right) \preceq \frac{1}{2} B_{norm}(\sigma) B_{norm}(\sigma') \left(M_{Id_{U_\sigma}} + M_{Id_{U_{\sigma'}}}\right)$$

Summing this equation over all $\sigma, \sigma' \in \mathcal{L}_V$ such that $U_\sigma \neq V$ and $U_{\sigma'} \neq V$,

$$\begin{aligned}
&\sum_{\sigma,\sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} H_{Id_U}(\sigma,\sigma') M_\sigma M_{\sigma'^T} \preceq \sum_{\sigma,\sigma' \in \mathcal{L}_V : U_\sigma \neq V, U_{\sigma'} \neq V} B_{norm}(\sigma) B_{norm}(\sigma') M_{Id_{U_\sigma}} \\
&\preceq \sum_{U \in \mathcal{I}_{mid} : w(U) > w(V)} \sum_{\sigma \in \mathcal{L}_V : U_\sigma = U} \sum_{\sigma' \in \mathcal{L}_V : V_\sigma \neq V} B_{norm}(\sigma) B_{norm}(\sigma') H_{Id_U}(\sigma,\sigma') M_{Id_U} \\
&\preceq \sum_{U \in \mathcal{I}_{mid} : w(U) > w(V)} |Aut(U)| W(U,V) M_{Id_U}
\end{aligned}$$

Since all of the coefficient matrices have SOS-symmetry, we can replace M_{Id_U} by $\frac{1}{|Aut(U)|} Id_{Sym,U}$ and this completes the proof. □

Using this lemma, we can show the following theorem:

Theorem 5.9. *Let G be the following directed graph:*

1. *The vertices of G are the index shapes $V \in \mathcal{I}_{mid}$*
2. *For each $U, V \in \mathcal{I}_{mid}$ such that $w(U) > w(V)$, we have an edge $e = (V, U)$ with weight $w(e) = \frac{2W(U,V)}{\lambda_V}$*

For all $V \in \mathcal{I}_{mid}$,

$$Id_{Sym,V} \preceq 2 \sum_{U \in \mathcal{I}_{mid} : w(U) \geq w(V)} \left(\sum_{P : P \text{ is a path from } V \text{ to } U \text{ in } G} \prod_{e \in E(P)} w(e) \right) M^{fact}(H_{Id_U})$$

Proof sketch. This can be proved by iteratively applying Lemma 5.8. □

5.0.1 Handling Non-multilinear Matrix Indices*

To be added.

References

[1] B. Barak, S. Hopkins, J. Kelner, P. Kothari, A. Moitra, A. Potechin. A Nearly Tight Sum-of-Squares Lower Bound for the Planted Clique Problem. FOCS 2016

[2] K. Ahn, D. Medarametla, A. Potechin. Graph Matrices: Norm Bounds and Applications.

A Proof that the Leftmost and Rightmost Minimum Vertex Separators are Well-defined

In this section, we give a general proof that the leftmost and rightmost minimum vertex separators are well-defined.

Lemma A.1. *For any two distinct vertex separators S_1 and S_2 of α, there exist vertex separators S_L and S_R of α such that:*

1. *S_L is a vertex separator of U_α and S_1 and a vertex separator of U_α and S_2.*
2. *S_R is a vertex separator of S_1 and V_α and a vertex separator of S_2 and V_α.*
3. *$w(S_L) + w(S_R) \leq w(S_1) + w(S_2)$*

Proof. Take S_L to be the set of vertices $v \in V(\alpha) \cap (S_1 \cup S_2)$ such that there is a path from U_α to v which doesn't intersect $S_1 \cup S_2$ before reaching v. Similarly, take S_R to be the set of vertices $v \in V(\alpha) \cap (S_1 \cup S_2)$ such that there is a path from V_α to v which doesn't intersect $S_1 \cup S_2$ before reaching v.

Now observe that S_L is a vertex separator between U_α and S_1. To see this, note that for any path P from U_α to a vertex $v \in S_1$, either P intersects S_L before reaching v or P does not intersect S_L before reaching v. In the latter case, $v \in S_L$. Thus, in either case, P intersects S_L. Following similar logic, S_L is also a vertex separator between U_α and S_2, S_R is a vertex separator between S_1 and V_α, and S_R is also a vertex separator between S_2 and V_α.

To show that $w(S_L) + w(S_R) \leq w(S_1) + w(S_2)$, observe that $w(S_L) + w(S_R) = w(S_R \cup S_R) + w(S_L \cap S_R)$ and $w(S_1) + w(S_2) = w(S_1 \cup S_2) + w(S_1 \cap S_2)$. Thus, to show that $w(S_L) + w(S_R) \leq w(S_1) + w(S_2)$, it is sufficient to show that

1. $S_L \cup S_R \subseteq S_1 \cup S_2$
2. $S_L \cap S_R \subseteq S_1 \cap S_2$

For the first statement, note that by definition any vertex in $S_L \cup S_R$ must be in $S_1 \cup S_2$. For the second statement, note that if $v \in S_L \cap S_R$ then there is a path from U_α to v which does not intersect any other vertices in $S_1 \cup S_2$ and there is a path from v to V_α which does not intersect any other vertices in $S_1 \cup S_2$. Combining these paths, we obtain a path P from U_α to V_α such that v is the only vertex in P which is in $S_1 \cup S_2$. This implies that $v \in S_1 \cap S_2$ as otherwise either S_1 or S_2 would not be a vertex separator between U_α and V_α. □

Corollary A.2. *The leftmost and rightmost minimum vertex separators between U_α and V_α are well-defined.*

Proof. Assume that there is no minimum leftmost vertex separator. If so, then there exists a minimum vertex separator S_1 between U_α and V_α such that

1. There does not exist a minimum vertex separator S' of α such that S' is also a minimum vertex separator of U_α and S_1 (otherwise we would take S' rather than S)

2. There exists a minimum vertex separator S_2 of α such that S' is not a minimum vertex separator of U_α and S_2 (as otherwise S_1 would be the leftmost minimum vertex separator)

Now let S_L and S_R be the vertex separators of α obtained by applying Lemma A.1 to S_1 and S_2. Since S_1 and S_2 are minimum vertex separators of α, we must have that $w(S_L) = w(S_R) = w(S_1) = w(S_2)$. Since S_L is a vertex separator of U_α and S_2, $S_L \neq S_1$. However, S_L is a vertex separator of U_α and S_1, which contradicts our choice of S_1.

Thus, there must be a leftmost minimum vertex separator of α. Following similar logic, there must be a rightmost minimum vertex separator of α as well. □

B Proofs with Canonical Maps

In this section, we give alternative proofs of Lemmas 2.78 and 3.14 using canonical maps.

Definition B.1 (Canonical Maps)**.** *For each shape α and each ribbon R of shape α, we arbitrarily choose a canonical map $\phi_R : V(\alpha) \to V(R)$ such that $\phi_R(H_\alpha) = H_R$, $\phi_R(U_\alpha) = A_R$, and $\phi_R(V_\alpha) = B_R$. Note that there are $|Aut(\alpha)|$ possible choices for this map.*

B.1 Proof of Lemma 2.78

Lemma B.2.

$$M^{orth}_\tau(H) = \sum_{\sigma \in Row(H), \sigma' \in Col(H)} H(\sigma, \sigma')|Decomp(\sigma, \tau, \sigma'^T)|M_{\sigma \circ \tau \circ \sigma'^T}$$

Proof. Observe that there is a bijection between ribbons R with shape $\sigma \circ \tau \circ \sigma'^T$ together with an element $\pi \in Decomp(\sigma, \tau, \sigma')$ and triples of ribbons (R_1, R_2, R_3) such that

1. R_1, R_2, R_3 have shapes σ, τ, and σ'^T, respectively.

2. $V(R_1) \cap V(R_2) = A_{R_2} = B_{R_1}$, $V(R_2) \cap V(R_3) = A_{R_3} = B_{R_2}$, and $V(R_1) \cap V(R_3) = A_{R_2} \cap B_{R_2}$

To see this, note that given such ribbons R_1, R_2, R_3, the ribbon $R = R_1 \circ R_2 \circ R_3$ has shape $\sigma \circ \tau \circ \sigma'^T$. Further note that we have two bijective maps from $V(\sigma \circ \tau \circ \sigma'^T)$ to $V(R)$. The first map is ϕ_R. The second map is $\phi_{R_1} \circ \phi_{R_2} \circ \phi_{R_3}$. Using this, we can take $\pi = \phi_R^{-1}(\phi_{R_1} \circ \phi_{R_2} \circ \phi_{R_3})$

Conversely, given a ribbon R of shape $\sigma \circ \tau \circ \sigma'^T$ and an element $\pi \in Decomp(\sigma, \tau, \sigma')$, let $R_1 = \phi_R(\pi(\sigma))$, let $R_2 = \phi_R(\pi(\tau))$, and let $R_3 = \phi_R(\pi(\sigma'^T))$. Note that this is well defined because for any element $\pi' \in Aut(\sigma) \times Aut(\tau) \times Aut(\sigma'^T)$, $\phi_R(\pi\pi'(\sigma)) = \phi_R(\pi(\pi'(\sigma))) = \phi_R(\pi(\sigma))$. Similarly, $\phi_R(\pi\pi'(\tau)) = \phi_R(\pi(\tau))$ and $\phi_R(\pi\pi'(\sigma'^T)) = \phi_R(\pi(\sigma'^T))$.

To confirm that this is bijection, we have to show that these two maps are inverses of each other. Given R_1, R_2, and R_3, applying these two maps gives us ribbons $R'_1 = \phi_R\phi_R^{-1}(\phi_{R_1} \circ \phi_{R_2} \circ \phi_{R_3})(H_\sigma) = R_1$, $R'_2 = \phi_R\phi_R^{-1}(\phi_{R_1} \circ \phi_{R_2} \circ \phi_{R_3})(H_\tau) = R_2$, and $R'_3 = \phi_R\phi_R^{-1}(\phi_{R_1} \circ \phi_{R_2} \circ \phi_{R_3})(H_{\sigma'^T}) = R_3$. Conversely, given R and an element $\pi \in Decomp(\sigma, \tau, \sigma')$ (which we represent by an element $\pi \in Aut(\sigma \circ \tau \circ \sigma'^T)$), applying these two maps gives us the ribbon

$$R' = \phi_R(\pi(\sigma)) \circ \phi_R(\pi(\tau)) \circ \phi_R(\pi(\sigma'^T)) = \phi_R\pi(\sigma \circ \tau \circ \sigma'^T) = R$$

and gives us the map

$$\phi_R^{-1}(\phi_{\phi_R(\pi(\sigma))} \circ \phi_{\phi_R(\pi(\tau))} \circ \phi_{\phi_R(\pi(\sigma'^T))})$$

Now observe that both $\phi_R\pi$ and $\phi_{\phi_R(\pi(\sigma))}$ give bijective maps from σ to the ribbon $\phi_R\pi(\sigma)$ so $\phi^{-1}_{\phi_R(\pi(\sigma))}\phi_R\pi \in Aut(\sigma)$. Following similar logic for τ and σ'^T, in $Decomp(\sigma, \tau, \sigma')$ this map is equivalent to $\phi_R^{-1}(\phi_R\pi) = \pi$ □

B.2 Proof of Lemma 3.14

Definition B.3 (Rigorous definition of intersection patterns)**.** *We define an intersection pattern P on composable shapes γ, τ, γ'^T to consist of the shape $\gamma \circ \tau \circ \gamma'^T$ together with a non-empty set of constraint edges $E(P)$ on $V(\gamma \circ \tau \circ \gamma'^T)$ such that:*

1. *For all vertices $u, v, w \in V(\gamma \circ \tau \circ \gamma'^T)$, if $(u, v), (v, w) \in E(P)$ then $(u, w) \in E(P)$*
2. *$E(P)$ does not contain a path between two vertices of γ, two vertices of τ, or two vertices of γ'^T. This ensures that when we consider γ, τ, γ' individually, their vertices are distinct.*
3. *Defining $V_*(\gamma) \subseteq V(\gamma)$ to be the vertices of γ which are incident to an edge in $E(P)$, U_γ is the unique minimum-weight vertex separator between U_γ and $V_*(\gamma) \cup V_\gamma$*
4. *Similarly, defining $V_*(\gamma'^T) \subseteq V(\gamma'^T)$ to be the vertices of γ'^T which are incident to an edge in $E(P)$, $V_{\gamma'^T}$ is the unique minimum-weight vertex separator between $V_*(\gamma'^T) \cup U_{\gamma'^T}$ and $V_{U_{\gamma'^T}}$*

5.* *All edges in $E(P)$ are between vertices of the same type.*

Definition B.4. *We say that two intersection patterns P, P' on shapes γ, τ, γ'^T are equivalent (which we write as $P \equiv P'$) if there is an automorphism $\pi \in Aut(\gamma) \times Aut(\tau) \times Aut(\gamma'^T)$ such that $\pi(P) = P'$ (i.e. if $E(P)$ and $E(P')$ are the constraint edges for P and P' respectively then $\pi(E(P)) = E(P')$).*

Definition B.5. *Given composable shapes γ, τ, γ'^T, we define $\mathcal{P}_{\gamma,\tau,\gamma'^T}$ to be the set of all possible intersection patterns P on γ, τ, γ'^T (up to equivalence)*

Definition B.6. *Given composable (but not properly composable) ribbons R_1, R_2, R_3 of shapes γ, τ, γ', we define the intersection pattern $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$ induced by R_1, R_2, R_3 as follows:*

1. *Take the canonical maps $\phi_{R_1} : V(\gamma) \to V(R_1)$, $\phi_{R_2} : V(\tau) \to V(R_2)$, and $\phi_{R_3} : V(\gamma'^T) \to V(R_3)$*

2. *Given vertices $u \in V(\gamma)$ and $v \in V(\tau)$, add a constraint edge between u and v if and only if $\phi_{R_1}(u) = \phi_{R_2}(v)$. Similarly, given vertices $u \in V(\gamma)$ and $w \in V({\gamma'}^T)$, add a constraint edge between u and w if and only if $\phi_{R_1}(u) = \phi_{R_3}(w)$ and given vertices $v \in V(\tau)$ and $w \in V({\gamma'}^T)$, add a constraint edge between v and w if and only if $\phi_{R_2}(v) = \phi_{R_3}(w)$.*

Definition B.7. *Given an intersection pattern $P \in \mathcal{P}_{\gamma,\tau,{\gamma'}^T}$, we define $V(\gamma \circ \tau \circ {\gamma'}^T)/E(P)$ to be $V(\gamma \circ \tau \circ {\gamma'}^T)$ where all of the edges in $E(P)$ are contracted (i.e. if $(u, v) \in E(P)$ then $u = v$ and $u = v$ only appears once).*

Definition B.8. *Given an intersection pattern $P \in \mathcal{P}_{\gamma,\tau,{\gamma'}^T}$, we define τ_P to be the shape such that:*

1. $V(H_{\tau_P}) = V(\gamma \circ \tau \circ {\gamma'}^T)/E(P)$
2. $E(H_{\tau_P}) = E(\gamma) \cup E(\tau) \cup E({\gamma'}^T)$
3. $U_{\tau_P} = U_\gamma$
4. $V_{\tau_P} = V_{{\gamma'}^T}$

Definition B.9. *Given an intersection pattern $P \in \mathcal{P}_{\gamma,\tau,{\gamma'}^T}$, we make the following definitions:*

1. *We define $Aut(P) = \{\pi \in Aut(\gamma \circ \tau \circ {\gamma'}^T) : \pi(E(P)) = E(P)\}$*
2. *We define $Aut_{pieces}(P) = \{\pi \in Aut(U_\gamma) \times Aut(\tau) \times Aut({\gamma'}^T) : \pi(E(P)) = E(P)\}$*
3. *We define $N(P) = |Aut(P)/Aut_{pieces}(P)|$*

Lemma B.10. *For all composable σ, τ, and ${\sigma'}^T$ (inclulding improper τ),*

$$M^{fact}_{\tau}(e_\sigma e^T_{\sigma'}) - M^{orth}_{\tau}(e_\sigma e^T_{\sigma'}) = \sum_{\substack{\sigma_2,\gamma:\gamma \text{ is non-trivial},\\ \sigma_2 \cup \gamma = \sigma}} \frac{1}{|Aut(U_\gamma)|} \sum_{P \in \mathcal{P}_{\gamma,\tau,Id_{V_\tau}}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'})$$
$$+ \sum_{\substack{\sigma'_2,\gamma':\gamma' \text{ is non-trivial},\\ \sigma'_2 \cup \gamma' = \sigma'}} \frac{1}{|Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{Id_{U_\tau},\tau,{\gamma'}^T}} N(P) M^{orth}_{\tau_P}(e_\sigma e^T_{\sigma'_2})$$
$$+ \sum_{\substack{\sigma_2,\gamma:\gamma \text{ is non-trivial},\\ \sigma_2 \cup \gamma = \sigma}} \sum_{\substack{\sigma'_2,\gamma':\gamma' \text{ is non-trivial},\\ \sigma'_2 \cup \gamma' = \sigma'}} \frac{1}{|Aut(U_\gamma)| \cdot |Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{\gamma,\tau,{\gamma'}^T}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'_2})$$

Proof. This lemma follows from the following bijection. Consider the third term

$$\sum_{\substack{\sigma_2,\gamma:\gamma \text{ is non-trivial},\\ \sigma_2 \cup \gamma = \sigma}} \sum_{\substack{\sigma'_2,\gamma':\gamma' \text{ is non-trivial},\\ \sigma'_2 \cup \gamma' = \sigma'}} \frac{1}{|Aut(U_\gamma)| \cdot |Aut(U_{\gamma'})|} \sum_{P \in \mathcal{P}_{\gamma,\tau,{\gamma'}^T}} N(P) M^{orth}_{\tau_P}(e_{\sigma_2} e^T_{\sigma'_2})$$

On one side, we have the following data:

1. Ribbons R_1, R_2, and R_3 such that
   (a) R_1, R_2, R_3 have shapes σ, τ, and ${\sigma'}^T$, respectively.

(b) $A_{R_2} = B_{R_1}$ and $A_{R_3} = B_{R_2}$

(c) $(V(R_1) \cup V(R_2)) \cap V(R_3) \neq A_{R_3}$ and $(V(R_2) \cup V(R_3)) \cap V(R_1) \neq B_{R_1}$

2. An ordering $O_{S'}$ on the leftmost minimum vertex separator S' between A_{R_1} and $V_* \cup B_{R_1}$.

3. An ordering $O_{T'}$ on the rightmost minimum vertex separator S' between $V_* \cup A_{R_3}$ and B_{R_3}.

On the other side, we have the following data

1. An intersection pattern $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$ where γ and ${\gamma'}^T$ are non-trivial.

2. Ribbons R'_1, R'_2, R'_3 of shapes σ_2, τ_P, ${\sigma'_2}^T$ such that $V(R'_1) \cap V(R'_2) = A_{R'_2} = B_{R'_1}$, $V(R'_2) \cap V(R'_3) = B_{R'_2} = A_{R'_3}$, and $V(R'_1) \cap V(R'_3) = A_{R'_2} \cap B_{R'_2}$

3. An element $\pi \in Aut(P)/Aut_{pieces}(P)$

To see this bijection, given R_1, R_2, R_3, we again implement our strategy for analyzing intersection terms. Recall that V_* is the set of vertices in $V(R_1) \cup V(R_2) \cup V(R_3)$ which have an unexpected equality with another vertex, S' is the leftmost minimum vertex separator between A_{R_1} and $B_{R_1} \cup V_*$, and T' is the rightmost minimum vertex separator between $A_{R_3} \cup V_*$ and B_{R_3}.

1. Decompose R_1 as $R_1 = {R'}_1 \circ R_4$ where ${R'}_1$ is the part of R_1 between A_{R_1} and $(S', O_{S'})$ and R_4 is the part of R_1 between $(S', O_{S'})$ and $B_{R_1} = A_{R_2}$. Decompose R_3 as $R_5 \cup R'_3$ where R_5 is the part of R_3 between A_{R_3} and $(T', O_{T'})$ and R'_3 is the part of R_3 between $(T', O_{T'})$ and B_{R_3}

2. Take the intersection pattern P and the ribbon R'_2 induced by R_4, R_2, and R_5.

3. Observe that we have two bijective maps from $V(\gamma \circ \tau \circ {\gamma'}^T)/E(P)$ to $V(R_4) \cup V(R_2) \cup V(R_5)$. The first map is $\phi_{R_4} \circ \phi_{R_2} \circ \phi_{R_5}$ and the second map is $\phi_{R'_2}$. We take $\pi = \phi^{-1}_{R'_2}(\phi_{R_4} \circ \phi_{R_2} \circ \phi_{R_5})$.

Conversely, given an intersection pattern $P \in \mathcal{P}_{\gamma,\tau,\gamma'^T}$, R'_1, R'_2, R'_3, and an element $\pi \in Aut(P)/Aut_{pieces}(P)$:

1. Take $R_4 = \phi_{R'_2}\pi(V(\gamma))$, $R_2 = \phi_{R'_2}\pi(V(\tau))$, and $R_5 = \phi_{R'_2}\pi(V({\gamma'}^T))$.

2. Take $R_1 = R'_1 \cup R_4$ and take $R_3 = R_5 \cup R'_3$.

3. Take O_S and O_T based on $B_{R'_1} = A_{R_4}$ and $B_{R_5} = A_{R'_3}$.

To confirm that this is a bijection, we need to show that these maps are inverses of each other.

If we apply the first map and then the second, we obtain the following:

1. We obtain the ribbons

(a) $R''_1 = R'_1 \circ \phi_{R'_2}\phi^{-1}_{R'_2}(\phi_{R_4} \circ \phi_{R_2} \circ \phi_{R_5})(V(\gamma))$

(b) $R''_2 = \phi_{R'_2}\phi^{-1}_{R'_2}(\phi_{R_4} \circ \phi_{R_2} \circ \phi_{R_5})(V(\tau))$

(c) $R''_3 = \phi_{R'_2}\phi^{-1}_{R'_2}(\phi_{R_4} \circ \phi_{R_2} \circ \phi_{R_5})(V({\gamma'}^T)) \circ R'_3$

where

(a) R'_1 is the part of R_1 between A_{R_1} and $(S', O_{S'})$ where S' is the minimum vertex separator between A_{R_1} and $V_* \cup B_{R_1}$.
(b) R_4 is the part of R_1 between $(S', O_{S'})$ and B_{R_1}
(c) R'_2 is the ribbon of shape τ_P induced (along with the intersection pattern P) by R_1, R_2, and R_3.
(d) R_5 is the part of R_3 between A_{R_3} and $(T', O_{T'})$.
(e) R'_3 is the part of R_3 between $(T', O_{T'})$ and B_{R_3}

This implies that $R''_1 = R'_1 \circ R_4 = R_1$, $R''_2 = R_2$, and $R''_3 = R_5 \circ R'_3 = R_3$. Since the second map leaves R'_1 and R'_3 unchanged, we recover the orderings O_S and O_T as well.

Conversely, if we apply the second map, we have that $R_1 = R'_1 \circ \phi_{R'_2}\pi(V(\gamma))$, $R_2 = \phi_{R'_2}\pi(V(\tau))$, and $R_3 = \phi_{R'_2}\pi(V(\gamma'^T)) \circ R'_3$ and we have the orderings O_S and O_T corresponding to $B_{R'_1}$ and $A_{R'_3}$ respectively. If we apply the first map,

1. R'_1 and R'_3 are preserved.
2. R''_2 and P'' are the ribbon and intersection pattern induced by the ribbons $\phi_{R'_2}\pi(\gamma)$, $\phi_{R'_2}\pi(\tau)$, and $\phi_{R'_2}\pi(\gamma'^T)$. To see that $R''_2 = R'_2$, observe that
$$R''_2 = \phi_{R'_2}\pi(V(\gamma)) \circ \phi_{R'_2}\pi(V(\tau)) \circ \phi_{R'_2}\pi(V(\gamma'^T)) = \phi_{R'_2}\pi(\gamma \circ \tau \circ \gamma'^T) = \phi_{R_2}(\gamma \circ \tau \circ \gamma'^T) = R'_2$$
To see that $P'' \equiv P$, observe that:
   (a) We have two bijective maps from $V(\gamma)$ to $V(\phi_{R'_2}\pi(\gamma))$. These two maps are $\phi_{R'_2}\pi$ and $\phi_{\phi_{R'_2}\pi(\gamma)}$.
   (b) We have two bijective maps from $V(\tau)$ to $V(\phi_{R'_2}\pi(\tau))$. These two maps are $\phi_{R'_2}\pi$ and $\phi_{\phi_{R'_2}\pi(\tau)}$.
   (c) We have two bijective maps from $V(\gamma'^T)$ to $V(\phi_{R'_2}\pi(\gamma'^T))$. These two maps are $\phi_{R'_2}\pi$ and $\phi_{\phi_{R'_2}\pi(\gamma'^T)}$.
   (d) For P'', the constraint edges are
$$\left(\phi^{-1}_{\phi_{R'_2}\pi(\gamma)}\phi_{R'_2}\pi \circ \phi^{-1}_{\phi_{R'_2}\pi(\tau)}\phi_{R'_2}\pi \circ \phi^{-1}_{\phi_{R'_2}\pi(\gamma'^T))}\phi_{R'_2}\pi\right)(E(P))$$
3. We have that
$$\pi'' = \phi^{-1}_{R'_2}(\phi_{\phi_{R'_2}\pi(V(\gamma))} \circ \phi_{\phi_{R'_2}\pi(V(\tau))} \circ \phi_{\phi_{R'_2}\pi(V(\gamma'^T))})$$
To see that $\pi'' \equiv \pi$, note that
$$\pi = \pi''\left(\phi^{-1}_{\phi_{R'_2}\pi(V(\gamma))}\phi_{R'_2}\pi \circ \phi^{-1}_{\phi_{R'_2}\pi(V(\tau))}\phi_{R'_2}\pi \circ \phi^{-1}_{\phi_{R'_2}\pi(V(\gamma'^T))}\phi_{R'_2}\pi)\right)$$

The analysis for the the first term is the same except that when γ' is trivial, we always take γ' to be the identity so $T = V(V_\tau) = V(U_{\sigma'^T})$ and the ordering O_T is given by $V_\tau = U_{\sigma'^T}$. Similarly, the analysis for the the second term is the same except that when γ is trivial, we always take γ to be the identity so $S = V(V_\sigma) = V(U_\tau)$ and the ordering O_S is given by $V_\sigma = U_\tau$. □