

---

# SMAC: Simultaneous Mapping and Clustering Using Spectral Decompositions

---

Chandrajit Bajaj<sup>1</sup> Tingran Gao<sup>2</sup> Zihang He<sup>3</sup> Qixing Huang<sup>1</sup> Zhenxiao Liang<sup>3</sup>

## Abstract

We introduce a principled approach for *simultaneous mapping and clustering* (SMAC) for establishing consistent maps across heterogeneous object collections (e.g., 2D images or 3D shapes). Our approach takes as input a heterogeneous object collection and a set of maps computed between some pairs of objects, and outputs a homogeneous object clustering together with a new set of maps possessing optimal intra- and inter-cluster consistency. Our approach is based on the spectral decomposition of a data matrix storing all pairwise maps in its blocks. We additionally provide tight theoretical guarantees for the accuracy of SMAC under established noise models. We also demonstrate the usefulness of our approach on synthetic and real datasets.

## 1. Introduction

Establishing maps (e.g. pointwise correspondences) across object collections is a fundamental problem spanning many scientific domains. High-quality maps facilitating information propagation and transformation are key to applications ranging from 3D reconstruction with partial scans (Huber & Hebert, 2001), data-driven geometry completion and reconstruction (Pauly et al., 2005), texture transfer (Schreiner et al., 2004; Kraevoy & Sheffer, 2004), to comparative biology (Boyer et al., 2011; Gao et al., 2017), joint data-analysis (Huang et al., 2011; Kim et al., 2012; Wang et al., 2013; 2014; Huang et al., 2014), and data exploration and organization (Kim et al., 2012; Huang et al., 2014).

High quality object maps are generally difficult to compute. Prior work on map computation focused on optimizing maps between pairs of objects; see (van Kaick et al., 2010) for

a standard survey and (Kim et al., 2011; Mandad et al., 2017) for some recent advances. Despite the significant progress, state-of-the-art techniques tends to hit a barrier on the quality of maps that are computed in a pairwise manner. Building upon the availability of big-data, a recent line of research (Kim et al., 2012; Nguyen et al., 2011; Ovsjanikov et al., 2012; Huang et al., 2012; Huang & Guibas, 2013; Huang et al., 2014; Chen et al., 2014; Zhou et al., 2015a,b; Shen et al., 2016; Leonardos et al., 2017; Huang et al., 2017) considered computing many pairwise maps jointly among a collection of objects. The promise of these approaches hinges upon the observation that one way to obtain a high quality pairwise map between dissimilar objects is to choose a path connecting these objects but consisting of consecutive similar shapes: maps between similar objects are typically of higher quality, and so is the resulted composed map. From a regularization perspective, joint map computation leverages the generic consistency of a network of maps among multiple objects, in which composition of maps along cycles are expected to be close to the identity map.

However, the performance of these data-driven approaches relies predominantly on the *homogeneity* of the object collection, i.e. the input objects fall into the same category or sub-category (e.g. Chairs, Cars, and Human models). In the presence of *heterogeneous* data, where the input objects fall into multiple underlying categories, applying existing data-driven approaches without the category label information tend to produce unsatisfactory results. In this setting, even though existing methods are able to suppress the noise in intra-cluster maps within a single cluster, jointly computed maps for the entire object collection leads are often significantly worse. One explanation is that high fraction of incorrect inter-cluster maps tends to “contaminate” the regularization effect of intra-cluster maps. A natural resolution is to employ a two-stage cascading strategy that identifies the underlying clusters before computing the intra- and inter-cluster maps. Unfortunately, such clustering requires accurate quantification of the object similarities, which is a difficult problem in its own right. Meanwhile, the error produced in the clustering stage is unlikely remedied by the consistency-based regularization.

In this paper, we propose to solve the mapping and clustering problems *simultaneously*. Instead of explicitly relying on certain pairwise similarity and/or map distortion

---

<sup>1</sup>Department of Computer Science, The University of Texas at Austin <sup>2</sup>Department of Statistics, The University of Chicago <sup>3</sup>Institute for Interdisciplinary Information Sciences, Tsinghua University. Correspondence to: Chandrajit Bajaj <bajaj@cs.utexas.edu>, Qixing Huang <huangqx@cs.utexas.edu>.

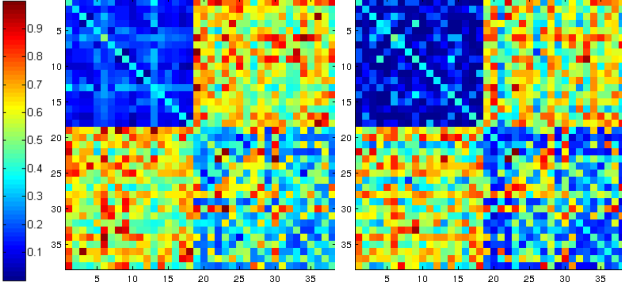


Figure 1. The pairwise map distortion scores (Left) and the pairwise cycle-consistency scores (Right) derived from 3-cycles. The scores are normalized by the maximum score in each matrix.

scores, we identify the underlying clusters based on the consistency of intra- and inter-cluster maps, inspired by the observation that maps tend to be more consistent along cycles within a cluster than across clusters. This discrepancy has been observed in many different contexts, and appears to be a consequence of the energy landscape of almost all optimization-based pairwise matching algorithms. The matching energy functional between objects in the same underlying cluster tends to have simple energy landscapes with easily identifiable global optimums, resulting in fairly cycle-consistent intra-cluster maps; in contrast, the highly non-convex energy landscape for dissimilar objects from different clusters leads to more “random” maps due to random initialization and/or multiple strong local minimums, for which cycle-consistency is much less often observed. This map consistency argument is the foundation of our *simultaneous mapping and clustering* (SMAC) algorithm.

### 1.1. Motivating Example

We validate the map consistency argument through a motivating example (see Figure 1) on a real dataset from SHREC07 Watertight benchmark (Giorgi et al., 2007). This dataset consists of 38 shapes: the first 18 are Human models and the remaining 20 are Fourleg models (e.g., Horses and Dogs). Each shape is represented as a discrete metric space with 1024 sample points generated from farthest-point sampling (Eldar et al., 1997). We compute pairwise *blended intrinsic maps* (Kim et al., 2011) for all objects in this collection and use these maps to compute two similarity scores for each object pair: a *map distortion score* that measures the squared sum of geodesic distortions across all point-pairs (c.f. (Bronstein et al., 2006)), and a *cycle-consistency score* that is the median value (among other options) of the distortion scores of all 3-cycles to which the pair belongs, where the distortion of a 3-cycle is defined as the squared sum of the geodesic distances between each point and its image propagated along the 3-cycle.

Figure 1 illustrates the distributions of the map distortion scores (Left) and the cycle-consistency scores (Right) on the 38 models. The cycle-consistency scores clearly reveal the

underlying cluster structure and in fact better separates the two clusters of models (Human vs. Fourleg) than the map-distortion scores (intra-cluster blocks in the right figure are darker in blue). The superior cluster separation is verified by comparing the results of spectral clustering (Ng et al., 2002; Lei & Rinaldo, 2015) using the two similarity scores: spectral clustering based on the cycle-consistency scores recovers the two underlying clusters perfectly, whereas the same procedure using the map distortion scores incorrectly puts two Fourleg models in the cluster of Human models. This motivating example illustrates the effectiveness and superiority of the map consistency score as a quantification of object similarity.

### 1.2. Approach Overview

Motivated by the example above, we propose an algorithm for *simultaneous mapping and clustering* (SMAC). Our SMAC algorithm takes as input (i) an object collection that falls into multiple clusters, and (ii) some noisy maps pre-computed between object pairs, and outputs the underlying clusters together with improved maps between all pairs of objects. Our SMAC algorithm builds upon the equivalence between map consistency and the low-rank property of a data matrix with consistent maps in its blocks (c.f. (Huang & Guibas, 2013)). We show that this low-rank property still holds in the setting of multiple disjoint collections of consistent intra-cluster maps, though the rank is expected to be higher due to multiple clusters. Based on this observation, the first step of our approach simultaneously recovers the underlying clusters and intra-cluster maps by spectral decomposition. We show that properly “rounding off” the leading eigenvectors recovers the ground-truth clusters and intra-cluster maps in a single pass. We then construct inter-cluster maps from the recovered intra-cluster maps. Our theoretical analysis establishes sharp exact recovery conditions for both steps under a fairly general noise model, using some novel tight  $L^\infty$ -stability bounds for eigen-decompositions.

### 1.3. Related Works

Joint object matching, i.e., simultaneously estimating maps among a collection of objects, is an emerging field across many scientific problems. Earlier works use combinatorial optimizations (Nguyen et al., 2011; Kim et al., 2012; Huang et al., 2012). More recent works (Huang & Guibas, 2013; Huang et al., 2014; Chen et al., 2014; Wang & Singer, 2013) rely on convex or non-convex optimization techniques. However, all these methods assume that the underlying object collection is *homogeneous* (all objects belong to a single category). For a *heterogeneous* object collection where the objects fall into multiple distinctive clusters, existing methods usually rarely succeed in producing both high-quality intra- and inter-cluster maps.

Clustering and in particular spectral clustering is another well-studied topic. We refer to (Lei & Rinaldo, 2015; Rohe et al., 2011) for some recent advances and to (Filippone et al., 2008; Luxburg, 2007; Fortunato, 2010) for surveys. Our approach falls into the general category of graph-based clustering, but the pairwise information we utilize is of “functional” rather than “scalar” nature. Instead of the more common approach that derives affinity scores from the pairwise maps for clustering, our SMAC algorithm discovers the cluster structure based purely on the consistency of pairwise maps and demonstrates improved empirical performance. This strategy is reminiscent of *heterogeneous multireference alignment* (Boumal et al., 2017) and *simultaneous alignment and classification* (Lederman & Singer, 2016) for *synchronization problems* in Cryo-Electron Microscopy; in this context, our residue-based clustering strategy is closest in nature to *learning group actions* (Gao et al., 2016).

Our approach relies on tight  $L^\infty$ -type bounds on leading eigenvectors of perturbed matrices. Though the stability of eigenvalues and eigenspaces are well-studied, element-wise eigenvector stability appears to be a much harder problem; see recent survey (O’Rourke et al., 2016). We introduce new stability bounds to tackle this technical difficulty.

#### 1.4. Mathematical Notation

We use lower bold letters  $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \dots$  to denote vectors, and upper letters  $A, B, C, \dots$  for matrices. For a block matrix  $X \in \mathbb{R}^{n_1 m_1 \times n_2 m_2}$ , we use  $X_{ij} \in \mathbb{R}^{m_1 \times m_2}$  to denote its  $ij$ -th block; the  $ij$ -th element of a matrix  $A$  is denoted as  $a_{ij}$ . With  $\otimes$  we denote the Kronecker product. For a symmetric matrix  $A \in \mathbb{R}^{n \times n}$ , we always sort the eigenvalues in non-increasing order, i.e.,  $\lambda_1(A) \geq \dots \geq \lambda_n(A)$ . Matrix norms  $\|\cdot\|_{\mathcal{F}}, \|\cdot\|_1, \|\cdot\|_2$ , and  $\|\cdot\|_\infty$  will be used for a matrix  $A \in \mathbb{R}^{n_1 \times n_2}$ , of which  $\|\cdot\|_2 = \sigma_{\max}(A)$  is the maximum singular value of  $A$ , and the spectral norm  $\|\cdot\|_2$  is sometimes simplified as  $\|\cdot\|$ . We denote  $\mathcal{P}_m$  for the set of permutation matrices of dimension  $m \times m$ .

## 2. Algorithm

For simplicity, we focus on describing and analyzing our algorithm under the setting where pair-wise maps are given by permutations. In Section 4, we show how to modify the algorithm for partially similar objects.

We first describe the problem setup. Consider  $n$  objects  $\mathcal{S} = \{S_1, \dots, S_n\}$  each represented by  $m$  points. With  $\mathcal{G} = (\mathcal{S}, \mathcal{E})$  we denote an observation graph among  $\mathcal{S}$ . An initial map  $X_{ij}^{in} \in \mathcal{P}_m$  is pre-computed on each edge  $(i, j) \in \mathcal{E}$  using an off-the-shelf pairwise object matching algorithm from  $S_i$  to  $S_j$ . We also assume the objects in  $\mathcal{S}$  are partitioned into  $k \geq 2$  clusters, but  $k$  is unknown. Our goal is to identify the underlying clusters, and in the mean-

**Algorithm 1** PermSMAC: Simultaneously mapping and clustering

**Input:** Observation graph  $\mathcal{G} = (\mathcal{S}, \mathcal{E})$  and initial pairwise maps  $X_{ij}^{in}, (i, j) \in \mathcal{E}$

**Output:** Underlying clusters  $\mathcal{S} = c_1 \cup \dots \cup c_k$  and optimized pairwise maps  $X_{ij}, 1 \leq i, j \leq n$

- 1: **{Step 1}** Simultaneously compute the intra-cluster maps and extract the underlying clusters:
- 2:   **{Step 1.1}** Form data matrix based on (1).
- 3:   **{Step 1.2}** Compute the critical value  $r = \operatorname{argmax}_{2 \leq i \leq nm} \frac{\lambda_i - \lambda_{i+1}}{\lambda_i + \lambda_{i+1}}$ .
- 4:   **{Step 1.3}** Let  $U \in \mathbb{R}^{nm \times r}$  store the leading  $r$  eigenvectors of  $X$ . Compute pair-wise maps  $X_{ij}^*$  by solving (2)
- 5:   **{Step 1.4}** Use  $f_{ij}(X_{ij}^*)$  as the affinity score and apply single-linkage clustering to obtain the underlying clusters
- 6: **{Step 2}** compute the inter-cluster maps by solving (6)

while improve all pairwise maps between objects in  $\mathcal{S}$ . As a basis for identifying the underlying clusters, we assume that the intra-cluster maps are more accurate (in terms of cycle-consistency) than inter-cluster maps.

Our algorithm proceeds in two major steps. The first step simultaneously extracts the underlying clusters and computes intra-cluster maps. The second step then computes inter-cluster maps. Now we introduce these two steps in details. Algorithm 1 provides the pseudo code.

### 2.1. Step I: Extract underlying clusters and compute intra-cluster maps.

Motivated from prior works for map synchronization (Pachauri et al., 2013; Shen et al., 2016), we use a block data matrix  $X \in \mathbb{R}^{nm \times nm}$  to encode the input maps:

$$X_{ij} = \begin{cases} (X_{ij}^{in} - \frac{1}{m} \mathbf{1} \mathbf{1}^T) & (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Our approach is motivated by the empirical observation that the leading eigen-vectors of  $X$  reveal the underlying clusters and simultaneously denoise intra-cluster maps if intra-cluster maps are more accurate than inter-cluster maps. We provide the algorithmic details below; an analysis is provided in Section 3.

Given the data matrix, we first estimates the number of stable eigen-vectors as

$$r = \operatorname{argmax}_{m \leq i \leq nm} \frac{\lambda_i - \lambda_{i+1}}{|\lambda_i| + |\lambda_{i+1}|}.$$

Here we search within the range of  $[m, nm]$ , as we expect multiple underlying clusters. Let  $U \in \mathbb{R}^{nm \times r}$  store the

top  $r$  eigen-vectors of  $X$ , and divide  $U$  into  $n$  matrices  $U_1, \dots, U_n$  of shape  $m \times r$  such that  $U = (U_1^T, \dots, U_n^T)^T$ . We then compute the estimated map  $X_{ij}^*$  along each edge  $(i, j) \in \mathcal{E}$  by solving a linear assignment problem, i.e.,

$$X_{ij}^* = \min_{X \in \mathcal{P}_m} f_{ij}(X), \quad f_{ij}(X) := \|X \cdot U_i - U_j\|_{\mathcal{F}}^2 \quad (2)$$

for all  $1 \leq i, j \leq n$ . Note that (2) admits an exact solution via linear assignment. In fact,

$$X_{ij}^* = \operatorname{argmin}_{X \in \mathcal{P}_m} \|X U_i - U_j\|_{\mathcal{F}}^2 \quad (3)$$

$$= \operatorname{argmin}_{X \in \mathcal{P}_m} (\|U_i\|_{\mathcal{F}}^2 + \|U_j\|_{\mathcal{F}}^2 - 2\langle X U_i, U_j \rangle) \quad (4)$$

$$= \operatorname{argmax}_{X \in \mathcal{P}_m} \langle X, U_j U_i^T \rangle. \quad (5)$$

Intuitively,  $U_j U_i^T$  provides an approximation of the underlying map  $X_{ij}$ , and the linear assignment procedure projects this approximation onto the space of permutations.

For clustering, we treat the residual score  $f_{ij}(X_{ij}^*)$ ,  $1 \leq i, j \leq n$  as the distance measure between  $S_i$  and  $S_j$ , and apply single-linkage clustering (Gower & Ross, 1969) to obtain the underlying clusters. We set  $k = \lfloor \frac{r}{m-1} \rfloor$  as the number of desired clusters.

Empirically, we found that when the input inter-cluster maps are inaccurate, the quality of estimated inter-cluster maps appear to be much more noisy than estimated intra-cluster maps. This motivates us to re-estimate inter-cluster maps as a second step, described in Section 2.2 below.

## 2.2. Step II: Estimate inter-cluster maps.

We estimate the inter-cluster maps between each pair of clusters independently. Specifically, consider two clusters  $c_s$  and  $c_t$ . Let  $S_{i_s} \in c_s$  and  $S_{i_t} \in c_t$  be a pair of objects selected from each cluster, respectively. We optimize the inter-cluster map  $X_{st}^{\text{inter}}$ , represented as a pairwise object map  $X_{i_s i_t}$ , by solving the following linear assignment:

$$\begin{aligned} X_{st}^{\text{inter}} &= \operatorname{argmin}_{X \in \mathcal{P}_m} \sum_{i \in c_s, j \in c_t, (i, j) \in \mathcal{E}} \|X - X_{j i_t} X_{i j}^{\text{in}} X_{i i_s}\|_1 \\ &= \operatorname{argmax}_{X \in \mathcal{P}_m} \sum_{i \in c_s, j \in c_t, (i, j) \in \mathcal{E}} \langle X, X_{j i_t} X_{i j}^{\text{in}} X_{i i_s} \rangle. \end{aligned} \quad (6)$$

Note that it is possible to jointly optimize  $X_{st}^{\text{inter}}$  among all pairs of clusters. However, since the number of clusters is usually significantly smaller than the size of each cluster, we found the gain of doing so is insignificant.

## 3. Analysis

We first describe our noise model in Section 3.1. We then analyze our method under this model and present the exact

recovery conditions for both underlying clusters and the pairwise maps in Section 3.2. Our analysis is based on a set of new stability bounds of eigen-decompositions. The proofs of all Lemma's and Theorem's with technical details can be referred to in the supplemental material.

### 3.1. Model for Analysis

We consider two models, one for the pairwise map and the other one for the observation graph and the underlying cluster structure.

**Map model.** We generalize the map model described in (Shen et al., 2016) to multiple clusters: Suppose there are  $k$  underlying clusters. With  $c_s$ ,  $1 \leq s \leq k$  we denote the vertex indices of the  $s$ -th cluster. In other words, we have  $\{1, \dots, n\} = c_1 \cup c_2 \cup \dots \cup c_k$ . Given an observation graph, the input pairwise maps are independent, and they follow

$$X_{ij}^{\text{in}} = \begin{cases} I_m & \text{with probability } \eta_{ij} \\ U_{\mathcal{P}_m} & \text{with probability } 1 - \eta_{ij} \end{cases}$$

where  $U_{\mathcal{P}_m}$  denotes a random permutation matrix satisfying

$$E[U_{\mathcal{P}_m}] = \frac{1}{m} \mathbf{1} \mathbf{1}^T. \quad (7)$$

$\eta_{ij}$  depends on the edge type:  $\eta_{ij} = p$  if  $(i, j)$  is an intra-cluster edge, i.e.,  $(i, j) \in \mathcal{E} \cap (\cup_{1 \leq s \leq k} c_s \times c_s)$ , and  $\eta_{ij} = q$  if  $(i, j)$  is an inter-cluster edge, i.e.,  $(i, j) \in \mathcal{E} \cap (\cup_{1 \leq s \neq t \leq k} c_s \times c_t)$ . We assume  $p > q$ .

**Remark 1.** Note that one can also generalize the model by assuming there exist underlying permutations  $P_i$ ,  $1 \leq i \leq n$ , so that the ground-truth map  $X_{ij} = P_j P_i^T$ . Nevertheless, it turns out that the two models are identical (Shen et al., 2016). Hence we adopt the simpler form for convenience.

**Model for the observation graph and clusters.** To obtain concise and interpretable exact recovery conditions, we are particularly interested in analyzing our algorithm when the observation graph and underlying clusters are generated from the following established model: assume  $n = n_0 k$ , and the size of each underlying cluster  $|c_s| = n_0$ ,  $1 \leq s \leq k$ ; the observation graph is generated from the Erdős-Rényi  $G(n, t)$  with edge selection probability  $t$ . However, the stability bounds we introduce in the supplemental material can be used for analyzing more general noisy models, e.g., sizes of the underlying clusters are different.

### 3.2. Map and Cluster Recovery

We begin by analyzing the leading eigenvalues and eigenvectors of  $E[X]$  to gain insights on the necessary conditions for map and cluster recovery. To make our discussion more general, we first assume the underlying cluster and the observation graph are fixed. Consider then the following  $(p, q)$ -



reweighted normalized adjacency matrix  $\bar{A}(p, q) \in \mathbb{R}^{n \times n}$ :

$$(\bar{A}(p, q))_{ij} = \begin{cases} p & (i, j) \in \mathcal{E} \cap \cup_{1 \leq s \leq m} (c_s \times c_s), \\ q & (i, j) \in \mathcal{E} \cap \cup_{1 \leq s \neq t \leq m} (c_s \times c_t). \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

It is clear that

$$E[X] = \bar{A}(p, q) \otimes (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T)$$

and thus the non-zero eigenvalues of  $E[X]$  are non-zero eigenvalues of  $\bar{A}(p, q)$  with multiplicity  $m-1$ . Furthermore, let  $(\frac{1}{\sqrt{m}} \mathbf{1}, H_m)$  be an orthonormal basis for  $\mathbb{R}^m$ . Then the leading  $k(m-1)$  eigenvectors of  $E[X]$  are given by  $S_k \otimes H_m$ . This leads two conditions on the eigenvalues and eigenvectors of  $\bar{A}(p, q)$  for map and cluster recovery:

- **Eigenvalue separation.** Since our method leverages the largest eigengap, we assume that  $\lambda_k(\bar{A}(p, q)) - \lambda_{k+1}(\bar{A}(p, q))$  has the largest eigengap. Define

$$\gamma = \frac{\max_{1 \leq i \neq k \leq n-1} \lambda_i(\bar{A}(p, q)) - \lambda_{i+1}(\bar{A}(p, q))}{\lambda_k(\bar{A}(p, q)) - \lambda_{k+1}(\bar{A}(p, q))},$$

Then a necessary condition for map recovery is  $\gamma < 1$ .

- **Eigenvector separation.** We further assume that the underlying clusters can be recovered by reading off the rows of  $S_k$ . Formally, consider rows of  $S_k$  as coordinates of the corresponding objects, and define

$$d_{\text{intra}} = \max_{1 \leq s \leq k} \max_{i, j \in c_s} \|(e_i - e_j)^T S_k\|, \quad (9)$$

$$d_{\text{inter}} = \min_{1 \leq s < t \leq k} \min_{i \in c_s, j \in c_t} \|(e_i - e_j)^T S_k\|. \quad (10)$$

$d_{\text{intra}}$  and  $d_{\text{inter}}$  essentially measure the maximum distance within each cluster and the minimum distance between different clusters, respectively. Thus, a necessary condition for cluster recovery is that  $d_{\text{intra}} < \mu \cdot d_{\text{inter}}$  for some small constant  $\mu$ .

Under these two conditions, it is easy to see that when  $X \approx E[X]$ , we have  $U_i \approx (e_i^T S_k) \otimes H_m$  and  $U_j U_i^T \approx (e_j^T S_k S_k^T e_i) (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T)$ . It follows that both the underlying clusters and intra-cluster maps can be exactly recovered.

These two separation conditions are quite general. In fact, they hold for the noisy model described in Section 3.1. To gain some further insight, one can show that (c.f. (Le et al., 2017)) with high probability

$$\lambda_1(\bar{A}(p, q)) = \begin{cases} q + \frac{p-q}{k} + O(\frac{1}{\sqrt{nt}}) & i = 1 \\ \frac{p-q}{k} + O(\frac{1}{\sqrt{nt}}) & 2 \leq i \leq k \\ O(\frac{1}{\sqrt{nt}}) & k+1 \leq i \leq n \end{cases}$$

indicating that  $\gamma = o(1)$ . Moreover, under this model

$$\bar{A}(p, q) \approx (p - q) I_k \otimes (\mathbf{1}\mathbf{1}^T) + q(\mathbf{1}\mathbf{1}^T).$$

If  $p$  and  $q$  are well-separated, the top  $k$  eigenvectors of  $\bar{A}(p, q)$  approximate  $I_k \otimes \mathbf{1}$ , meaning  $d_{\text{intra}} \approx 0$  and  $d_{\text{inter}} \approx 1$ . Now we formally state the exact recovery conditions:

**Theorem 3.1. (Intra-Cluster and Map Recovery)** Assume  $t = \Omega(\frac{\log(n)}{n})$ . Consider the noise model described in Section 3.1. There exists an absolute constant  $c_{\text{intra}}$  such that PermSMAC recovers the underlying intra-cluster maps and the underlying clusters with high probability if

$$p - q \geq c_{\text{intra}} k \sqrt{\frac{\log(n)}{nt}}. \quad (11)$$

**Remark 2.** Note that the gap between  $p$  and  $q$  is used to ensure the recovery of the underlying clusters. Moreover, the recovery rate matches the information theoretic lower bound established in (Chen et al., 2016) up to  $O(\sqrt{\log(n)})$ , indicating the tightness of our condition for PermSMAC.

The following theorem provides an exact recovery condition for the inter-cluster maps. Compared with the previous lower bound, the lower bound on  $q$  for inter-cluster map recovery is significantly cruder. This shows the advantage of recovering inter-cluster maps as a separate step.

**Theorem 3.2.** There exists an absolute constant  $c_{\text{inter}} > 0$ , so that when

$$q \geq c_{\text{inter}} k \sqrt{\frac{\log(n)}{n^2 t}},$$

PermSMAC recovers the underlying inter-cluster maps with high probability.

## 4. Partial Matching

In this section we extend the algorithm to handle partially (as opposed to fully) similar objects. Each input object  $S_i$  ( $1 \leq i \leq n$ ) in this setting has  $m_i \in \mathbb{N}_+$  points, where the  $m_i$ 's vary across the collection. Consequently, the pairwise maps  $X_{ij}^{in} \in \{0, 1\}^{m_j \times m_i}$  are no longer permutation matrices since some rows and columns may only contain zero elements. We propose the following modified algorithm for SMAC in this partial matching setting.

**Step I: Extract underlying clusters and compute intra-cluster maps.** Forming the data matrix and leading eigenvector computation stay the same as in the full matching setting, except we replace  $X_{ij}^{in} - \frac{1}{m} \mathbf{1}\mathbf{1}^T$  by  $X_{ij}^{in}$  in (1). The first difference occurring in the partial matching setting is that we cannot apply (2) to obtain pair-wise maps and affinity scores for clustering. Our strategy is to apply single linkage clustering on the rows of  $U$ . Specifically, the distance measure between two points  $p, p'$  is given by

$\|\mathbf{u}_p - \mathbf{u}_{p'}\|^2$ , where  $\mathbf{u}_p$  is the corresponding row of  $p$  in  $U$ . The number of output clusters is set as  $r$ . Each output cluster of this single-linkage procedure collects a set of matched points among the input objects. We merge two clusters if the objects they belong to overlap. Suppose we finally obtain  $k$  clusters  $c_1, \dots, c_k$ . For each cluster  $c_i$ , we introduce a binary matrix  $Y_i \in \{0, 1\}^{n_i \times r_i}$ , whose columns encode the enclosed point clusters. Then it is easy to see that the blocks of  $Y_i Y_i^T$  describe intra-cluster maps. Note that  $r_i = m$  in the full setting, but in the partial setting  $r_i$  is usually larger than  $\max_{1 \leq i \leq n} m_i$ , due to partial similarity.

**Step II: Compute inter-cluster maps.** In the partial setting, we encode the inter-cluster map from cluster  $c_s$  to cluster  $c_t$  as a matrix  $X_{st} \in \{0, 1\}^{r_s \times r_t}$ . Consider a object pair  $(i, j) \in \mathcal{E}$ , where  $i \in c_s$  and  $j \in c_t$ . With  $E_{i,s}$  and  $E_{j,t}$  we denote the index matrices that extract the corresponding blocks in  $Y_s$  and  $Y_t$ . It is easy to see that the entries  $Y_t^T E_{j,t} X_{ij}^{in} E_{i,s}^T Y_s$  provide cues for the inter-cluster map  $X_{st}$ . Similar to the full map setting, we compute

$$C_{st} = \sum_{i \in c_s, j \in c_t, (i,j) \in \mathcal{E}} Y_t^T E_{j,t} X_{ij}^{in} E_{i,s}^T Y_s.$$

Since the inter-cluster maps may not be permutation matrices either, we apply a simple thresholding to obtain the inter-cluster maps:

$$X_{st} = C_{st} > \beta_{st},$$

where  $\beta_{st}$  is set as 0.9 times the maximum element of  $C_{st}$  in our experiments.

## 5. Experimental Results

In this section, we evaluate our approach on both synthetic (Section 5.1) and real datasets (Section 5.2 and Section 5.3). For baseline comparison, we consider state-of-the-art approaches for clustering and joint mapping in each domain.

### 5.1. Experimental Results on Synthetic Instances

We apply the model described in Section 3 to generate the synthetic data sets for our experiments. Below we summarize how the procedure depends on the model parameters:

- **The observation graph  $\mathcal{G}$ .** We employ a standard two-community stochastic block model (Abbe et al., 2016) which enables us to fully control the vertex degree and the spectral gap. We use this model to generate three observation graphs  $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$ . All of them have  $n = 300$  vertices, but the vertex degrees and spectral gaps vary. Specifically,  $\mathcal{G}_1$  is the clique graph.  $\mathcal{G}_2$  is a sparse graph, whose average vertex degree is 50 and the spectral gap is 0.1. The average vertex degree and spectral gap of  $\mathcal{G}_3$  are 50 and 0.5, respectively. In

our experiment, we treat  $\mathcal{G}_1$  as the default observation graph. We also study the influence of the observation graphs on the performance of our algorithm.

- **Number of clusters  $k$ .** Without loss of generality, we allocate each object into a underlying cluster with probability  $\frac{1}{k}$ . For each observation graph, we generate and fix one underlying cluster throughout our experiments.
- **Other parameters  $m, \gamma, p, q$ .** We fix the number of points on each object as  $m = 30$ . In the partial matching setting, we follow the protocol (Chen et al., 2014) to generate the input objects so that the expected size of each object is  $m\gamma$ . We sample the ratio of correct inter-cluster maps  $q = \exp(-\frac{i}{10})$ ,  $15 \leq i \leq 50$ . Since  $p > q$ , we sample the ratio of correct intra-cluster maps so that  $p - q = \frac{i}{100}$ ,  $1 \leq i \leq 30$ .

We now study the empirical phase transition curves when varying  $p$  and  $q$  under different  $k, \gamma$  and observation graphs.

**Varying  $k$ .** The first two rows of Figure 2 show the phase transition curves of map and cluster recovery for  $k = 2, 4, 6$ . Across all configurations, our approach tolerates a significant portion of noise in the input maps. The fraction of noise that our approach can handle reduces as  $k$  increases, which is consistent with the exact recovery conditions in Section 3. In addition, phase transitions with respect to mapping recovery and cluster recovery roughly align. The subtle differences are two-fold: when  $p$  and  $q$  are close, cluster recovery breaks as there is no cue for clustering; likewise, map recovery breaks when  $q$  approaches 0.

**Versus mapping only.** Figure 2 compares our approach to state-of-the-art map recovery technique (Huang & Guibas, 2013). SMAC clearly exhibits a clear advantage in the regime, where  $q$  is small and  $p$  is significantly larger than  $q$ . This is expected through our exact recovery condition.

**Varying observation graph.** Figure 3 shows phase transition curves of map and cluster recovery when varying the observation graphs ( $k = 2, \gamma = 1$ ). Our approach tolerates a larger fraction of incorrect maps for larger vertex degrees. Moreover, when vertex degrees are comparable, a small spectral gap means higher recovery rate.

**Varying  $\gamma$ .** Figure 4 shows the phase transition curves when varying the overlapping ratio  $\gamma$ . We again show three configurations, i.e.,  $\gamma = 1, \gamma = 0.8$  and  $\gamma = 0.6$ . Still, our approach can tolerate a large rate of noise in the input maps. Moreover, the rate reduces as  $\gamma$  becomes smaller. This is expected, as low overlapping ratio introduces weak signal for mapping and cluster recovery.

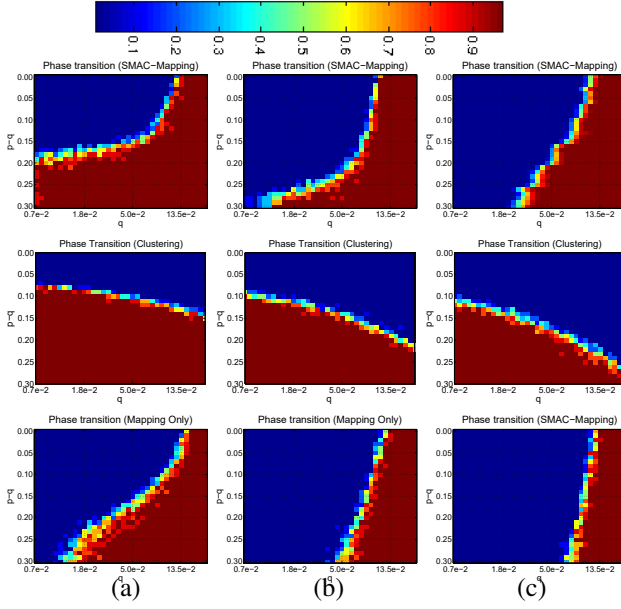


Figure 2. Phase transition curves when varying  $q$  and  $p - q$  under different configurations of  $k$ . The first and second rows show mapping and cluster recovery of PermsMAC, respectively. The third row shows mapping recovery of (Huang & Guibas, 2013). (a)  $k = 2$ . (b)  $k = 4$ . (c)  $k = 6$ . Red means more success.

## 5.2. Experimental Results on 3D Shapes

We proceed to evaluate SMAC on 3D shapes. We consider two datasets for this task. The first dataset collects four categories of 3D shapes from SHREC07-Watertight (Giorgi et al., 2007), namely, Human, Fourleg, Armadillo and Teddy. These categories appear to have both similar global structures and local geometric details. However, the inter-category variability is salient. The second dataset is more fine-grained. It has 10 underlying shape collections from FAUST training dataset (Bogo et al., 2014), where each collection consists of different poses of the same human subject (10 poses per collection). For evaluating shape maps, we follow the protocol of (Kim et al., 2011) by collecting statistics on the geodesic distortion of predicted correspondences with respect to human annotated feature correspondences.

**Mapping performance.** Figure 5(c) plots the accuracy of predicted correspondences of our approach versus the input. For a detailed assessment, we separate the statistics of intra-cluster maps and inter-cluster maps. We consider two approaches for baseline comparison: the first applies (Huang & Guibas, 2013) to the entire dataset, and the second applies (Huang & Guibas, 2013) to each category in isolation then applies the third step of our approach to compute inter-cluster maps. The second baseline may be considered as a performance upper bound. Our proposed SMAC is significantly better than mapping without clustering (which is seriously affected by the noise in inter-cluster maps) and competitive against the second baseline.

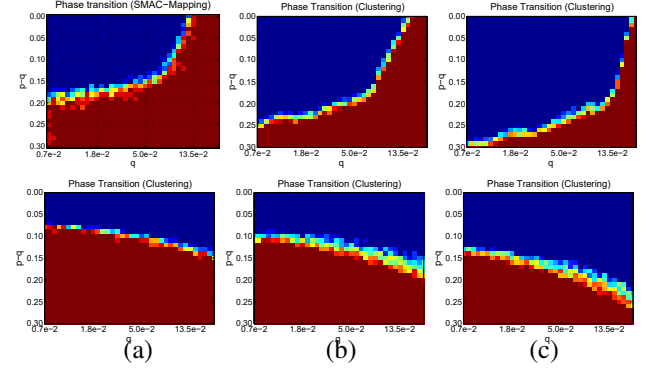


Figure 3. Phase transition curves of map recovery and cluster recovery when varying the observation graph. The top row shows mapping recovery, and the bottom row shows cluster recovery. (a)  $\mathcal{G}_1$ . (b)  $\mathcal{G}_2$ . (c)  $\mathcal{G}_3$ .

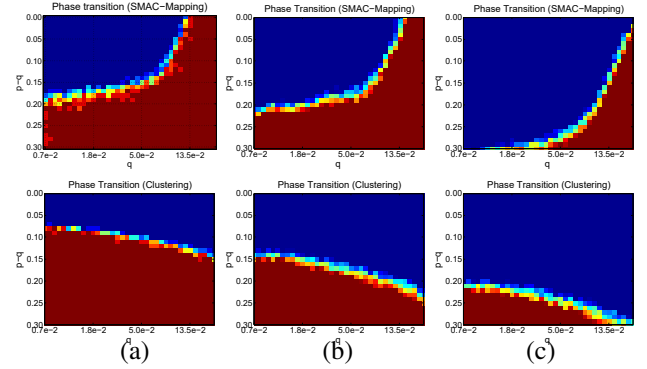


Figure 4. Phase transition curves of map recovery and cluster recovery when varying the overlapping ratio. The top row shows mapping recovery, and the bottom row shows cluster recovery. (a)  $\gamma = 1$ . (b)  $\gamma = 0.8$ . (c)  $\gamma = 0.6$ .

**Clustering performance.** As shown in Table 1, our approach correctly identifies all underlying clusters in SHREC07 and FAUST. We also applied two baseline clustering approaches on the same dataset. The first approach performs  $k$ -means on the spectral shape descriptors (Rustamov, 2007). This approach only yields 84.6% and 72.0%, respectively. The second approach utilizes the mapping distortion as an affinity measure and applies spectral clustering. This approach yields 94.9% and 74.0%, respectively, which are better than the first baseline. However, our approach is still better, which shows the advantage of using the cycle-consistency constraint for clustering.

## 5.3. Experimental Results on 2D Images

Finally, we evaluate our approach on two datasets of 2D images. The first dataset (Figure 6(Left)) consists of 600 internet images of Notre Dame. These images naturally fall into 3 categories, each of which collects images from a similar camera pose (Snaveely et al., 2006). The second dataset (Figure 6(Right)) collects 600 internet images of 4 landmark churches in Europe (Amiens Cathedral (200 im-

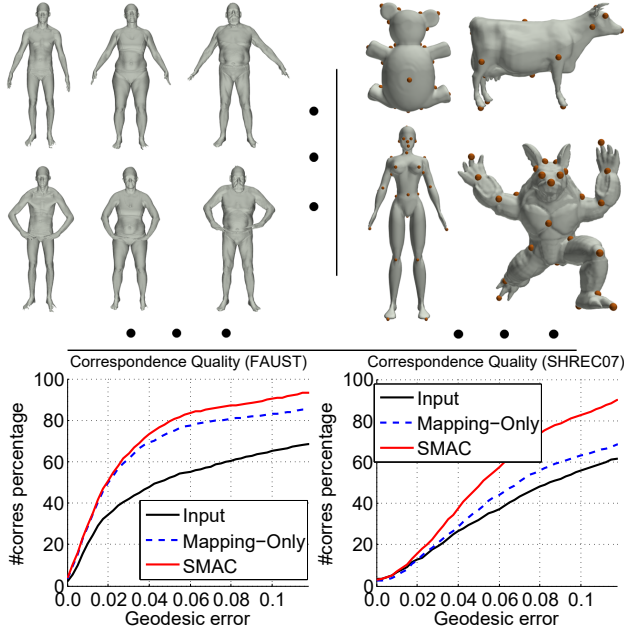


Figure 5. Experimental results on 3D shapes. We consider two datasets: (left) Human/Fourleg/Armadillo/Teddy, and (right) Different poses of xxx human subjects. For each dataset, we show (top) snapshots of each dataset, and (bottom) error distribution of predicted intra-cluster and inter-cluster shape maps.

	SHREC07	FAUST	Notre Dame	Church
SMAC	100%	94.0%	99.3%	96.1%
Distort.	94.9%	74.0%	94.3%	91.9%
Descrip.	84.6%	72.0%	84.3%	86.7%

Table 1. Classification accuracy of our method and baseline approaches on the four datasets show in Figure 5 and Figure 6. Distort. and Descrip. stand for using the pairwise mapping distortion score and global object descriptor for clustering, respectively.

ages), York Minster (200 images), Duomo (100 images) and Westminster Abbey (100 images)). As inter-cluster maps do not make much sense here, we only evaluate clustering results and intra-cluster maps in this experiment. We sample 400 SIFT features (Lowe, 2004) for each image and apply SIFT flow (Liu et al., 2011) to establish pairwise correspondences between the features. We manually mark feature correspondences for evaluation.

**Mapping performance.** Figure 6 compares our approach with the two baseline approaches introduced in Section 5.2. The relative performance is consistent. Specifically, due to the small-overlapping region across different clusters, inter-cluster maps are rather noisy, so applying joint-mapping directly leads to sub-optimal results. In addition, our approach is competitive against the approach of computing intra-cluster and inter-cluster maps in a sequential manner.

**Clustering performance.** Finally, we evaluate our approach in terms of clustering accuracy. We choose two base-

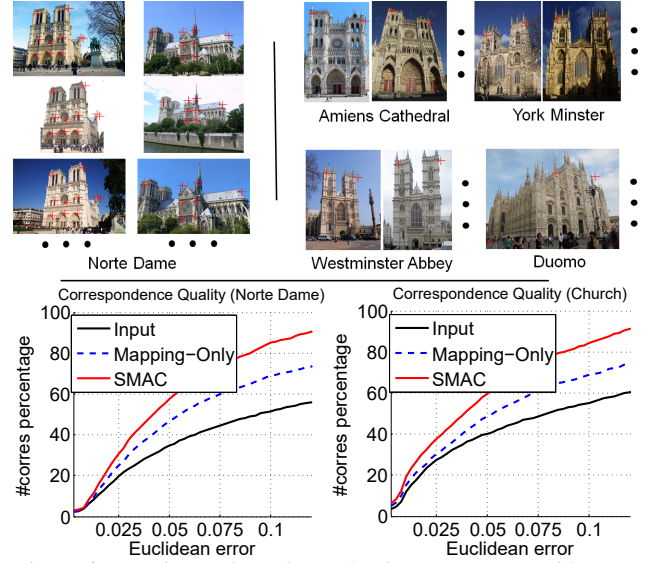


Figure 6. Experimental results on 2D images. We consider two datasets: (Left) Internet images of Norte Dame from three diverse camera poses, and (Right) Internet images of five churches from similar view points. Visualization is the same as Figure 5.

line approaches, where the first baseline approach performs  $k$ -means on image descriptors. In this case, we employ GIST (Oliva & Torralba, 2001). The second baseline uses the residual of SIFT flow as the affinity score for clustering. As shown in Table 1, our approach leads to a clustering accuracy of 99.3% and 96.1% on Notre Dame and Church, respectively. They are higher than those of the top performing baselines, i.e., 94.5% and 92.1%, respectively.

## 6. Conclusions

We have introduced SMAC for simultaneously computing consistent maps across a heterogeneous data collection and identifying the underlying clusters. The key idea is to leverage the higher self-consistency within intra-cluster maps than inter-cluster maps. Enforcing this variation of consistency allows us to denoise the input maps in a sequential manner while simultaneously identifying the underlying cluster structures. The approach is based on spectral decomposition, for which we provided tight exact recovery conditions for both the input maps and the underlying clusters. Experimental results on synthetic data sets justify our exact recovery conditions, and experimental results on real data sets demonstrate the efficacy of our approach.

**Acknowledgement.** This research by Chandrajit Bajaj was supported in part by a grant from the NIH, R01 GM117594. Tingran Gao acknowledges support from Simons Math+X Investigators Award 400837, DARPA D15AP00109, and NSF IIS 1546413. Qixing Huang would like to acknowledge support for this research from NSF DMS-1700234, NSF CI-P-1729486, and NSF IIS-1618648.



## References

- Abbe, E., Bandeira, A. S., and Hall, G. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, 62(1):471–487, 2016.
- Bogo, F., Romero, J., Loper, M., and Black, M. J. FAUST: Dataset and evaluation for 3D mesh registration. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Piscataway, NJ, USA, June 2014. IEEE.
- Boumal, N., Bendory, T., Lederman, R. R., and Singer, A. Heterogeneous multireference alignment: a single pass approach. *arXiv preprint arXiv:1710.02590*, 2017.
- Boyer, D. M., Lipman, Y., St. Clair, E., Puente, J., Patel, B. A., Funkhouser, T., Jernvall, J., and Daubechies, I. Algorithms to Automatically Quantify the Geometric Similarity of Anatomical Surfaces. *Proceedings of the National Academy of Sciences*, 108(45):18221–18226, 2011. doi: 10.1073/pnas.1112822108.
- Bronstein, A. M., Bronstein, M. M., and Kimmel, R. Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching. In *PNAS*, volume 103, pp. 1168–1172, 2006.
- Chen, Y., Guibas, L. J., and Huang, Q. Near-optimal joint object matching via convex relaxation. In *ICML*, pp. 100–108, 2014.
- Chen, Y., Suh, C., and Goldsmith, A. J. Information recovery from pairwise measurements. *IEEE Trans. Information Theory*, 62(10):5881–5905, 2016.
- Eldar, Y., Lindenbaum, M., Porat, M., and Zeevi, Y. Y. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Processing*, 6(9):1305–1315, 1997.
- Filippone, M., Camastra, F., Masulli, F., and Rovetta, S. A survey of kernel and spectral methods for clustering. *Pattern Recog.*, 41(1):176–190, January 2008.
- Fortunato, S. Community detection in graphs. *Physics Reports*, 486(3-5):75 – 174, 2010.
- Gao, T., Brodzki, J., and Mukherjee, S. The Geometry of Synchronization Problems and Learning Group Actions. *arXiv preprint arXiv:1610.09051*, 2016.
- Gao, T., Yapuncich, G. S., Daubechies, I., Mukherjee, S., and Boyer, D. M. Development and Assessment of Fully Automated and Globally Transitive Geometric Morphometric Methods, With Application to a Biological Comparative Dataset With High Interspecific Variation. *The Anatomical Record*, pp. n/a–n/a, 2017. ISSN 1932-8494. Submitted.
- Giorgi, D., Biasotti, S., and Paraboschi, L. Shape retrieval contest 2007: Watertight models track, 2007.
- Gower, J. C. and Ross, G. J. S. Minimum spanning trees and single linkage cluster analysis (ref: 69v18 p106-110). *Applied Statistics*, 18:54–64, 1969.
- Huang, Q. and Guibas, L. J. Consistent shape maps via semidefinite programming. *Comput. Graph. Forum*, 32(5):177–186, 2013.
- Huang, Q., Koltun, V., and Guibas, L. Joint shape segmentation with linear programming. *ACM Trans. Graph.*, 30(6):125:1–125:12, December 2011.
- Huang, Q., Zhang, G., Gao, L., Hu, S., Butscher, A., and Guibas, L. J. An optimization approach for extracting and encoding consistent maps in a shape collection. *ACM Trans. Graph.*, 31(6):167:1–167:11, 2012.
- Huang, Q., Wang, F., and Guibas, L. J. Functional map networks for analyzing and exploring large shape collections. *ACM Trans. Graph.*, 33(4):36:1–36:11, 2014.
- Huang, X., Liang, Z., Bajaj, C., and Huang, Q. Translation synchronization via truncated least squares. In *NIPS*, 2017.
- Huber, D. F. and Hebert, M. Fully automatic registration of multiple 3d data sets. *Image and Vision Computing*, 21: 637–650, 2001.
- Kim, V., Li, W., Mitra, N., DiVerdi, S., and Funkhouser, T. Exploring collections of 3d models using fuzzy correspondences. *ACM Trans. Graph.*, 31(4):54:1–54:11, July 2012.
- Kim, V. G., Lipman, Y., and Funkhouser, T. Blended intrinsic maps. *ACM Trans. Graph.*, 30(4):79:1–79:12, July 2011.
- Kraevoy, V. and Sheffer, A. Cross-parameterization and compatible remeshing of 3d models. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH ’04, pp. 861–869, 2004.
- Le, C. M., Levina, E., and Vershynin, R. Concentration and regularization of random graphs. *Random Struct. Algorithms*, 51(3):538–561, 2017.
- Lederman, R. R. and Singer, A. A representation theory perspective on simultaneous alignment and classification. *arXiv preprint arXiv:1607.03464*, 2016.
- Lei, J. and Rinaldo, A. Consistency of spectral clustering in stochastic block models. *Annals of Statistics*, 43(1): 215–237, February 2015.
- Leonardos, S., Zhou, X., and Daniilidis, K. Distributed consistent data association via permutation synchronization. In *ICRA*, pp. 2645–2652. IEEE, 2017.

- Liu, C., Yuen, J., and Torralba, A. Sift flow: Dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):978–994, 2011.
- Lowe, D. G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, nov 2004.
- Luxburg, U. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, December 2007.
- Mandad, M., Cohen-Steiner, D., Kobbelt, L., Alliez, P., and Desbrun, M. Variance-minimizing transport plans for inter-surface mapping. *ACM Transactions on Graphics*, 36(4), 2017.
- Ng, A. Y., Jordan, M. I., and Weiss, Y. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pp. 849–856, 2002.
- Nguyen, A., Ben-Chen, M., Welnicka, K., Ye, Y., and Guibas, L. J. An optimization approach to improving collections of shape maps. *Comput. Graph. Forum*, 30(5):1481–1491, 2011.
- Oliva, A. and Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vision*, 42(3):145–175, May 2001.
- O’Rourke, S., Vu, V., and Wang, K. Eigenvectors of random matrices: A survey. *J. Comb. Theory, Ser. A*, 144:361–442, 2016.
- Ovsjanikov, M., Ben-Chen, M., Solomon, J., Butscher, A., and Guibas, L. J. Functional maps: a flexible representation of maps between shapes. *ACM Trans. Graph.*, 31(4): 30:1–30:11, 2012.
- Pachauri, D., Kondor, R., and Singh, V. Solving the multi-way matching problem by permutation synchronization. In *NIPS*, pp. 1860–1868, 2013.
- Pauly, M., Mitra, N. J., Giesen, J., Gross, M., and Guibas, L. J. Example-based 3d scan completion. In *Proceedings of the Third Eurographics Symposium on Geometry Processing*, SGP ’05, 2005.
- Rohe, K., Chatterjee, S., and Yu, B. Spectral clustering and the high-dimensional stochastic blockmodel. *The Annals of Statistics*, 39(4):1878–1915, August 2011.
- Rustamov, R. M. Laplace-beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the Fifth Eurographics Symposium on Geometry Processing*, SGP ’07, pp. 225–233, 2007.
- Schreiner, J., Asirvatham, A., Praun, E., and Hoppe, H. Inter-surface mapping. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH ’04, pp. 870–877, 2004.
- Shen, Y., Huang, Q., Srebro, N., and Sanghavi, S. Normalized spectral map synchronization. In *NIPS*, pp. 4925–4933, 2016.
- Snively, N., Seitz, S. M., and Szeliski, R. Photo tourism: Exploring photo collections in 3d. *ACM Trans. Graph.*, 25(3):835–846, July 2006.
- van Kaick, O., Zhang, H. R., Hamarneh, G., and Cohen-Or, D. A survey on shape correspondence. In *EuroGraphics: State-of-the-Art Report*, pp. 61–82, 2010.
- Wang, F., Huang, Q., and Guibas, L. J. Image co-segmentation via consistent functional maps. In *ICCV*, pp. 849–856, 2013.
- Wang, F., Huang, Q., Ovsjanikov, M., and Guibas, L. J. Unsupervised multi-class joint image segmentation. In *CVPR*, pp. 3142–3149. IEEE Computer Society, 2014.
- Wang, L. and Singer, A. Exact and stable recovery of rotations for robust synchronization. *Information and Inference: A Journal of the IMA*, 2:145193, December 2013.
- Zhou, T., Lee, Y. J., Yu, S. X., and Efros, A. A. Flowweb: Joint image set alignment by weaving consistent, pixel-wise correspondences. In *CVPR*, pp. 1191–1200. IEEE Computer Society, 2015a.
- Zhou, X., Zhu, M., and Daniilidis, K. Multi-image matching via fast alternating minimization. *CoRR*, abs/1505.04845, 2015b.

# Supplemental Material to SMAC: Simultaneous Mapping and Clustering Using Spectral Decompositions

Chandrajit Bajaj<sup>1</sup>, Tingran Gao<sup>3</sup>, Zihang He<sup>2</sup>, Qixing Huang<sup>\*1</sup>, and Zhenxiao Liang<sup>†2</sup>

<sup>1</sup>Department of Computer Science, The University of Texas at Austin

<sup>2</sup>Institute for Interdisciplinary Information Sciences, Tsinghua University

<sup>3</sup>Department of Statistics, The University of Chicago

## Abstract

This document contains all the proofs to the main paper on SMAC: Simultaneous Mapping and Clustering Using Spectral Decomposition.

## 1 Organization of the Supplemental Material

We organize our paper’s supplemental material as follows. In Section 2, we present a set of tools for analyzing the stability of leading eigen-vectors of a perturbed data matrix. These stability bounds are expressed with respect to the algebraic constants of the original matrix and the perturbation matrix. In Section 3 we study how these stability bounds shape out under different noise models of the perturbation matrix. We then use these results to analyze the exact recovery conditions of PermSMAC in Section 4.

## 2 Stability of Eigen-decomposition in the Deterministic Setting

In this Section, we present our framework for analyzing the stability of eigen-decompositions. The framework is based on a few key lemmas regarding the stability of eigenvalues and eigenvectors (Section 2.1). Their proofs are deferred to Sections 2.2–2.5.

### 2.1 Key Lemmas

Given a symmetric matrix  $A \in \mathbb{R}^{n \times n}$ , we are interested in controlling the stability of the leading eigenvectors of a symmetric block matrix  $A \otimes (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T) \in \mathbb{R}^{nm \times nm}$  under a symmetric block noise matrix  $N \in \mathbb{R}^{nm \times nm}$ .

To state the eigen-decomposition stability problem. Let  $\lambda_i, 1 \leq i \leq n$  and  $\mathbf{s}_i, 1 \leq i \leq n$  be the eigenvalues and corresponding eigenvectors of  $A$ , respectively. Similarly, let  $\mu_i, 1 \leq i \leq nm$  and  $\mathbf{u}_i, 1 \leq i \leq nm$  be the eigenvalues and corresponding eigenvectors of  $\bar{A} = A \otimes (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T) + N$ , respectively. We are interested in bounding the difference between  $\lambda_i, 1 \leq i \leq k$  and  $\mu_{(i-1)m+j}, 1 \leq i \leq k, 1 \leq j \leq m$  as well as the difference between the column space  $(\mathbf{s}_1, \dots, \mathbf{s}_k) \otimes H_m$  and that of  $U_{k(m-1)} = (\mathbf{u}_1, \dots, \mathbf{u}_{k(m-1)})$ , where  $H_m \in \mathbb{R}^{m \times (m-1)}$  is a basis matrix of  $I - \frac{1}{m} \mathbf{1}\mathbf{1}^T$ . For convenience, we use  $S_k = (\mathbf{s}_1, \dots, \mathbf{s}_k)$  to denote the leading  $k$  eigenvectors of  $A$ , and use  $\bar{S}_k = (\mathbf{s}_{k+1}, \dots, \mathbf{s}_n)$  to collect its remaining eigenvectors. First, we present the following result regarding eigenvalue stability:

---

\*huangqx@cs.utexas.edu

†liangzx96@gmail.com

**Lemma 2.1. (Eigenvalue Stability.)** Denote  $E_{11} = (S_k \otimes H_m)^T N (S_k \otimes H_m)$ . Suppose

$$\|N\| + \|E_{11}\| < \lambda_k - \lambda_{k+1},$$

then we have for every  $1 \leq i \leq k, 1 \leq j \leq m-1, p = (i-1)(m-1) + j$ ,

$$-\|E_{11}\| \leq \sigma_p - \lambda_i \leq \|E_{11}\| + \frac{\|N\| + \|E_{11}\|}{\lambda_i - \lambda_{k+1}} (\|N\| - \|E_{11}\|). \quad (1)$$

**Proof:** See Section 2.2 □

**Remark 1.** Note that (1) is tighter than the well-known Weyl's inequality  $|\lambda_1 - \sigma_i| \leq \|N\|$ , as

$$\|E_{11}\| + \frac{\|N\| + \|E_{11}\|}{\lambda_i - \lambda_{k+1}} (\|N\| - \|E_{11}\|) < \|E_{11}\| + (\|N\| - \|E_{11}\|) = \|N\|.$$

In particular, as  $\|E_{11}\| \ll \|N\|$  in most cases, so our bound is significantly better than the Weyl's inequality when  $\|N\| + \|E_{11}\| \ll \lambda_k - \lambda_{k+1}$ .

Moreover, our result is also tighter than that of [Eldridge et al., 2018], which introduces a similar stability bound:

$$-\|E_{11}\| \leq \sigma_p - \lambda_i \leq \|E_{11}\| + \frac{\|N\|^2}{\lambda_i - \lambda_{k+1} - \|N\| + \|E_{11}\|}.$$

(1) is tighter as  $\lambda_i - \lambda_{k+1} - \|N\| + \|E_{11}\| \leq \lambda_i - \lambda_{k+1}$  and  $\|N\|^2 - \|E_{11}\|^2 \leq \|N\|^2$ .

**Remark 2.** To gain additional intuitions on (1), let us consider the special case where  $k = m = 1$ , and let  $\lambda_i(A)$  be the  $i$ -th greatest eigenvalues of symmetric matrix  $A$ . For the matrix function  $A(t) = A + tN$  depending on the parameter  $t$  with  $0 \leq t \leq 1$ , let  $\lambda_i(t) = \lambda_i(A(t))$ , and  $\mathbf{u}_i(t)$  corresponds to eigenvalue  $\lambda_i(t)$ . First of all, it is easy to show that (e.g., through differentiating  $A(t)\mathbf{u}_1(t) = \lambda_1(t)\mathbf{u}_1(t)$  and utilizing  $\|\mathbf{u}(t)\| = 1$ )

$$\lambda_1'(t) = \mathbf{u}_1(t)^T A'(t) \mathbf{u}_1(t) = \mathbf{u}_1(t)^T N \mathbf{u}_1(t). \quad (2)$$

Recall that  $E_{11} = S_k^T N S_k = \mathbf{u}_1(t)^T N \mathbf{u}_1(t)$  is a scalar in this special case. Hence the first order expansion of  $\lambda_1(t)$  is  $\lambda_1(t) \approx \lambda(A) + tE_{11}$ . We now proceed to differentiate (2), which gives

$$\lambda_1''(t) = \mathbf{u}_1'(t)^T N \mathbf{u}_1(t) + \mathbf{u}_1(t)^T N \mathbf{u}_1'(t) = 2\mathbf{u}_1(t)^T N \mathbf{u}_1'(t). \quad (3)$$

Note that the expression of  $\mathbf{u}_1'(t)$  is given by

$$\mathbf{u}_1'(t) = \sum_{i=2}^n \frac{\mathbf{u}_i(t)^T N \mathbf{u}_1(t)}{\lambda(t) - \lambda_i(t)} \mathbf{u}_i(t).$$

Substituting it into formula (3) yields the second order approximation of  $\lambda(t)$  would be

$$\lambda_1(t) \approx \lambda_1(A) + E_{11}t + \sum_{i=2}^n \frac{|\mathbf{u}_i(t)^T N \mathbf{u}_1(t)|^2}{\lambda_1(A) - \lambda_i(A)} t^2,$$

or loosely speaking,

$$\lambda_1(t) \approx \lambda_1(A) + E_{11}t + \sum_{i=2}^n \frac{|\mathbf{u}_i(t)^T N \mathbf{u}_1(t)|^2}{\lambda_1(A) - \lambda_2(A)} t^2,$$

in which

$$\begin{aligned} \sum_{i=2}^n |\mathbf{u}_i(t)^T N \mathbf{u}_1(t)|^2 &= \sum_{i=1}^n |\mathbf{u}_i(t)^T (N \mathbf{u}_1(t))|^2 - (\mathbf{u}_1(t)^T N \mathbf{u}_1(t))^2 \\ &= |N \mathbf{u}_1(t)|^2 - E_{11}^2 \\ &\leq \|N\|^2 - \|E_{11}\|^2. \end{aligned}$$



Thus the second-order Taylor expansion of  $\lambda(t)$  implies

$$\|E_{11}\| + \frac{\|N\|^2 - \|E_{11}\|^2}{\lambda_1(A) - \lambda_2(A)}$$

is a good approximation for  $|\lambda_1(A + N) - \lambda_1(A)|$ , which is exactly what we proposed in formula (1).

However, it must be pointed out that this expansion is just an approximation rather than a real bound but it provides a good insight why (1) comes out. Besides, it would also be hard to generalize this expansion idea to  $k > 1$  since  $\|E_{11}\|$  will no longer be a scalar if  $k > 1$ .

To characterize the difference between  $S_k \otimes H_m$  and  $U_{k(m-1)} := (\mathbf{u}_1, \dots, \mathbf{u}_{k(m-1)})$ , we consider the following decomposition of  $U_{k(m-1)}$ :

$$U_{k(m-1)} = (S_k \otimes H_m)X + Y,$$

where

$$X \in \mathbb{R}^{k(m-1) \times k(m-1)}, \quad S_k^T Y = 0.$$

In other words,  $(S_k \otimes H_m)X$  is the projection of  $U_{k(m-1)}$  onto the column space of  $S_k \otimes H_m$ , and  $Y$  is the projection of  $U_{k(m-1)}$  onto the dual space  $\bar{S}_k \otimes H_m$ . Intuitively, we say  $U_{k(m-1)}$  is stable if  $Y$  is small and  $X$  is close to a unitary matrix (which defines all orthogonal basis of a linear space). The following Lemma provides a bound on the difference between  $X$  and a unitary matrix:

**Lemma 2.2. (Controlling  $X$ .)** Denote  $E_{11} = (S_k \otimes H_m)^T N (S_k \otimes H_m)$ . Suppose  $\lambda_k - \lambda_{k+1} > \|N\| + \|E_{11}\|$ . Then there exists a unitary matrix  $R \in O(k(m-1))$  so that

$$\|X - R\| \leq 1 - \sqrt{1 - \left( \frac{\|N\|}{\lambda_k - \lambda_{k+1} - \|E_{11}\|} \right)^2}. \quad (4)$$

In particular,

$$\|X - R\| \leq \frac{\|N\|^2}{(\lambda_k - \lambda_{k+1} - \|E_{11}\|)^2}. \quad (5)$$

**Proof:** See Appendix 2.3. □

It is easy to derive an upper bound on  $\|Y\|$  using Lemma 2.2. However, our analysis requires bounding individual blocks  $Y_i, 1 \leq i \leq n$ , making such bounds insufficient for our purpose. To this end, we introduce the following expression of  $Y$ , from which we will derive block-wise bounds.

**Lemma 2.3. (Controlling  $Y$ .)** Denote  $\Sigma_i = \text{diag}(\sigma_{(i-1)(m-1)+1}, \dots, \sigma_{i(m-1)}), 1 \leq i \leq k, \bar{\Lambda} = \text{diag}(\lambda_{k+1}, \dots, \lambda_n)$ , and  $\bar{S}_k = (\mathbf{s}_{k+1}, \dots, \mathbf{s}_n)$ . Let

$$B_i = (\bar{S}_k \otimes H_m)(\lambda_i I - \bar{\Lambda} \otimes I_{m-1})^{-1}(\bar{S}_k \otimes H_m)^T, \quad 1 \leq i \leq k. \quad (6)$$

Suppose  $\|N\| + \|E_{11}\| < \lambda_k - \lambda_{k+1}$ , then

$$Y_i := (\mathbf{y}_{(i-1)m+1}, \dots, \mathbf{y}_{im}) = \sum_{l=0}^{\infty} ((I - B_i N)^{-1} B_i)^{l+1} N (S_k \otimes H_m) X_i (\lambda_i I_m - \Sigma_i)^l. \quad (7)$$

**Proof:** See Appendix 2.4. □

We now present the last lemma which applies (7) to obtain a  $L^\infty$ -type bound on blocks of  $Y_i$ . Specifically, let  $E_b = \mathbf{e}_b \otimes H_m \in \mathbb{R}^{n(m-1) \times (m-1)}$ . The following Lemma provides a way to bound  $\|E_b^T Y\|$ :

**Lemma 2.4. (Bounding  $L^\infty$ -norm of  $Y$ )** Given  $i \in \{1, \dots, k\}$ . Let  $B_i$  be defined by (6). Suppose there are small constants  $\epsilon_1, \epsilon_2, \epsilon_3, \delta < 1$  such that the following four conditions are satisfied:

- $\|B_i\| \|N\| \leq \epsilon_1$ .

- $\|B_i\| \|\lambda_i I_m - \Sigma_i\| \leq \epsilon_2$ .
- $\exists j_0 \geq 0, m_0 \geq 0, s.t., \forall 0 \leq j \leq j_0, i_l \geq 0, 1 \leq l \leq j+1, \text{ where } 0 \leq \sum_{l=1}^{j+1} i_l \leq m_0,$

$$\max_{1 \leq b \leq n} \|E_b^T (B_i N)^{i_1} B_i \cdots (B_i N)^{i_{j+1}} B_i N (S_k \otimes H_m)\| \leq \epsilon_3 \|B_i\|^j \cdot \delta^{\sum_{l=1}^{j+1} i_l + 1}. \quad (8)$$

- $\epsilon_2 + \delta < 1$ .

Then

$$\max_{1 \leq b \leq n} \|E_b^T Y_i\| \leq \|X_i\| \cdot \left( \frac{\epsilon_1}{1 - \epsilon_1 - \epsilon_2} \cdot \left( \left( \frac{\epsilon_2}{1 - \epsilon_1} \right)^{j_0+1} + \left( \frac{\epsilon_1}{1 - \epsilon_2} \right)^{m_0} \right) + \frac{\delta}{1 - \epsilon_2 - \delta} \epsilon_3 \right). \quad (9)$$

*Proof:* See Appendix 2.5 □

**Remark 3.** As we will see later, the dominant term in (8) is  $\|E_1^T BNS\|$ , which can be controlled using standard concentration bound. The technical difficulty is how to extend it to high order moments. Later we will show how to achieve this goal by controlling power moments. Note that (8) does not incur a strong bound. For example, in randomized models we consider in this paper,  $\|B\| \|N\| = O\left(\frac{1}{\sqrt{\log(n)}}\right)$ . However, the right-hand side of (8) only decays at a geometric rate.

**Remark 4.** It turns out the major task in terms of controlling eigen-decomposition stability is to provide bounds on  $E_{11}$ ,  $N$  and the left-hand sides in (8). This is the goal of the next two sections.

## 2.2 Proof of Lemma 2.1

Denote  $\Lambda_k = (\lambda_1, \dots, \lambda_k)$  and  $\bar{\Lambda}_k = (\lambda_{k+1}, \dots, \lambda_n)$ . It is clear that we can decompose  $A$  as

$$A = S_k \Lambda_k S_k^T + \bar{S}_k \bar{\Lambda}_k \bar{S}_k^T.$$

For convenience we set  $p = (i-1)(m-1) + j$  and define  $L_m = I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T$ . To control  $|\lambda_i - \sigma_{(i-1)(m-1)+j}|, 1 \leq i \leq k, 1 \leq j \leq m-1$ , we consider an eigen-decomposition of  $A \otimes L_m + N$  with respect to basis spanned by  $S_k \otimes H_m$  and  $\bar{S}_k \otimes H_m$ . Introduce matrices

$$\begin{aligned} E_{11} &= (S_k \otimes H_m)^T N (S_k \otimes H_m), \\ E_{21} &= (\bar{S}_k \otimes H_m)^T N (S_k \otimes H_m), \\ E_{22} &= (\bar{S}_k \otimes H_m)^T N (\bar{S}_k \otimes H_m). \end{aligned}$$

Using unitary matrix  $S \otimes H_m$  to change the basis for matrices  $A \otimes L_m$  and  $N$ , respectively, we obtain that

$$\begin{aligned} (S \otimes H_m)^T (A \otimes L_m) (S \otimes H_m) &= \begin{bmatrix} \Lambda_k & 0 \\ 0 & \bar{\Lambda}_k \end{bmatrix} \otimes I_{m-1} \\ (S \otimes H_m)^T N (S \otimes H_m) &= \begin{bmatrix} E_{11} & E_{21}^T \\ E_{21} & E_{22} \end{bmatrix} \end{aligned}$$

Observe that the matrix  $A \otimes L_m + N - \mu I$  is congruent to

$$B(\mu) = \begin{bmatrix} \Lambda_k \otimes I_{m-1} + E_{11} - \mu I + E_{21}^T (\mu I - \bar{\Lambda}_k - E_{22})^{-1} E_{21} & 0 \\ 0 & \bar{\Lambda}_k \otimes I_{m-1} + E_{22} - \mu I \end{bmatrix}$$

and has the same inertia as well, which can be verified from the identity

$$\begin{aligned} & (S \otimes H_m)^T (A \otimes L_m + N - \mu I) (S \otimes H_m) \\ &= \begin{bmatrix} \Lambda_k \otimes I_{m-1} + E_{11} - \mu I & E_{21}^T \\ E_{21} & \bar{\Lambda}_k \otimes I_{m-1} + E_{22} - \mu I \end{bmatrix} \\ &= \begin{bmatrix} I & P \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \Lambda_k \otimes I_{m-1} + E_{11} - \mu I + E_{21}^T (\mu I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21} & 0 \\ 0 & \bar{\Lambda}_k \otimes I_{m-1} + E_{22} - \mu I \end{bmatrix} \begin{bmatrix} I & 0 \\ P^T & I \end{bmatrix}^{-1}, \end{aligned}$$

in which  $P = E_{21}^T(\mu I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1}$ . Hence by letting  $\mu = \sigma_p$  we have  $\lambda_p(A \otimes L_m + N - \sigma_p I) = 0$  and then  $\lambda_p(B(\sigma_p)) = 0$ , and further more

$$\sigma_p = \lambda_p \left( \begin{bmatrix} \Lambda_k \otimes I_{m-1} & 0 \\ 0 & \bar{\Lambda}_k \otimes I_{m-1} + E_{22} \end{bmatrix} + \begin{bmatrix} E_{11} + E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21} & 0 \\ 0 & 0 \end{bmatrix} \right).$$

However, the Weyl's inequality told that

$$\begin{aligned} \sigma_p - \lambda_i &= \sigma_p - \lambda_p(\Lambda \otimes I_{m-1}) \geq \lambda_{\min} \left( \begin{bmatrix} E_{11} + E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21} & 0 \\ 0 & 0 \end{bmatrix} \right) \\ &\geq \min\{\lambda_{\min}(E_{11} + E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21}), 0\} \\ &\geq -\|E_{11}\| \end{aligned} \quad (10)$$

in which we used the fact that  $\sigma_p - \bar{\Lambda}_k \otimes I_{m-1} - E_{22}$  is positive definite, and

$$\begin{aligned} \sigma_p - \lambda_i &= \sigma_p - \lambda_p(\Lambda \otimes I_{m-1}) \leq \lambda_{\max} \left( \begin{bmatrix} E_{11} + E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21} & 0 \\ 0 & 0 \end{bmatrix} \right) \\ &\leq \|E_{11} + E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21}\| \\ &\leq \|E_{11} + E_{21}^T((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21}\| \end{aligned} \quad (11)$$

It remains to bound  $E_{21}^T(\sigma_p I - \bar{\Lambda}_k \otimes I_{m-1} - E_{22})^{-1} E_{21}$ . Towards this end, we consider an arbitrary value  $\mu > \lambda_{\max}(N)$ . It is clear that

$$\mu I - \begin{bmatrix} E_{11} & E_{21}^T \\ E_{21} & E_{22} \end{bmatrix} \quad (12)$$

is a positive definite matrix, and  $\mu I - E_{22}$  is also positive definite. From the identity

$$\mu I - \begin{bmatrix} E_{11} & E_{21}^T \\ E_{21} & E_{22} \end{bmatrix} = \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} \mu I - E_{11} - E_{21}^T(\mu I - E_{22})^{-1} E_{21} & 0 \\ 0 & \mu I - E_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ Q^T & I \end{bmatrix}^{-1}$$

in which  $Q = E_{21}^T(E_{22} - \mu I)^{-1}$ , we can see that  $\mu I - E_{11} - E_{21}^T(\mu I - E_{22})^{-1} E_{21}$  is positive definite as well. Applying it to (12), we obtain that

$$\mu I - E_{11} - E_{21}^T(\mu I - E_{22})^{-1} E_{21} \succeq 0,$$

or equivalently,

$$E_{11} + E_{21}^T(\mu I - E_{22})^{-1} E_{21} \preceq \mu I. \quad (13)$$

Similarly, for  $\mu < \lambda_{\min}(N)$ , we have

$$E_{11} + E_{21}^T(\mu I - E_{22})^{-1} E_{21} \succeq \mu I. \quad (14)$$

Letting  $\mu \rightarrow \|N\|$  and  $\mu \rightarrow -\|N\|$  respectively, we obtain that

$$\begin{aligned} E_{21}^T(\|N\|I - E_{22})^{-1} E_{21} &\preceq \|N\|I - E_{11}, \\ E_{21}^T(\|N\|I + E_{22})^{-1} E_{21} &\preceq \|N\|I + E_{11}. \end{aligned}$$

Here the second inequality arise by taking negative of (14).

As thus for any non-negative number  $\alpha, \beta$ , we have

$$E_{11} + E_{21}^T(\alpha(\|N\|I - E_{22})^{-1} + \beta(\|N\|I + E_{22})^{-1})^{-1} E_{21} \preceq (\alpha + \beta)\|N\|I + (1 + \beta - \alpha)E_{11}.$$

If we in addition have

$$E_{21}^T((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \preceq E_{21}^T(\alpha(\|N\|I - E_{22})^{-1} + \beta(\|N\|I + E_{22})^{-1})^{-1} E_{21}, \quad (15)$$

for some  $\alpha, \beta \geq 0$ , then we reach an upper bound

$$\lambda_{\max} \left( E_{11} + E_{21}^T ((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \right) \leq (\alpha + \beta) \|N\| + |1 + \beta - \alpha| \|E_{11}\| \quad (16)$$

To ensure inequality (15), we need to show that the following inequality

$$\frac{1}{\sigma_p - \lambda_{k+1} - x} \leq \frac{\alpha}{\|N\| - x} + \frac{\beta}{\|N\| + x}$$

holds for every eigenvalue  $x$  of  $E_{22}$ , or in a stronger sense, for all  $-\|N\| \leq x \leq \|N\|$ . Define

$$\theta = \beta/\alpha, \varepsilon = \|N\|, a = \|E_{11}\|$$

and then the above condition changes to

$$\alpha \geq \frac{\varepsilon^2 - x^2}{(\sigma_p - \lambda_{k+1} - x)((1 + \theta)\varepsilon + (1 - \theta)x)}$$

for all  $-\varepsilon \leq x \leq \varepsilon$ . Hence we can take  $\alpha$  as

$$\begin{aligned} \alpha_\theta &= \max_{-\varepsilon \leq x \leq \varepsilon} \frac{\varepsilon^2 - x^2}{(\sigma_p - \lambda_{k+1} - x)((1 + \theta)\varepsilon + (1 - \theta)x)} \\ &= \frac{\varepsilon}{(\sigma_p - \lambda_{k+1})(1 + \theta)} \max_{-1 \leq x \leq 1} \frac{1 - x^2}{(1 - \frac{\varepsilon}{\sigma_p - \lambda_{k+1}}x)(1 + \frac{1 - \theta}{1 + \theta}x)} \end{aligned}$$

and substitute it into inequality (16) we obtain

$$\lambda_{\max} \left( E_{11} + E_{21}^T ((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \right) \leq (1 + \theta)\alpha_\theta \varepsilon + |1 + (\theta - 1)\alpha_\theta| a.$$

But it is easy to see that  $\alpha_\theta \leq \alpha_0 \leq \frac{2\varepsilon}{\sigma_p - \lambda_{k+1} + \varepsilon} \leq \frac{2\varepsilon}{\lambda_k - \lambda_{k+1}} \leq 1$ , so the absolute value sign can be removed from the inequality above, and substitute the expression of  $\alpha_\theta$  into it:

$$\lambda_{\max} \left( E_{11} + E_{21}^T ((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \right) \leq (1 + \theta)\alpha_\theta \varepsilon + (\theta - 1)\alpha_\theta a + a \quad (17)$$

To simplify our computation, we introduce the following trigonometric function notations:

$$\begin{aligned} \sin \phi &= \frac{1 - \theta}{1 + \theta}, & -\frac{\pi}{2} &\leq \phi \leq \frac{\pi}{2}; \\ \sin \psi &= \frac{\varepsilon}{\sigma_p - \lambda_{k+1}}, & 0 &\leq \psi < \frac{\pi}{6}; \\ \sin \omega &= \frac{a}{\varepsilon}, & 0 &\leq \omega \leq \frac{\pi}{2} \end{aligned} \quad (18)$$

Applying proposition 1 under these notations, we can rewrite  $\alpha_\theta$  as

$$\alpha_\theta = \frac{\sin \psi}{1 + \theta} \cdot \frac{1}{\cos^2 \frac{\phi - \psi}{2}}, \quad (19)$$

and substitute it into (17) to obtain

$$\begin{aligned} \lambda_{\max} \left( E_{11} + E_{21}^T ((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \right) &\leq (1 + \theta)\alpha_\theta \varepsilon + (\theta - 1)\alpha_\theta a + a \\ &= a + \frac{\varepsilon \sin \psi}{\cos^2 \frac{\phi - \psi}{2}} (1 - \sin \omega \sin \phi) \end{aligned} \quad (20)$$



By optimizing over  $\theta$ , namely over  $\phi$ , and applying proposition 2, we have

$$\begin{aligned}
& \lambda_{\max} \left( E_{11} + E_{21}^T ((\sigma_p - \lambda_{k+1})I - E_{22})^{-1} E_{21} \right) \\
& \leq a + \varepsilon \sin \psi \min_{-\pi/2 \leq \phi \leq \pi/2} \frac{1 - \sin \omega \sin \phi}{\cos^2 \frac{\phi - \psi}{2}} \\
& = a + \varepsilon \frac{\sin \psi \cos^2 \omega}{1 + \sin \omega \sin \psi} \\
& = \|E_{11}\| + \frac{\|N\|^2 - \|E_{11}\|^2}{\sigma_p - \lambda_{k+1} + \|E_{11}\|} \\
& \leq \|E_{11}\| + \frac{\|N\|^2 - \|E_{11}\|^2}{\lambda_i - \lambda_{k+1}}, \tag{21}
\end{aligned}$$

where in the last step we used the fact  $\sigma_p - \lambda_i \geq -\|E_{11}\|$  that has been proved before.  $\square$

**Proposition 1.** *Given two real number  $\alpha, \beta$  with  $-\frac{\pi}{2} \leq \alpha, \beta \leq \frac{\pi}{2}$ , we have*

$$\max_{-1 \leq x \leq 1} \frac{1 - x^2}{(1 + x \sin \alpha)(1 + x \sin \beta)} = \frac{1}{\cos^2 \frac{\alpha + \beta}{2}}$$

*Proof.* It is trivial when  $x = \pm 1$  so we always assume  $-1 < x < 1$  in the following. Taking transformation  $x = \frac{1-p^2}{1+p^2}, p \in \mathbb{R} \setminus \{0\}$ , we have

$$\begin{aligned}
& \frac{(1 + x \sin \alpha)(1 + x \sin \beta)}{1 - x^2} \\
& = \frac{1}{4} [(1 - \sin \alpha)p + (1 + \sin \alpha)p^{-1}] [(1 - \sin \beta)p + (1 + \sin \beta)p^{-1}] \\
& = \frac{1}{4} [(1 - \sin \alpha)(1 - \sin \beta)p^2 + 2(1 - \sin \alpha \sin \beta) + (1 + \sin \alpha)(1 + \sin \beta)p^{-2}] \\
& \geq \frac{1}{2} (1 - \sin \alpha \sin \beta + \cos \alpha \cos \beta) \\
& = \frac{1}{2} (1 + \cos(\alpha + \beta)) \\
& = \cos^2 \frac{\alpha + \beta}{2}.
\end{aligned}$$

The proposition follows immediately.  $\square$

**Proposition 2.** *Given  $\phi, \psi, \omega$  as define in [\(18\)](#), we claim that*

$$\min_{-\pi/2 \leq \phi \leq \pi/2} \frac{1 - \sin \omega \sin \phi}{\cos^2 \frac{\phi - \psi}{2}} = \frac{\cos^2 \omega}{1 + \sin \omega \sin \psi}$$

*Proof.* Note that

$$\begin{aligned}
& 1 - \sin \omega \sin \phi \\
& = 1 - \sin \omega \sin(\phi - \psi + \psi) \\
& = 1 - \sin \omega (\sin \psi \cos(\phi - \psi) + \cos \psi \sin(\phi - \psi)) \\
& = 1 - \sin \omega \left( \sin \psi \left( \cos^2 \frac{\phi - \psi}{2} - \sin^2 \frac{\phi - \psi}{2} \right) + 2 \cos \psi \sin \frac{\phi - \psi}{2} \cos \frac{\phi - \psi}{2} \right).
\end{aligned}$$

Define  $p = \tan \frac{\phi - \psi}{2}$ .  $p$  can be taken over interval  $[\tan(-\pi/4 - \psi/2), \tan(\pi/4 - \psi/2)]$  as taking  $\phi$  over  $[-\pi/2, \pi/2]$ . Some trigonometric calculation shows that  $[\tan(-\pi/4 - \psi/2), \tan(\pi/4 - \psi/2)]$  is equal to

$[-\frac{\cos \psi}{1-\sin \psi}, \frac{\cos \psi}{1+\sin \psi}]$ . Using the identity we just proved, we have

$$\begin{aligned} & \frac{1 - \sin \omega \sin \phi}{\cos^2 \frac{\phi - \psi}{2}} \\ &= p^2 + 1 - \sin \omega ((1 - p^2) \sin \psi + 2p \cos \psi) \\ &= (1 + \sin \omega \sin \psi) p^2 - 2p \sin \omega \cos \psi + (1 - \sin \omega \sin \psi) \end{aligned}$$

We can see it is a quadratic function about  $p$ , so it has a minimum at  $p_0 = \frac{\sin \omega \cos \psi}{1 + \sin \omega \sin \psi}$ . It is clear that  $-\frac{\cos \psi}{1 - \sin \psi} \leq p_0 \leq \frac{\cos \psi}{1 + \sin \psi}$ , hence  $p_0$  can be certainly taken. Further more, this quadratic function has a minimum value  $\frac{\cos^2 \omega}{1 + \sin \omega \sin \psi}$  on this point, which completes our proof.  $\square$

### 2.3 Proof of Lemma 2.2

Also denote  $I - \frac{1}{m} \mathbf{1}\mathbf{1}^T$  as  $L_m$ . We only prove the first inequality, since the second inequality can be inferred from

$$1 - \sqrt{1 - u^2} \leq u, \quad -1 \leq u \leq 1.$$

Multiply both sides of

$$(A \otimes L_m + N)((S_k \otimes H_m)X + Y) = ((S_k \otimes H_m)X + Y)\Sigma$$

by  $(\bar{S}_k \otimes L_m)^T$ , yielding

$$\begin{aligned} & (\bar{S}_k \otimes H_m)^T (A \otimes L_m + N)((S_k \otimes H_m)X + Y) = (\bar{S}_k \otimes H_m)^T Y \cdot \Sigma \\ & \Leftrightarrow (\bar{S}_k \otimes H_m)^T N((S_k \otimes H_m)X + Y) = (\bar{S}_k \otimes H_m)^T Y \cdot \Sigma - ((\bar{S}_k \bar{\Lambda}) \otimes H_m)^T Y. \end{aligned} \quad (22)$$

Now we prove the following proposition, which will be used later:

**Proposition 3.** Denote  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_k)$  and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$ . Suppose  $\lambda_1 \geq \dots \lambda_k > \sigma_1 \geq \dots \geq \sigma_n$ , then for any  $Y \in \mathbb{R}^{n \times k}$ , we have

$$\|Y\Lambda - \Sigma Y\| \geq (\lambda_k - \sigma_1)\|Y\|.$$

*Proof:* Without losing generality, we can assume  $\sigma_n \geq 0$ , since we can always shift  $\lambda_i, 1 \leq i \leq k$  and  $\sigma_j, 1 \leq j \leq n$  by the same amount without changing the value of  $Y\Lambda - \Sigma Y$  and  $\lambda_k - \sigma_1$ . With this assumption, the proof directly follows from triangle inequality:

$$\|Y\Lambda - \Sigma Y\| \geq \|Y\Lambda\| - \|\Sigma Y\| \geq \|Y(\lambda_k I_k)\| - \|(\sigma_1 I_n)Y\| = (\lambda_k - \sigma_1)\|Y\|.$$

$\square$

Now let us come back to the proof of Lemma 2.2. Since the columns of  $U_{k(m-1)} = (S_k \otimes H_m)X + Y$  are orthogonal, we have

$$\begin{aligned} \|N\| &\geq \|(\bar{S}_k \otimes H_m)^T N\| \geq \|(\bar{S}_k \otimes H_m)^T N((S_k \otimes H_m)X + Y)\| \\ &= \|(\bar{S}_k \otimes H_m)^T Y \cdot \Sigma - (\bar{\Lambda} \otimes I_{m-1}) \cdot (\bar{S}_k \otimes H_m)^T Y\| \\ &\geq \min_{1 \leq i \leq km} |\sigma_i - \lambda_{k+1}| \|(\bar{S}_k \otimes H_m)^T Y\| \quad (\text{Applying Proposition 3}) \\ &= \min_{1 \leq i \leq km} |\sigma_i - \lambda_{k+1}| \|Y\| \\ &\geq (\lambda_k - \lambda_{k+1} - \|E_{11}\|)\|Y\|, \end{aligned} \quad (23)$$

where the last equality is due to Lemma 2.1

As  $V^T V = X^T X + Y^T Y = I_{k(m-1)}$ . It follows that

$$\sqrt{1 - \left( \frac{\|N\|}{\lambda_k - \lambda_{k+1} - \|E_{11}\|} \right)^2} \leq \sqrt{1 - \|Y\|^2} = \sigma_{\min}(X) \leq \sigma_{\max}(X) \leq 1.$$

Let  $X = V_X \Sigma_X W_X^T$  be the SVD of  $X$ . Define  $R = V_X W_X^T$ . We have

$$\begin{aligned} \|X - R\| &= \|V_X \Sigma_X W_X^T - V_X W_X^T\| \\ &\leq \|\Sigma_X - I\| \\ &\leq 1 - \sqrt{1 - \left( \frac{\|N\|}{\lambda_k - \lambda_{k+1} - \|E_{11}\|} \right)^2}, \end{aligned}$$

which ends the proof.  $\square$

## 2.4 Proof of Lemma 2.3

First, we have

$$\begin{aligned} &(\bar{S}_k \otimes H_m)(\lambda_i I - \bar{\Lambda} \otimes I_{m-1} - E_{22})^{-1}(\bar{S}_k \otimes H_m)^T \\ &= (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}} \left( I - (\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}} (\bar{S}_k \otimes H_m)^T N (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}} \right)^{-1} \\ &\quad (\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}} (\bar{S}_k \otimes H_m)^T \\ &= \sum_{k=0}^{\infty} \left( (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-1} (\bar{S}_k \otimes H_m)^T N \right)^k (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-1} (\bar{S}_k \otimes H_m)^T \\ &= (I - B_i N)^{-1} B_i, \end{aligned}$$

in which

$$B_i = (\bar{S}_k \otimes H_m)(\lambda_i I - \bar{\Lambda} \otimes I_{m-1})^{-1} (\bar{S}_k \otimes H_m)^T$$

In the above argument, since

$$\|(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}} (\bar{S}_k \otimes H_m)^T N (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1}))^{-\frac{1}{2}}\| \leq \frac{\|N\|}{\lambda_k - \lambda_{k+1}} < 1,$$

we can safely apply Taylor expansion.

Now let us consider each column of  $Y$ . Denote  $p = (i-1) \cdot (m-1) + j$ . By solving linear system for  $X, Y$ , we have

$$\begin{aligned} \mathbf{y}_p &= (\bar{S}_k \otimes H_m)(\sigma_p I - (\bar{\Lambda} \otimes I_{m-1} + E_{22}))^{-1} (\bar{S}_k \otimes H_m)^T N (S_k \otimes H_m) \mathbf{x}_p \\ &= (\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1} + E_{22}) - (\lambda_i - \sigma_p)I)^{-1} (\bar{S}_k \otimes H_m)^T N (S_k \otimes H_m) \mathbf{x}_p \\ &= \sum_{k=0}^{\infty} (\bar{S}_k \otimes H_m) \left( (\lambda_i I - (\bar{\Lambda} \otimes I_{m-1} + E_{22}))^{-1} \right)^{k+1} (\bar{S}_k \otimes H_m)^T N (S_k \otimes H_m) \mathbf{x}_p (\lambda_i - \sigma_p)^k \\ &= \sum_{k=0}^{\infty} \left( (\bar{S}_k \otimes H_m)(\lambda_i I - ((\bar{\Lambda} \otimes I_{m-1}) + E_{22}))^{-1} (S \otimes H_m)^T \right)^{k+1} N (\bar{S}_k \otimes H_m) \mathbf{x}_p (\lambda_i - \sigma_p)^k \quad (24) \\ &= \sum_{k=0}^{\infty} \left( (I - B_i N)^{-1} B_i \right)^{k+1} N (S_k \otimes H_m) \mathbf{x}_p (\lambda_i - \sigma_p)^k \end{aligned}$$

Like before, from  $\|N\| + \|E_{11}\| < \lambda_k - \lambda_{k+1}$  it is easy to check that

$$\begin{aligned}
& (\lambda_i - \sigma_p) \|(\bar{S}_k \otimes H_m)(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1} + E_{22}))^{-1}(\bar{S}_k \otimes H_m)^T\| \\
& \leq (\lambda_i - \sigma_p) \|(\lambda_i I - (\bar{\Lambda} \otimes I_{m-1} + E_{22}))^{-1}\| \\
& \leq \|E_{11}\| \times \frac{1}{\lambda_i - \lambda_{k+1} - \|N\|} \\
& \leq \frac{\|E_{11}\|}{\lambda_k - \lambda_{k+1} - \|N\|} < 1,
\end{aligned}$$

which shows that the power series above would be surely convergent.

Putting this in the matrix form leads to

$$Y_{\cdot,i} = \sum_{l=0}^{\infty} \left( (I - B_i N)^{-1} B_i \right)^{l+1} N(S_k \otimes H_m) X_{\cdot,i} (\lambda_i I - \Sigma_k^{(i)})^l.$$

□

## 2.5 Proof of Lemma 2.4

First of all, since  $\|B_i N\| \leq \|B_i\| \cdot \|N\| < 1$ , it is clear that

$$(I - B_i N)^{-1} B_i = \sum_{j \geq 0} (B_i N)^j B_i. \quad (25)$$

Consider the following three terms:

$$\begin{aligned}
\delta_1 &:= \sum_{j \geq j_0+1} ((I - B_i N)^{-1} B_i)^{j+1} N(S_k \otimes H_m) X_i (\lambda_i I_m - \Sigma_i)^j. \\
\delta_2 &:= \sum_{j=0}^{j_0} \sum_{m \geq m_0+1} \sum_{\sum_{l=1}^{j+1} i_l = m} \left( \prod_{l=1}^{j+1} (B_i N)^{i_l} B \right) N(S_k \otimes H_m) X_i (\lambda_i I_m - \Sigma_i)^j. \\
\delta_3 &:= \sum_{j=0}^{j_0} \sum_{m=0}^{m_0} \sum_{\sum_{l=1}^{j+1} i_l = m} \left( \prod_{l=1}^{j+1} (B_i N)^{i_l} B \right) N(S_k \otimes H_m) X_i (\lambda_i I_m - \Sigma_i)^j.
\end{aligned} \quad (26)$$

It is clear that

$$Y_i = \delta_1 + \delta_2 + \delta_3.$$

Now we bound  $E_b^T \delta_k, 1 \leq k \leq 3$ . First of all, we have

$$\begin{aligned}
\|E_b^T \delta_1\| &\leq \|\delta_1\| \leq \sum_{j \geq j_0+1} \|(I - B_i N)^{-1} B_i\|^{j+1} \|N\| \|S_k \otimes H_m\| \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&\leq \sum_{j \geq j_0+1} \left( \frac{\|B_i\|}{1 - \|B_i\| \|N\|} \right)^{j+1} \|N\| \|S_k \otimes H_m\| \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&= \frac{\|B_i\| \|N\| \|X\|}{1 - \|B_i\| \|N\|} \sum_{j \geq j_0+1} \left( \frac{\|B_i\| \|\lambda_i I_m - \Sigma_i\|}{1 - \|B_i\| \|N\|} \right)^j \\
&= \frac{\epsilon_1}{1 - \epsilon_1 - \epsilon_2} \cdot \left( \frac{\epsilon_2}{1 - \epsilon_1} \right)^{j_0+1} \|X_i\|.
\end{aligned} \quad (27)$$



Similarly,

$$\begin{aligned}
\|E_b^T \delta_2\| &\leq \|\delta_2\| \leq \sum_{j=0}^{j_0} \sum_{m \geq m_0+1} \sum_{\sum_{l=1}^{j+1} i_l = m} \left( \prod_{l=1}^{j+1} (\|B_i\| \|N\|)^{i_l} \|B_i\| \right) \|N\| \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&= \sum_{j=0}^{j_0} \sum_{m \geq m_0+1} \sum_{\sum_{l=1}^{j+1} i_l = m} (\|B_i\|^{m+j+1} \|N\|^{m+1}) \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&= \sum_{j=0}^{j_0} \sum_{m \geq m_0+1} C_{m+j}^j (\|B_i\|^{m+j+1} \|N\|^{m+1}) \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&= \sum_{j=0}^{j_0} (\|B_i\| \|\lambda_i I_m - \Sigma_i\|)^j \sum_{m \geq m_0+1} C_{m+j}^j (\|B_i\| \|N\|)^{m+1} \|X_i\| \\
&\leq \|X_i\| \sum_{m \geq m_0+1} \epsilon_1^{m+1} \sum_{j=0}^{j_0} \epsilon_2^j C_{m+j}^j.
\end{aligned} \tag{28}$$

Since when  $-1 < x < 1$ ,

$$\frac{1}{(1-x)^{j+1}} = \sum_{m=0}^{\infty} x^m C_{m+j}^m.$$

It follows that

$$\begin{aligned}
\|E_b^T \delta_2\| &\leq \|X_i\| \sum_{m \geq m_0+1} \epsilon_1^{m+1} \sum_{j=0}^{j_0} \epsilon_2^j C_{m+j}^j \\
&\leq \|X_i\| \sum_{m \geq m_0+1} \epsilon_1^{m+1} \frac{1}{(1-\epsilon_2)^{m+1}} \\
&= \|X_i\| \left( \frac{\epsilon_1}{1-\epsilon_2} \right)^{m_0} \frac{\epsilon_1}{1-\epsilon_1-\epsilon_2}.
\end{aligned} \tag{29}$$

We use a similar strategy to bound  $\delta_3$ . In fact,

$$\begin{aligned}
\|E_b^T \delta_3\| &\leq \sum_{j=0}^{j_0} \sum_{m=0}^{m_0} \sum_{\sum_{l=1}^{j+1} i_l = m} \|E_b^T \left( \prod_{l=1}^{j+1} (B_i N)^{i_l} B_i \right) N(S_k \otimes H_m)\| \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&\leq \sum_{j=0}^{j_0} \sum_{m=0}^{m_0} \sum_{\sum_{l=1}^{j+1} i_l = m} \|B_i\|^j \epsilon_3 \delta^{m+1} \|X_i\| \|\lambda_i I_m - \Sigma_i\|^j \\
&\leq \|X_i\| \epsilon_3 \sum_{j=0}^{j_0} \epsilon_2^j \sum_{m=0}^{m_0} C_{m+j}^j \delta^{m+1} \\
&= \|X_i\| \epsilon_3 \sum_{m=0}^{m_0} \delta^{m+1} \sum_{j=0}^{j_0} C_{m+j}^j \epsilon_2^j \\
&\leq \|X_i\| \epsilon_3 \sum_{m=0}^{m_0} \delta^{m+1} \frac{1}{(1-\epsilon_2)^{m+1}} \\
&\leq \frac{\delta}{1-\epsilon_2-\delta} \cdot \epsilon_3 \cdot \|X_i\|.
\end{aligned} \tag{30}$$

Combing (27), (29) and (30), we have

$$\max_{1 \leq b \leq n} \|E_b^T Y_i\| \leq \|X_i\| \cdot \left( \frac{\epsilon_1}{1-\epsilon_1-\epsilon_2} \cdot \left( \left( \frac{\epsilon_2}{1-\epsilon_1} \right)^{j_0+1} + \left( \frac{\epsilon_1}{1-\epsilon_2} \right)^{m_0} \right) + \frac{\delta}{1-\epsilon_2-\delta} \epsilon_3 \right),$$

which ends the proof.  $\square$

### 3 Stability of Eigen-decomposition in Randomized Noise Models

#### 3.1 Key Lemmas

The next Lemma provides various bounds regarding the random matrix model discussed above: Given a constant  $c \geq 1$ . Suppose  $d_{\min} \geq 2c \log(n)$  and  $n \geq 10$ . Consider a random matrix  $N \in \mathbb{R}^{nm \times nm}$  whose blocks are given by

$$N_{ij} = \begin{cases} \frac{1}{\sqrt{d_i d_j}} X_{ij} & (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

where  $X_{ij}, (i, j) \in \mathcal{E}$  are independent random matrices that satisfy

$$\mathbb{E}[X_{ij}] = 0, \quad \|X_{ij}\| \leq K_{ij},$$

where  $K_{ij}, (i, j) \in \mathcal{E}$  are positive constants. Let  $K = \max_{(i,j) \in \mathcal{E}} K_{ij}$ , then

**Lemma 3.1.** *With probability at least  $1 - \frac{m}{n^{2c}}$ ,*

$$\|N\| \leq \frac{CK}{\sqrt{d_{\min}}}$$

for some absolute constant  $C$ .

**Lemma 3.2.**

*Given  $l$  positive integer  $t_1, \dots, t_l$ , and suppose  $l \leq \frac{\log n}{\log \log n}$ ,  $d_{\min} > \log n$ , define a vector  $\omega$  as*

$$\omega_i = \frac{1}{\sqrt{d_i}} \sum_{j \in \mathcal{N}(i)} \frac{1}{\sqrt{d_j}},$$

and setting a constant  $\|B\|_{\omega}$  associated to  $B$  and  $\omega$  as

$$\|B\|_{\omega} := \|abs(B) \cdot \omega\|_{\infty}$$

in which  $abs(B)$  is obtained by taking absolute values elementwisely from  $B$ , then we have

$$\begin{aligned} \|(e_i^T \otimes I_m) \prod_{j=1}^l ((B^{t_j} \otimes I_m) N) S_k\|_{\infty} &\leq \left( \sqrt{\frac{\log n}{d_{\min}}} \cdot CK \right)^l \cdot \|B\|_{\omega} \cdot \|B\|_{\infty}^{t-l/2-1} \cdot \|B\|_{\max}^{l/2} \cdot \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \|U_k\| \\ &\leq \left( \sqrt{\frac{\log n}{d_{\min}}} CK \right)^l \|B\|_{\infty}^{t-l/2} \|B\|_{\max}^{l/2} \sqrt{\frac{d_{\max}}{|\mathcal{E}|}} \|U_k\| \end{aligned} \quad (31)$$

with probability exceeding  $1 - n^{-1/l}$ . Recall that in our random model we have  $S_k = s \otimes U_k$ . Note that we put no special assumption on  $B$  here. Namely,  $B$  could be different from  $B_i$  defined in Section (2.1).

**Lemma 3.3.** *Given noise matrix  $N$  under the model in (3.1), and*

$$K \leq \frac{1}{C} \sqrt{\frac{d_{\min}}{\|B_i\| \cdot \|B_i\|_{\infty} \cdot \log n}},$$

then we have the block-wise bound on  $Y_i$

$$\max_{1 \leq b \leq n} \|(e_b^T \otimes I_m) Y_i\| \leq \frac{\|B_i\|_{\omega}}{\|B_i\|_{\infty}} \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \leq \sqrt{\frac{d_{\max}}{|\mathcal{E}|}}$$

with probability exceeding  $1 - \frac{1}{n^c}$  for some absolute constant  $C$  depending on  $c$ . Recall that  $B_i$  and  $Y_i$  were given by

$$B_i := \left( \bar{S}_k (\lambda_i I - \bar{\Lambda})^{-1} \bar{S}_k^T \right) \otimes I_m,$$

$$Y_i := (\mathbf{y}_{(i-1)m+1}, \dots, \mathbf{y}_{im}) = \sum_{l=0}^{\infty} ((I - B_i N)^{-1} B_i)^{l+1} N(S_k \otimes I_m) X_i (\lambda_i I_m - \Sigma_i)^l,$$

in Section (2.1) respectively, where  $1 \leq i \leq k$ .

### 3.2 Proof of Lemma 3.1

We provide a bound for  $\|N\|$  by giving an estimation of  $\mathbb{E}[\text{tr}(N^b)]$  for even positive integers  $b$ .

**Proposition 4.** *Given two positive integer  $b, w$  such that  $1 \leq w \leq \frac{b}{2} + 1$ , let  $G$  be a undirected complete graph of  $w$  vertices then the number of cycles of length  $b$  in  $G$  satisfying the following properties can be bounded above by  $C^b b^{b/2}$  (or equivalently  $C' b^{b/2+1-w} w!$  for some slightly different absolute constant  $C'$ ) for some absolute constant  $C$ :*

- No self-loop in the cycle;
- Each vertex of  $G$  appears;
- Each edge appearing in the cycle would appear at least two times.

*Proof.* Define the following variables associated to  $w$  vertices:

- The degrees (not consider the multiplicity of edge) of  $w$  vertices  $d_1, \dots, d_w$  in the cycle, which is actually the degree of some vertex in the induced undirected graph by the cycle;
- The multi-degrees of  $w$  vertices  $D_1, \dots, D_w$  in the cycle, which is actually the number of times some vertex appears in the cycle sequence.

We say a leg in the cycle is innovative if the undirected edge of that leg didn't arose before when traversing the cycle, and non-innovative if otherwise. It is clear that

$$d_1 + \dots + d_w = 2j, \quad D_1 + \dots + D_w = b, \quad d_t \leq D_t.$$

Thus the number of the possibilities of  $\{D_t\}$  could be bounded by  $2^b$ , so it suffices to consider only fixed  $\{D_t\}$ . Define  $T_p$  such that  $T_p/p$  is the number of vertices with multi-degree  $p$ , or equivalently, the number of vertices in the cycle having multi-degree  $p$ . It is clear that  $\sum_{p=1}^{\infty} T_p = b$ .

We try to show that it is possible to reconstruct the cycle by recording a group of arrays. In detail, supposing some valid cycle is  $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_b \rightarrow i_1$ , define an array  $\alpha$  of length  $b$  such that  $\alpha_t := 1$  if  $i_t i_{t+1}$  ( $i_{b+1} = i_1$  here) is innovative while  $\alpha_t := 0$  if otherwise, then the following arrays would be recorded:

- A non-negative integer array  $\mathbf{c}$  of length  $w$  such that  $q = c_1 + \dots + c_w \leq b$  (to be modified);
- A bit array  $\mathbf{v} = v_1 \dots v_b$  of length  $b$ , in which there are exactly  $q$  ones and  $b - q$  zeros.  $v_t$  will be forced to be 1 if  $\alpha_t = 1$ .
- Non-negative integer  $E$ , which is in fact an encoding of the cycle using some method we would show in the following.

We define  $E$  by

$$E = L + L_{max} M$$

in which  $L_{max} := \binom{q}{c_1, \dots, c_w}$ , and  $L$  and  $M$  are non-integer numbers which will be defined next.

To define  $L$ , an integer array  $\mathbf{u}$  would be introduced which is of length  $q$  with  $t$  appearing exactly  $c_t$  times for  $1 \leq t \leq w$ .  $L$  is simply the encoding of  $\mathbf{u}$ , thus it could be bounded by

$$0 \leq L < L_{max} := \binom{q}{c_1, \dots, c_w} \leq C_1^b \frac{q^q}{c_1^{c_1} \dots c_w^{c_w}}$$

for some absolute constant  $C_1$  from the Stirling's formula and the condition  $q \leq b$ .

Now we would describe how to calculate the number  $M$  from the known cycle and all other information stored. First decode  $\mathbf{u}$  from  $L$  and  $c_1, \dots, c_w$ . Let  $m_1 = i_1 - 1$ ,  $m_{1,max} = w$ . Consider  $1 \leq t \leq b-1$  and suppose  $i_1 \rightarrow \dots \rightarrow i_t$  is known. If  $v_t = 1$  is the  $l$ -th one in  $\mathbf{v}$  ( $l \leq q$ ),  $u_l$  will be used to indicated the multi-degree of vertex  $i_{t+1}$ , thus a integer  $m_{t+1}$  such that  $0 \leq m_{t+1} < T_{u_l}/u_l$  can uniquely determine  $i_{t+1}$ . In this case let  $m_{t+1,max} = T_{u_l}/u_l$ . If  $v_t = 0$ , from the definition of  $\mathbf{v}$  the leg  $i_t i_{t+1}$  must be a non-innovative leg. Thus an integer  $m_{t+1}$  such that  $0 \leq m_{t+1} < d_{i_t} \leq D_{i_t}$  can uniquely determine  $i_{t+1}$ . In this case let  $m_{t+1,max} = D_{i_t}$ . In the end, encode  $M$  as

$$M = m_1 + m_2 m_{1,max} + \dots + m_b m_{b-1,max} \dots m_{1,max}.$$

From the construction of  $M$ , clearly we can reconstruct the original cycle from  $M$  together with all other stored information  $\mathbf{D}, \mathbf{v}, \mathbf{c}, L$ . On the contrary, given one cycle, for each valid  $\mathbf{v}$  we can find such a (unique)  $M(\mathbf{v})$  and  $E(\mathbf{v})$  depending on  $\mathbf{v}$  since  $\mathbf{u}$  can be completely determined by  $\mathbf{v}$  and the cycle.

Here we would construct a special  $\mathbf{v}$  and show that  $E(\mathbf{v})$  is small enough to carry our proof forward. Let  $\mathbf{v}$  be all blanks except on the position  $v_t$  such that  $\alpha_t = 1$  ( $v_t$  will be forced to set as 1 in such case). As for all other non-innovative legs  $i_t i_{t+1}$ , we set  $v_t := 1$  if  $b/D_{i_{t+1}} < D_{i_t}$  while  $v_t := 0$  if otherwise. Clearly

$$M < M_{max} := m_{1,max} \dots m_{b,max}.$$

For each degree value  $p$ , the term  $T_p/p$  will appear  $c_p$  times in the sequence  $m_{t,max}$  from the definition, and further more we denote by  $f_p$  the number of times that the degree  $p$  appears in the sequence  $m_{t,max}$ . By such a construction, we have

$$\begin{aligned} E(\mathbf{v}) &< L_{max} M_{max} \\ &\leq C_1^b \frac{q^q}{c_1^{c_1} \dots c_w^{c_w}} \times w \prod_{\substack{i_t i_{t+1} \\ \text{innovative}}} \frac{T_{D_{i_{t+1}}}}{D_{i_{t+1}}} \prod_{\substack{i_t i_{t+1} \\ \text{non-innovative}}} \min\{D_{i_t}, b/D_{i_{t+1}}\} \\ &= C_1^b \frac{q^q}{c_1^{c_1} \dots c_w^{c_w}} \times w \left( \prod_{p=1}^w (T_p/p)^{c_p} \right) \cdot \left( \prod_{p=1}^w p^{f_p} \right) \\ &= C_2^b \left( \prod_{p=1}^w (q/p)^{c_p} \right) \cdot \left( \prod_{p=1}^w (T_p/c_p)^{c_p} \right) \cdot \left( \prod_{p=1}^w p^{f_p} \right) \\ &\leq C_3^b \left( \prod_{p=1}^w (b/p)^{c_p} \right) \cdot \left( \prod_{p=1}^w p^{f_p} \right) \tag{32} \\ &= C_3^b \prod_{\substack{i_t i_{t+1} \\ \text{innovative}}} \frac{b}{D_{i_{t+1}}} \prod_{\substack{i_t i_{t+1} \\ \text{non-innovative}}} \min\{D_{i_t}, b/D_{i_{t+1}}\} \\ &= C_3^b \prod_{t=1}^b \sqrt{\frac{b D_{i_t}}{D_{i_{t+1}}}} \prod_{\substack{i_t i_{t+1} \\ \text{innovative}}} \sqrt{\frac{b}{D_{i_t} D_{i_{t+1}}}} \prod_{\substack{i_t i_{t+1} \\ \text{non-innovative}}} \min \left\{ \sqrt{\frac{b}{D_{i_t} D_{i_{t+1}}}}, \sqrt{\frac{D_{i_t} D_{i_{t+1}}}{b}} \right\} \\ &\leq C_3^b b^{b/2} \prod_{\substack{i_t i_{t+1} \\ \text{innovative}}} \sqrt{\frac{b}{D_{i_t} D_{i_{t+1}}}} \prod_{\substack{i_t i_{t+1} \\ \text{non-innovative}}} \min \left\{ \sqrt{\frac{b}{D_{i_t} D_{i_{t+1}}}}, \sqrt{\frac{D_{i_t} D_{i_{t+1}}}{b}} \right\} \tag{33} \\ &\leq C_3^b b^{b/2} \end{aligned}$$

for some slightly different absolute constants  $C_2, C_3$ , in which (32) comes from the fact that  $(T_p/c_p)^{c_p} \leq e^{T_p/e}$  and  $T_1 + \dots + T_w = q \leq b$ . To explain the last step in the transformation above, consider each edge  $\{r, s\}$  in the cycle. It have been given that  $\{r, s\}$  will appear at least 2 times in the cycle, supposing it appears exact  $k \geq 2$  times, then it will contribute a factor

$$\sqrt{\frac{b}{D_r D_s}} \min \left\{ \sqrt{\frac{b}{D_r D_s}}, \sqrt{\frac{D_r D_s}{b}} \right\}^{k-1} \leq 1$$

to the formula (33).

In this way, we can always encode the cycle into a tuple  $(\mathbf{D}, \mathbf{v}, \mathbf{c}, E)$  for some  $\mathbf{v}$  such that  $E < C_3^b b^{b/2}$  for some absolute constant  $C_3$ . The total number of such tuples can be clearly bounded by  $C^b b^{b/2}$  for some slightly different absolute constant  $C$ .

In the end, by Stirling's formula we can write the ratio between  $b^{b/2}$  and  $b^{b/2+1-w} w!$  as

$$\frac{b^{b/2}}{b^{b/2+1-w} w!} \leq C_4^b \left( \frac{b}{w} \right)^w$$

for some absolute constant  $C_4$ . However simple calculus gives that  $(b/w)^w \leq e^{w/e} \leq e^{b/e}$ . Hence the number of valid cycles can be also bounded by  $C' b^{b/2+1-w} w!$  for some slightly different constant  $C'$ .  $\square$

**Proposition 5.** *Suppose  $S(n)$  is the set consisting of all permutations on  $\{1, \dots, n\}$ . Let  $(i_1, i_2), \dots, (i_b, i_1)$  be a cycle of length  $b$  in the graph. Let  $e_1, \dots, e_j$  are the  $j$  distinct undirected edges appearing in the cycle, with occurrence times  $\alpha_1, \dots, \alpha_j$ , and  $f_1, \dots, f_w$  are the  $w$  distinct vertices appearing in the cycle, with occurrence times  $\beta_1, \dots, \beta_w$ . Suppose all  $\alpha$  are at least 2, then we have*

$$\mathbb{E}_{\sigma \in U_{S(n)}} [N_{\sigma(i_1)\sigma(i_2)} \dots N_{\sigma(i_b)\sigma(i_1)}] \leq \frac{K^{b-2j}}{d_{\min}^{b-w+1} (n-1) \dots (n-w+1)}$$

*Proof.* Define  $G$  as a graph such that  $(s, t)$  is an edge of  $G$  if and only if  $(f_s, f_t)$  appears in the cycle  $(i_1, \dots, i_b, i_1)$ . We have

$$\begin{aligned} & \mathbb{E}_{\sigma \in U_{S(n)}} [N_{\sigma(i_1)\sigma(i_2)} \dots N_{\sigma(i_b)\sigma(i_1)}] \\ & \leq \mathbb{E}_{\sigma \in U_{S(n)}} [|N_{\sigma(i_1)\sigma(i_2)} \dots N_{\sigma(i_b)\sigma(i_1)}|] \\ & \leq \mathbb{E}_{\sigma \in U_{S(n)}, X} \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} |X_{\sigma(e_1)}|^{\alpha_1} \dots |X_{\sigma(e_j)}|^{\alpha_j} \right] \\ & \leq K^{\alpha_1 + \dots + \alpha_j - 2j} \mathbb{E}_{\sigma \in U_{S(n)}} \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} \mathbb{E}_X [X_{\sigma(e_1)}^2 \dots X_{\sigma(e_j)}^2] \right] \\ & \leq K^{\alpha_1 + \dots + \alpha_j - 2j} \mathbb{E}_{\sigma \in U_{S(n)}} \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} \delta(\sigma; 1, \dots, w) \right], \end{aligned}$$

in which  $\delta(\sigma; p_1, \dots, p_l) := 1$  if corresponding  $X_{\sigma(f_{p_s}), \sigma(f_{p_t})}$  are non-zero for all  $(p_s, p_t) \in G$ , and  $\delta(\sigma; p_1, \dots, p_l) := 0$  otherwise. It is easy to see that if  $\delta(\sigma; P) = 1$  for some  $\sigma$  then all subset  $P' \subseteq P$  also satisfy  $\delta(\sigma; P') = 1$ . It suffices to bound  $\mathbb{E}_{\sigma \in U_{S(n)}} \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} \delta(\sigma; 1, \dots, w) \right]$ . To use induction to achieve this end, we relax our assumption such that now we have not a cycle but just a connected graph consisting of  $w$  vertices  $f_1, \dots, f_w$  with degrees  $\beta_1, \dots, \beta_w$ . We can find a vertex such that the graph is still connected after removing this vertex. Without loss of generality, suppose it is  $f_w$ , and one of its adjacent vertices is  $f_{w-1}$ . Fixing  $\sigma(f_{w-1})$ , there are at most  $d_{\sigma(f_{w-1})}$  choices for  $\sigma(f_w)$  to make  $\delta(\sigma; 1, \dots, w)$  not vanish. Thus we have

$$\begin{aligned} & \mathbb{E}_{\sigma} \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} \delta(\sigma; 1, \dots, w) \right] \\ & \leq \mathbb{E}_{\sigma(f_1), \dots, \sigma(f_{w-1})} \left[ d_{\sigma(f_1)}^{-\beta_1} \dots d_{\sigma(f_{w-1})}^{-\beta_{w-1}} \delta(\sigma; 1, \dots, w-1) \mathbb{E}_{\sigma(f_w)} \left[ d_{\sigma(f_w)}^{-\beta_w} \delta(\sigma; w-1, w) \right] \right] \\ & \leq d_{\min}^{-\beta_w} \mathbb{E}_{\sigma(f_1), \dots, \sigma(f_{w-1})} \left[ \frac{d_{\sigma(f_{w-1})}}{n-w+1} d_{\sigma(f_1)}^{-\beta_1} \dots d_{\sigma(f_{w-1})}^{-\beta_{w-1}} \delta(\sigma; 1, \dots, w-1) \right] \\ & \leq \frac{1}{d_{\min}^{\beta_r} (n-w+1)} \mathbb{E}_{\sigma} \left[ d_{\sigma(f_1)}^{-\beta_1} \dots d_{\sigma(f_{w-2})}^{-\beta_{w-2}} d_{\sigma(f_{w-1})}^{-\beta_{w-1}+1} \delta(\sigma; 1, \dots, w-1) \right]. \end{aligned}$$

Observe that the last line has the form of recursion in the sense that we now have a connected graph consisting of  $f_1, \dots, f_{w-1}$  with degrees  $\beta_1, \dots, \beta_{w-2}, \beta_{w-1} + 1$ , hence we obtain that

$$\begin{aligned} \mathbb{E}_\sigma \left[ \frac{1}{d_{\sigma(f_1)}^{\beta_1} \dots d_{\sigma(f_w)}^{\beta_w}} \delta(\sigma; 1, \dots, w) \right] &\leq \frac{1}{d_{\min}^{\beta_1 + \dots + \beta_w - w + 1} (n-1) \dots (n-w+1)} \\ &= \frac{1}{d_{\min}^{b-w+1} (n-1) \dots (n-w+1)} \end{aligned}$$

□

**Proposition 6.**

$$\mathbb{E}[\text{tr}(N^b)] \leq C^b \frac{n}{d_{\min}^{b/2}}$$

for some absolute constant  $C$ .

*Proof.* We have

$$\mathbb{E}[\text{tr}(N^b)] = \sum_{1 \leq i_1, \dots, i_b \leq n} \mathbb{E}[N_{i_1 i_2} \dots N_{i_b i_1}], \quad (34)$$

which is a sum over all cycles  $i_1 \rightarrow \dots \rightarrow i_b \rightarrow i_1$  of length  $b$ . According to the independence of  $N_{ij}$ , all terms are zero but those terms in which every  $\{i_s, i_{s+1}\}$  appears at least two times. Consider cycles with  $w$  distinct vertices and  $j$  distinct edges appearing. Let  $\mathcal{V}(b, w, j)$  be the set of all such cycles. Using Proposition 4,  $|\mathcal{V}(b, w, j)|$  will be bounded by

$$C_1^b b^{b/2+1-w} w! \times \binom{n}{w} = C_1^b b^{b/2+1-w} n(n-1) \dots (n-w+1)$$

where  $C_1$  is an absolute constant. While Proposition 5 gives the expected contribution to sum (34) over cycles of the certain shape, which means

$$\begin{aligned} &\sum_{(i_1, \dots, i_b) \in \mathcal{V}(b, w, j)} \mathbb{E}[N_{i_1 i_2} \dots N_{i_b i_1}] \\ &= \sum_{\sigma \in U_{S(n)}} \mathbb{E} \left[ \sum_{(i_1, \dots, i_b) \in \mathcal{V}(b, w, j)} \mathbb{E}[N_{\sigma(i_1)\sigma(i_2)} \dots N_{\sigma(i_b)\sigma(i_1)}] \right] \\ &= \sum_{(i_1, \dots, i_b) \in \mathcal{V}(b, w, j)} \sum_{\sigma \in U_{S(n)}} \mathbb{E}[N_{\sigma(i_1)\sigma(i_2)} \dots N_{\sigma(i_b)\sigma(i_1)}] \\ &\leq C_1^b b^{b/2+1-w} n(n-1) \dots (n-w+1) \times \frac{K^{b-2j}}{d_{\min}^{b-w+1} (n-1) \dots (n-w+1)} \\ &\leq C_1^b n d_{\min}^{-b/2} \times \left( \frac{K^2 b}{d_{\min}} \right)^{b/2-w+1} \\ &\leq C^b \frac{n}{d_{\min}^{b/2}} \end{aligned}$$

for some absolute constant  $C$  in which we used the fact that  $j \geq w-1$  and  $b = O(\log n)$ ,  $K = O\left(\sqrt{\frac{d_{\min}}{\log n}}\right)$ .

**Theorem 3.1.** Suppose  $K = O\left(\sqrt{\frac{d_{\min}}{\log n}}\right)$ , then

$$\|N\| \leq \frac{C}{\sqrt{d_{\min}}}$$

holds with high probability for some absolute constant  $C$ .



*Proof.* Setting  $b = 4\lceil \log n \rceil$  which is an even number, we write the constant in Proposition (6) as  $C_1$  here, then the Proposition (6) together with the Markov's inequality gives

$$\begin{aligned} \Pr \left[ \|N\| > \frac{3C_1}{\sqrt{d_{\min}}} \right] &= \Pr \left[ \|N\|^b > \frac{3^b C_1^b}{d_{\min}^{b/2}} \right] \\ &\leq \frac{n}{3^b} < \frac{1}{n^4}. \end{aligned}$$

Hence  $\|N\| \leq \frac{C}{\sqrt{d_{\min}}}$  with high probability in which  $C = 3C_1$  is an absolute constant.  $\square$

### 3.3 Proof of Lemma 3.2

*Proof.* First we have

$$\|B\|_{\omega} = \|abs(B) \cdot \omega\|_{\infty} \leq \|\omega\|_{\infty} \|B\|_{\infty} \leq \sqrt{\frac{d_{\max}}{d_{\min}}} \|B\|_{\infty}$$

Hence it suffices to prove (31) by showing that

$$\|(\mathbf{e}_1^T \otimes I_m) \prod_{j=1}^l ((B^{t_j} \otimes I_m) N) S_k\|_{\infty} \leq \left( \sqrt{\frac{\log n}{d_{\min}}} \cdot CK \right)^l \cdot \|B\|_{\omega} \cdot \|B\|_{\infty}^{t-l/2-1} \cdot \|B\|_{\max}^{l/2} \cdot \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \|U_k\| \quad (35)$$

holds with probability exceeding  $1 - \frac{1}{n^c}$ . Denote by  $a_{pq}^{(j)}$  the element of matrix  $B^{t_j}$  in  $p$ -th row and  $q$ -th column. In this way, we can expand the left side of (35) as

$$\left\| \sum_{\alpha, \beta} a_{1\alpha_1}^{(1)} N_{\alpha_1\beta_1} a_{\beta_1\alpha_2}^{(2)} \dots a_{\beta_{l-1}\alpha_l}^{(l)} N_{\alpha_l\beta_l} (s_{\beta_l} U_k) \right\| \quad (36)$$

in which we used the fact that  $S_k = \mathbf{s} \otimes U_k$ . Here the summation is taken over all possible integer arrays  $\alpha$  and  $\beta$  which are both of length  $l$ . Recall that  $s_{\beta_l} = \sqrt{\frac{d_{\beta_l}}{|\mathcal{E}|}}$  and  $N_{p,q} = \frac{1}{\sqrt{d_p d_q}} X_{p,q}$  if  $(p, q) \in \mathcal{E}$  while  $N_{p,q} = 0$  otherwise. We thereby denote by  $\mathcal{H}$  as the set consisting of all pairs  $(\alpha, \beta)$  such that  $(\alpha_j, \beta_j) \in \mathcal{E}$  for all  $j$ . Hence it suffices to prove that

$$\begin{aligned} &\frac{1}{\sqrt{|\mathcal{E}|}} \left\| \sum_{(\alpha, \beta) \in \mathcal{H}} \frac{a_{1,\alpha_1}^{(1)} a_{\beta_1,\alpha_2}^{(2)} \dots a_{\beta_{l-1},\alpha_l}^{(l)}}{\sqrt{d_{\alpha_1} d_{\beta_1} \dots d_{\alpha_{l-1}} d_{\beta_{l-1}} d_{\alpha_l}}} \cdot X_{\alpha_1\beta_1} \dots X_{\alpha_l\beta_l} U_k \right\| \\ &\leq \left( \sqrt{\frac{\log n}{d_{\min}}} \cdot CK \right)^l \cdot \|B\|_{\omega} \cdot \|B\|_{\infty}^{t-l/2-1} \cdot \|B\|_{\max}^{l/2} \cdot \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \|U_k\|, \end{aligned} \quad (37)$$

or in a stronger sense,

$$\begin{aligned} &\left\| \sum_{(\alpha, \beta) \in \mathcal{H}} \frac{a_{1,\alpha_1}^{(1)} a_{\beta_1,\alpha_2}^{(2)} \dots a_{\beta_{l-1},\alpha_l}^{(l)}}{\sqrt{d_{\alpha_1} d_{\beta_1} \dots d_{\alpha_{l-1}} d_{\beta_{l-1}} d_{\alpha_l}}} \cdot X_{\alpha_1\beta_1} \dots X_{\alpha_l\beta_l} \right\| \\ &\leq C^l (\log n)^{l/2} \cdot \|B\|_{\omega} \cdot \|B\|_{\infty}^{t-l/2-1} \cdot \|B\|_{\max}^{l/2}. \end{aligned} \quad (38)$$

since  $\|X_{pq}\| \leq K$  for all  $p, q$ . To this end, we employ the power moment method, which needs us to show

$$\begin{aligned} &\mathbb{E} \left[ \left\| \left( \sum_{(\alpha, \beta) \in \mathcal{H}} \prod_{r=1}^{2k} \frac{a_{1,\alpha_1}^{(1)} a_{\beta_1,\alpha_2}^{(2)} \dots a_{\beta_{l-1},\alpha_l}^{(l)}}{\sqrt{d_{\alpha_1} d_{\beta_1} \dots d_{\alpha_{l-1}} d_{\beta_{l-1}} d_{\alpha_l}}} \cdot X_{\alpha_1^{(r)}\beta_1^{(r)}} \dots X_{\alpha_l^{(r)}\beta_l^{(r)}} \right)^{2k} \right\| \right] \\ &\leq (\log n)^{kl} \cdot \|B\|_{\omega}^{2k} \cdot \|B\|_{\infty}^{2kt-kl-2k} \cdot \|B\|_{\max}^{kl}. \end{aligned} \quad (39)$$

holds for all integer  $k$  such that  $kl \leq O(\log n)$ . The left side of (39) can be furthermore relaxed and expanded to

$$\sum_{(\hat{\alpha}, \hat{\beta}) \in \mathcal{H}^{2k}} \prod_{r=1}^{2k} \frac{|a_{1, \alpha_1^{(r)}}^{(1)} a_{\beta_1^{(r)}, \alpha_2^{(r)}}^{(2)} \dots a_{\beta_{l-1}^{(r)}, \alpha_l^{(r)}}^{(l)}|}{\sqrt{d_{\alpha_1^{(r)}} d_{\beta_1^{(r)}} \dots d_{\alpha_{l-1}^{(r)}} d_{\beta_{l-1}^{(r)}} d_{\alpha_l^{(r)}}}} \cdot \left| \mathbb{E} \left[ \prod_{r=1}^{2k} X_{\alpha_1^{(r)} \beta_1^{(r)}} \dots X_{\alpha_l^{(r)} \beta_l^{(r)}} \right] \right| \quad (40)$$

where  $\hat{\alpha} = (\alpha^{(1)}, \dots, \alpha^{(2k)})$ ,  $\hat{\beta} = (\beta^{(1)}, \dots, \beta^{(2k)})$ . Here  $\mathcal{H}^{2k}$  is defined such that  $(\hat{\alpha}, \hat{\beta}) \in \mathcal{H}^{2k}$  if and only if  $(\alpha^{(r)}, \beta^{(r)}) \in \mathcal{H}$ . An important observation is that  $\mathbb{E}[X_{\alpha_1 \beta_1} \dots X_{\alpha_l \beta_l}]$  is non-zero if and only if all unordered pairs  $\{\alpha_j, \beta_j\}$  will appear at least two times among these  $2k$  pairs since all  $X_{pq}$  are independent for distinct  $\{p, q\}$  and  $\mathbb{E}[X_{pq}] = 0$ . Let  $\mathcal{J}_{2k} \subseteq \mathcal{H}^{2k}$  be the set consisting of all  $(\hat{\alpha}, \hat{\beta})$  such that every  $(\alpha_j^{(r)}, \beta_j^{(r)})$  would appear at least two times among these  $2kl$  unordered pairs, which means it suffices to prove that

$$\sum_{(\hat{\alpha}, \hat{\beta}) \in \mathcal{J}_{2k}} \prod_{r=1}^{2k} \frac{|a_{1, \alpha_1^{(r)}}^{(1)} a_{\beta_1^{(r)}, \alpha_2^{(r)}}^{(2)} \dots a_{\beta_{l-1}^{(r)}, \alpha_l^{(r)}}^{(l)}|}{\sqrt{d_{\alpha_1^{(r)}} d_{\beta_1^{(r)}} \dots d_{\alpha_{l-1}^{(r)}} d_{\beta_{l-1}^{(r)}} d_{\alpha_l^{(r)}}}} \leq (\log n)^{kl} \cdot \|B\|_{\omega} \cdot \|B\|_{\infty}^{2kt - kl - 2k} \cdot \|B\|_{\max}^{2kl}. \quad (41)$$

To estimate the sum in the left side of (41), we introduce the concept of summing graph.

**Definition 1.** A summing graph  $\mathcal{G}_s = (\mathcal{V}_s, \mathcal{E}_1, \mathcal{E}_2)$  is formally defined as following:

1.  $\mathcal{V}_s$  contains  $1 + n$  nodes, in which one node is special and called the root node while the remaining  $n$  nodes are used to represent  $n$  variables that we are summing over. In fact, the root node represents the fixed index 1 in  $a_{1, \alpha_1^{(r)}}^{(1)}$  in (41).
2.  $\mathcal{E}_1$  is a undirected edge set on vertex set  $\mathcal{V}_s$ . If  $(u, v) \in \mathcal{E}_1$ , then the variables represented by  $u, v$  are adjacent in  $\mathcal{G}$ .
3.  $\mathcal{E}_2$  is an undirected labeled edge set on vertex set  $\mathcal{V}_s$  and  $\mathcal{E}_2$  can contain multiple edges. Each edge in  $\mathcal{E}_2$  has the form  $(p, q, n)$  where  $p, q$  is the end node while  $n$  is an integer label.

A summing graph  $\mathcal{G}_s$  combined with a vector  $\theta$  can induce a sum  $\Sigma(\mathcal{G}_s, \theta)$  as defined below:

$$\Sigma(\mathcal{G}_s, \theta) = \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_s)} \prod_{(p, q, n) \in \mathcal{E}_2} [B^n]_{v_p, v_q} / F(\mathbf{v}, \theta)$$

in which we define

$$\mathcal{U}(\mathcal{G}_s) = \{\mathbf{v} \in \{1, \dots, n\}^{|\mathcal{V}_s|} : \text{for every } (p, q) \in \mathcal{E}_1, v_p, v_q \text{ are adjacent in } \mathcal{G}\},$$

$$F(\mathcal{G}_s, \mathbf{v}, \theta) = \prod_{p \in \mathcal{V}_s} \frac{1}{d_{v_i}^{\theta_p/2}} \prod_{p \notin \mathcal{V}_s} \frac{1}{d_{\min}^{\theta_p/2}}.$$

Return to the original problem. For some  $(\hat{\alpha}, \hat{\beta}) \in \mathcal{J}_{2k}$ , we can divide all  $2kl$  terms of form  $\{\alpha_j^{(r)}, \beta_j^{(r)}\}$  into a number of groups such that each group includes the same unordered pairs. By the definition of  $\mathcal{J}_{2k}$  each group has a size of at least 2. Without loss of generality we could assume all groups are of size 2 or 3. For example, if some  $\{p, q\}$  occurs 7 times among  $2kl$  unordered pairs, we can divide these 7 terms into 3 groups of size 2, 2, 3 respectively. Let the number of groups with size 2 be  $kl - 3u$ , then the number of groups with size 3 would be  $2u$  since there are  $2kl$  terms in total. To obtain identical relations between ordered pairs  $(\alpha_j^{(r)}, \beta_j^{(r)})$ , we can use a bit string of length  $2kl$  to indicate whether the corresponding  $(\alpha_j^{(r)}, \beta_j^{(r)})$  has the same order as its first appearance. Fixing  $u$ , the number of possible groupings would be

$$\frac{(2kl)!}{2^{kl-3u} 6^{2u} (kl-3u)!(2u)!} = \frac{(2kl)!}{2^{kl} 4.5^u (kl-3u)!(2u)!}.$$

For some grouping configuration with  $kl - u$  groups together with a  $2kl$  bit string like mentioned above, a summing group  $\mathcal{G}_1$  could be constructed as below:

- $\mathcal{V}(\mathcal{G}_1)$  has  $2(kl-u)+1$  nodes. In particular, one of them represents 1, which is the root node, while each of others represents exactly one element in  $kl-u$  groups. More precisely, each node can be regarded as a set of variables of form  $\alpha_i^{(r)}$  or  $\beta_i^{(r)}$  so that all variables in the same node are forced to have the same value.
- The edge joining nodes representing  $\alpha_i^{(r)}, \beta_i^{(r)}$  would be in  $\mathcal{E}_1(\mathcal{G}_1)$  for all possible  $i, r$ . Thus each node except the root node is associated with exact one edge in  $\mathcal{E}_1(\mathcal{G})$ .
- Starting from an empty  $\mathcal{E}_2(\mathcal{G}_1)$ , add a labeled edge  $(\overline{\beta_i^{(r)}}, \overline{\alpha_{i+1}^{(r)}}, i+1)$  for all possible  $\beta_i^{(r)}, \alpha_{i+1}^{(r)}$ , and add a labeled edge  $(\mathbf{root}, \alpha_1^{(r)}, 1)$  for all possible  $\alpha_1^{(r)}$ . Recall that  $\mathcal{E}_2(\mathcal{G}_1)$  is a unordered edge set containing multiple edges.
- $\theta_1$  is defined as an integer vector of the same length as the number of nodes in  $\mathcal{G}_1$ . Let  $\theta_{1,0}$  be 0 for root node 0, and  $\theta_{1,p}$  be the number of times that variable represented by node  $p$  appears in  $\alpha_i^{(r)}$  and  $\beta_i^{(r)}$  except  $\beta_l^{(r)}$ . It is clear that  $\theta_{1,p} \geq 2$  for node  $p$  where  $p$  is not the root node and does not contain variable of form  $\beta_l^{(r)}$ . Hence there are at most  $2k$  indices  $p$  such that  $\theta_p = 1$ .

In this way, we can easily verify that

$$\sum_{(\alpha, \beta) \in \mathcal{J}_{2k}} \prod_{r=1}^{2k} \frac{|a_{1, \alpha_1^{(r)}}^{(1)} a_{\beta_1^{(r)}, \alpha_2^{(r)}}^{(2)} \cdots a_{\beta_{l-1}^{(r)}, \alpha_l^{(r)}}^{(l)}|}{\sqrt{d_{\alpha_1^{(r)}} d_{\beta_1^{(r)}} \cdots d_{\alpha_{l-1}^{(r)}} d_{\beta_{l-1}^{(r)}} d_{\alpha_l^{(r)}}}} \leq \sum_{\mathcal{G}_1} \Sigma(\mathcal{G}_1, \theta_1)$$

in which  $\mathcal{G}_1$  is taken over all possible configurations of grouping.

Next, we will provide an estimation on  $\Sigma(\mathcal{G}_1, \theta_1)$  by induction where  $\mathcal{G}_1$  contains  $2(kl-u)+1$  nodes.

Claim:

$$\Sigma(\mathcal{G}_1, \theta_1) \leq \|B\|_{\omega}^{2k} \cdot \|B\|_{\infty}^{2kt-kl-2k} \cdot \|B\|_{\max}^{kl} \cdot d_{\min}^{-(l-1)k-u}.$$

**Proposition 7.** *We can remove some edges from  $\mathcal{E}_2(\mathcal{G}_1)$  to obtain a graph  $\mathcal{G}_2$  satisfying*

- $\mathcal{G}_2$  is a tree with edge set  $\mathcal{E}_1(\mathcal{G}_2) \cup \mathcal{E}_2(\mathcal{G}_2)$ .
- $\mathcal{E}_2(\mathcal{G}_2)$  does not contain multiple edges any more.
- Except the root node, each node in  $\mathcal{G}_2$  is associated with exactly one edge in  $\mathcal{E}_1$ .

Define the difference between two graphs  $\mathcal{G}$  and  $\mathcal{G}'$  as

$$\Delta(\mathcal{G}', \mathcal{G}) := \sum_{(p, q, n) \in \mathcal{E}_2(\mathcal{G}') \setminus \mathcal{E}_2(\mathcal{G})} n$$

for graph  $\mathcal{G} \subseteq \mathcal{G}'$ , then we have

$$\Sigma(\mathcal{G}_1, \theta_1) \leq \|B\|_{\max}^{\Delta(\mathcal{G}_1, \mathcal{G}_2)} \Sigma(\mathcal{G}_2, \theta_2).$$

Since there are  $2kl$  edges in  $\mathcal{E}_2(\mathcal{G}_1)$  but only  $kl-u$  edges in  $\mathcal{E}_2(\mathcal{G}_2)$ , we have

$$\Delta(\mathcal{G}_1, \mathcal{G}_2) \geq kl+u.$$

In the following, we will construct a sequence of summing graphs  $\mathcal{G}_2, \dots, \mathcal{G}_m$  in a inductive way, where  $\mathcal{G}_m$  is a summing graph with only one node, the root node. Specifically, we impose an inductive assumption which holds for  $\mathcal{G}_1$  and  $\mathcal{G}_2$ :  $\mathcal{G}_z$  is a tree over edge set  $\mathcal{E}_1(\mathcal{G}_z) \cup \mathcal{E}_2(\mathcal{G}_z)$  for  $z = 2, \dots, m$ .

Suppose  $\mathcal{G}_z (z \geq 2)$  is given and has more than one nodes, then we can choose from tree  $\mathcal{G}_z$  a leaf node  $\pi_z$  which is not the root node. By inductive assumption there exists some  $\tau_z$  with  $(\pi_z, \tau_z) \in \mathcal{E}_1(\mathcal{G}_z)$  and some  $\nu_z$  with  $(\tau_z, \nu_z) \in \mathcal{E}_2(\mathcal{G}_z)$ , and  $\mathcal{G}_{z+1}$  is still a tree where  $\mathcal{G}_{z+1}$  is obtained by removing  $\{\pi_z, \tau_z\}$  and their associated edges from  $\mathcal{G}_z$ . There are two possible bounds:

1.  $\theta_{z,\tau_z} \geq 2$ . In this case, let  $\theta_{z+1}$  be the same as  $\theta_z$  on all components except  $\theta_{z+1,\tau_z} = \theta_{z,\tau_z} - 2$ . Then we have

$$\begin{aligned}
\Sigma(\mathcal{G}_z, \theta_z) &= \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_z)} \prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_z)} |[B^n]_{v_p, v_q}| / F(\mathcal{G}_z, \mathbf{v}, \theta_z) \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \sum_{v_{\tau_z}} \sum_{v_{\pi_z} \in \mathcal{N}(v_{\tau_z})} \prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}| |[B^{n(\tau_z, \nu_z)}]_{v_{\tau_z}, v_{\nu_z}}| / \left( F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1}) d_{v_{\tau_z}} \right) \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \sum_{v_{\tau_z}} \sum_{v_{\pi_z} \in \mathcal{N}(v_{\tau_z})} \frac{1}{d_{\tau_z}} |[B^{n(\tau_z, \nu_z)}]_{v_{\pi_z}, v_{\nu_z}}| \\
&= \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \sum_{v_{\tau_z}} |[B^{n(\tau_z, \nu_z)}]_{v_{\pi_z}, v_{\nu_z}}| \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \|B\|_{\infty}^{\Delta(\mathcal{G}_z, \mathcal{G}_{z+1})} \\
&= \Sigma(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1}) \|B\|_{\infty}^{\Delta(\mathcal{G}_z, \mathcal{G}_{z+1})}
\end{aligned}$$

2. It is trivial that  $\theta_{z,\pi_z} \geq 1$  and  $\theta_{z,\tau_z} \geq 1$ . So we can always let  $\theta_{z+1}$  be the same as  $\theta_z$  on all components except  $\theta_{z+1,\pi_z} = \theta_{z,\pi_z} - 1$  and  $\theta_{z+1,\tau_z} = \theta_{z,\tau_z} - 1$ . Then we have

$$\begin{aligned}
\Sigma(\mathcal{G}_z, \theta_z) &= \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_z)} \prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_z)} |[B^n]_{v_p, v_q}| / F(\mathcal{G}_z, \mathbf{v}, \theta_z) \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \sum_{v_{\tau_z}} \sum_{v_{\pi_z} \in \mathcal{N}(v_{\tau_z})} \prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}| |[B^{n(\tau_z, \nu_z)}]_{v_{\tau_z}, v_{\nu_z}}| / \left( F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1}) \sqrt{d_{\tau_z} d_{\pi_z}} \right) \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \sum_{v_{\tau_z}} \sum_{v_{\pi_z} \in \mathcal{N}(v_{\tau_z})} \frac{1}{\sqrt{d_{\tau_z} d_{\pi_z}}} |[B^{n(\tau_z, \nu_z)}]_{v_{\pi_z}, v_{\nu_z}}| \\
&= \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \sum_{v_{\tau_z}} \omega_{\tau_z} |[B^{n(\tau_z, \nu_z)}]_{v_{\pi_z}, v_{\nu_z}}| \\
&\leq \sum_{\mathbf{v} \in \mathcal{U}(\mathcal{G}_{z+1})} \frac{\prod_{(p,q,n) \in \mathcal{E}_2(\mathcal{G}_{z+1})} |[B^n]_{v_p, v_q}|}{F(\mathcal{G}_{z+1}, \mathbf{v}, \theta_{z+1})} \|B\|_{\infty}^{\Delta(\mathcal{G}_z, \mathcal{G}_{z+1})-1} \|B\|_{\omega} \\
&= \Sigma(\mathcal{G}_{z+1}, \theta_{z+1}) \|B\|_{\infty}^{\Delta(\mathcal{G}_z, \mathcal{G}_{z+1})-1} \|B\|_{\omega}
\end{aligned}$$

Since at first  $\theta_{2,p} \geq 1$  for all non-root nodes  $p$  and  $\theta_{z+1,p} < \theta_{z,p}$  happens only if node  $p$  is removed from  $\mathcal{G}_z$ , the second bound always works. Besides, we have bounds

$$\Sigma(\mathcal{G}_z, \theta_z) \leq \Sigma(\mathcal{G}_{z+1}, \theta_{z+1}) \|B\|_{\infty}^{\Delta(\mathcal{G}_z, \mathcal{G}_{z+1})-1} \min\{\|B\|_{\infty}, \|B\|_{\omega}\} \quad (42)$$

if  $\theta_{z,\tau_z} \geq 2$ .

$$\Sigma(\mathcal{G}_m, \theta_m) = d_{\min}^{-\sum_i \theta_{m,i}/2}.$$

Now let's calculate  $\sum_i \theta_{m,i}$ . Initially we have

$$\sum_i \theta_{2,i} = \sum_i \theta_{1,i} = 2k(2l-1).$$

The inductive steps give  $\sum_i \theta_{z+1,i} = \sum_i \theta_{z,i} - 2$  and there are  $kl - u$  inductive steps in total. Hence

$$\sum_i \theta_{m,i} = 2k(2l-1) - 2(kl - u) = 2kl - 2k + 2u,$$

which means  $\Sigma(\mathcal{G}_m, \boldsymbol{\theta}_m) = d_{\min}^{-kl+k+u}$ .

Note that there are at most  $2k$  nodes  $p$  with  $\theta_{2,p} = 1$ , which means the bound (42) holds for all inductive steps except  $2k$  steps. In this way, together with the fact  $\|B\|_{\max} \leq \|B\|_{\infty}$  and  $\Delta(\mathcal{G}_1, \mathcal{G}_m) = t$ ,  $\Sigma(\mathcal{G}_1, \boldsymbol{\theta}_1)$  could be expressed as

$$\begin{aligned}\Sigma(\mathcal{G}_1, \boldsymbol{\theta}_1) &\leq \left(\frac{\|B\|_{\omega}}{\|B\|_{\infty}}\right)^{2k} \|B\|_{\max}^{\Delta(\mathcal{G}_1, \mathcal{G}_2)} \|B\|_{\infty}^{\Delta(\mathcal{G}_2, \mathcal{G}_m)} \Sigma(\mathcal{G}_m, \boldsymbol{\theta}_m) \\ &\leq \left(\frac{\|B\|_{\omega}}{\|B\|_{\infty}}\right)^{2k} \|B\|_{\max}^{kl} \|B\|_{\infty}^{\Delta(\mathcal{G}_1, \mathcal{G}_m) - kl} d_{\min}^{-kl+k+u} \\ &\leq \|B\|_{\omega}^{2k} \|B\|_{\max}^{kl} \|B\|_{\infty}^{t-kl-2k} d_{\min}^{-k(l-1)+u}\end{aligned}$$

For  $\mathcal{G}_m$  which contains a single root node, we have

$$\Sigma(\mathcal{G}_m, \boldsymbol{\theta}_m) = d_{\min}^{-\sum_i \theta_{m,i}/2} = d_{\min}^{-k(l-1)-u},$$

where the last equation comes from the fact  $\sum_i \theta_{1,i} = 2k(2l-1)$  and  $\sum_i \theta_{z+1,i} = \sum_i \theta_{z,i} - 2$  if and only an edge in  $\mathcal{E}_2(\mathcal{G}_z)$  was removed, but there are exact  $kl - u$  edges in  $\mathcal{E}_2(\mathcal{G}_1)$  so that

$$\sum_i \theta_{m,i} = \sum_i \theta_{1,i} - 2(kl - u) = 2kl - 2k + 2u.$$

Note that the third cases would appear at most  $2k$  times and the second cases appear exactly  $kl$  times, which means

$$\Sigma(\mathcal{G}_1, \boldsymbol{\theta}_1) \leq \|B\|_{\max}^{kl} \|B\|_{\omega}^{2k} \|B\|_{\infty}^{\Delta(\mathcal{G}_1, \mathcal{G}_m) - kl - 2k} \Sigma(\mathcal{G}_m, \boldsymbol{\theta}_m)$$

Return to calculate (41). We have

$$\begin{aligned}&\sum_{(\hat{\alpha}, \hat{\beta}) \in \mathcal{J}_{2k}} \prod_{r=1}^{2k} \frac{|a_{1, \alpha_1^{(r)}}^{(1)} a_{\beta_1^{(r)}, \alpha_2^{(r)}}^{(2)} \cdots a_{\beta_{l-1}^{(r)}, \alpha_l^{(r)}}^{(l)}|}{\sqrt{d_{\alpha_1^{(r)}} d_{\beta_1^{(r)}} \cdots d_{\alpha_{l-1}^{(r)}} d_{\beta_{l-1}^{(r)}} d_{\alpha_l^{(r)}}}} \\ &\leq \sum_{\mathcal{G}_1} \Sigma(\mathcal{G}_1, \boldsymbol{\theta}_1) \\ &= \sum_{u \geq 0} \sum_{|\mathcal{V}(\mathcal{G}_1)| = kl - u + 1} \Sigma(\mathcal{G}_1, \boldsymbol{\theta}_1) \\ &\leq \|B\|_{\omega}^{2k} \|B\|_{\max}^{kl} \|B\|_{\infty}^{t-kl-2k} d_{\min}^{-k(l-1)} \sum_{u \geq 0} \frac{(2kl)!}{2^{kl} 4.5^u (kl-3u)! (2u)!} d_{\min}^{-u}\end{aligned}$$

For the last line above, we have

$$\frac{(2kl)!}{(kl-3u)! (2u)! d_{\min}^u} \leq \frac{(2kl)!}{(kl-3u)! (2u)! (kl)^u} \leq (Ckl)^{kl}$$

for some absolute constant  $C$ . Thus

$$\sum_{(\hat{\alpha}, \hat{\beta}) \in \mathcal{J}_{2k}} \prod_{r=1}^{2k} \frac{|a_{1, \alpha_1^{(r)}}^{(1)} a_{\beta_1^{(r)}, \alpha_2^{(r)}}^{(2)} \cdots a_{\beta_{l-1}^{(r)}, \alpha_l^{(r)}}^{(l)}|}{\sqrt{d_{\alpha_1^{(r)}} d_{\beta_1^{(r)}} \cdots d_{\alpha_{l-1}^{(r)}} d_{\beta_{l-1}^{(r)}} d_{\alpha_l^{(r)}}}} \leq (C \log n)^{kl} \|B\|_{\omega}^{2k} \|B\|_{\max}^{kl} \|B\|_{\infty}^{t-kl-2k} d_{\min}^{-k(l-1)}$$

for some absolute constant  $C$ . □

### 3.3.1 A Lower Bound on $\|B\|_\infty$

In the simple case, it is clear that  $B = L^\dagger$ ,  $BL = I - \mathbf{s}\mathbf{s}^T$ ,  $B\mathbf{s} = 0$ . In the following I will explore the relationship between  $\|B\|_\infty$  and  $d_{\max}/d_{\min}$ . Without loss of generality, suppose  $d_1 = d_{\max}$  and  $d_2 = d_{\min}$ , then consider the vector which is a linear combination of the first column of  $L$  denoted as  $\mathbf{col}_1(L)$  and  $\mathbf{s}$ :

$$\mathbf{v} = \mathbf{col}_1(L) - \frac{1 - 1/\sqrt{d_1 d_2}}{\sqrt{d_1} + \sqrt{d_2}} \mathbf{s}.$$

In this way, it can be easily verified that

$$\|\mathbf{v}\|_\infty = \frac{\sqrt{d_2} + 1/\sqrt{d_2}}{\sqrt{d_1} + \sqrt{d_2}}$$

and

$$B\mathbf{v} = B\mathbf{col}_1(L) = \mathbf{col}_1(I - \mathbf{s}\mathbf{s}^T),$$

so

$$\|B\mathbf{v}\|_\infty \geq 1 - \frac{d_1}{|\mathcal{E}|} \geq \frac{1}{2}$$

since  $|\mathcal{E}| \geq 2d_1$ . Hence we have

$$\|B\|_\infty \geq \frac{\sqrt{d_1} + \sqrt{d_2}}{2(\sqrt{d_2} + 1/\sqrt{d_2})} \geq \frac{1}{4} \sqrt{\frac{d_1}{d_2}} \geq \frac{1}{4} \sqrt{\frac{d_{\max}}{d_{\min}}}.$$

### 3.4 Proof of Lemma 3.3

Suppose

$$K \leq C_1 \sqrt{\frac{d_{\min}}{\|B_i\| \cdot \|B_i\|_\infty \cdot \log n}},$$

the inequality (31) turns into

$$\|(\mathbf{e}_b^T \otimes I_m) \prod_{j=1}^l ((B_i^{t_j} \otimes I_m) N) S_k\|_\infty \leq C^l \|B_i\|_\infty^{t-l} \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \|U_k\|.$$

Thus we have

$$\|(\mathbf{e}_b^T \otimes I_m) (B_i N)^{i_1} B_i \dots (B_i N)^{i_{j+1}} B_i N (S_k \otimes I_m)\| \leq \sqrt{\frac{d_{\min}}{|\mathcal{E}|}} \|U_k\| \cdot \|B_i\|_\infty^j (C_1 C)^{\sum_{l=1}^{j+1} i_l + 1}.$$

On the other hand, by choosing small enough absolute constant  $C_1$ , we have

$$\|B_i\| \|N\| \leq \|B_i\| \frac{C_2 K}{\sqrt{d_{\min}}} \leq \frac{1}{10},$$

$$\|B_i\| \|\lambda_i I_m - \Sigma_i\| \leq \|B_i\| \|N\| \leq \frac{1}{10}.$$

Thus applying Lemma 2.4 provides the desired conclusion.

## 4 Proof of Exact Recovery Conditions Under the Full Setting

### 4.1 Proof of Theorem 3.1

We prove Theorem 3.1 by establishing the concentration under the un-normalized data, i.e., which is identical to set  $d_i = nt$ . To begin with, let us rewrite the data matrix under the proposed model of mapping and



observation graph. Since we assume the observation graph and the input pair-wise maps are generated through independent procedures, it follows that

$$\bar{X}_{ij} = \frac{1}{nt} \cdot \begin{cases} I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T & \text{with probability } \eta_{ij}t \\ U_{P_m} - \frac{1}{m} \mathbf{1}\mathbf{1}^T & \text{with probability } (1 - \eta_{ij})t \\ 0 & \text{with probability } (1 - t) \end{cases} \quad (43)$$

Here  $\eta_{ij}, 1 \leq i, j \leq n$  form a matrix  $(p - q)(I_k \otimes (\mathbf{1}\mathbf{1}^T)) + q\mathbf{1}\mathbf{1}^T$ .

(43) gives rise to

$$E[\bar{X}_{ij}] = \frac{\eta_{ij}}{n} (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T), \quad (44)$$

and

$$N_{ij} := \bar{X}_{ij} - E[\bar{X}_{ij}] = \frac{1}{nt} \cdot \begin{cases} (1 - \eta_{ij}t)(I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T) & \text{with probability } \eta_{ij}t \\ U_{P_m} - \frac{1}{m} \mathbf{1}\mathbf{1}^T - \eta_{ij}t(I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T) & \text{with probability } (1 - \eta_{ij})t \\ -\eta_{ij}t(I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T) & \text{with probability } (1 - t) \end{cases} \quad (45)$$

It is obvious that  $N_{ij}\mathbf{1} = 0$  and  $N_{ij}^T\mathbf{1} = 0$ . Moreover,

$$\|N_{ij}\| \leq \frac{1 + \eta_{ij}t}{nt} \leq \frac{2}{nt}.$$

Decompose  $\bar{X} = E[\bar{X}] + N$ , it follows that

$$E[\bar{X}] = \frac{1}{n} ((p - q)(I_k \otimes (\mathbf{1}\mathbf{1}^T)) + q(\mathbf{1}\mathbf{1}^T)) \otimes (I_m - \frac{1}{m} \mathbf{1}\mathbf{1}^T). \quad (46)$$

Following the convention of notation, let

$$A = \frac{1}{n} ((p - q)(I_k \otimes (\mathbf{1}\mathbf{1}^T)) + q(\mathbf{1}\mathbf{1}^T)).$$

It is easy to check that the rank of  $A$  is  $k$ , and its top  $k$  eigenvalues are given by

$$\lambda_1(A) = q + \frac{p - q}{k}, \quad \lambda_i(A) = \frac{p - q}{k}, \quad 2 \leq i \leq k. \quad (47)$$

Let  $(\frac{1}{\sqrt{k}}\mathbf{1}, H_k)$  be an orthonormal basis for  $\mathbb{R}^k$ , then it is easy to see that the corresponding top  $k$  eigenvectors of  $A$  are given by

$$S_k = \sqrt{\frac{k}{n}} \mathbf{1} \otimes (\frac{1}{\sqrt{k}}\mathbf{1}, H_k).$$

Moreover,

$$\bar{S}_k \bar{S}_k^T = (I_{n_0} - \frac{1}{n_0} \mathbf{1}\mathbf{1}^T) \otimes I_k.$$

To apply Lemma 3.3, it is easy to see that

$$B_1 = \frac{1}{q + \frac{p - q}{k}} (I_{n_0} - \frac{1}{n_0} \mathbf{1}\mathbf{1}^T) \otimes I_k,$$

and

$$B_i = \frac{k}{p - q} (I_{n_0} - \frac{1}{n_0} \mathbf{1}\mathbf{1}^T) \otimes I_k, \quad 2 \leq i \leq k.$$

Denote

$$T = (I_{n_0} - \frac{1}{n_0} \mathbf{1}\mathbf{1}^T) \otimes I_k.$$

It is easy to check that

$$\|T\|_\omega = 2, \quad \|T\|_\infty = 2, \quad \|T\|_{\max} = 1.$$

Applying Lemma 3.3, we obtain the following stability bound on the top  $k(m - 1)$  eigen-vectors of  $\bar{X}$ :

**Lemma 4.1.** Let  $\bar{U} = (\bar{U}_1^T, \dots, \bar{U}_n^T)^T$  be the top  $k(m-1)$  eigen-vectors of  $\bar{X}$ . Then there exists a rotation matrix  $\bar{R} \in O(k(m-1))$  and a universal constant  $c$ , so that when

$$p - q \geq ck \sqrt{\frac{\log(n)}{nt}},$$

we have w.h.p,

$$\max_{1 \leq i \leq n} \|U_i - (e_i^T S_k \otimes H_m) \cdot \bar{R}\| \leq \frac{1}{6} \cdot \frac{1}{\sqrt{n}}. \quad (48)$$

**Complete the proof of Theorem 3.1.** Since the spectral norm bounds the difference between the corresponding rows, it follows from Lemma 4.1 that (1) the distance between the corresponding elements between each pair of objects in the embedding space is upper bounded by  $1/3$ , (2)  $d_{\text{intra}} < 1/3$ , and (3)  $d_{\text{inter}} > 2/3$ . This means both the intra-cluster maps and the underlying clusters can be recovered, which ends the proof.

## 4.2 Proof of Theorem 3.2

We prove a stronger recovery condition for inter-cluster maps. Note that inter-cluster map recovery solves the following linear assignment:

$$X_{st} = \operatorname{argmax}_{X \in \mathcal{P}_m} \langle X, C_{st} \rangle, \quad C_{st} = \sum_{(i,j) \in \mathcal{E}, i \in c_s, j \in c_t} X_{ji} X_{ij}^{\text{in}} X_{is} i. \quad (49)$$

We prove a stronger exact recovery condition as follows. To begin with, we define the minimum number of inter-cluster edges between one pair of clusters as.

$$N_{\text{inter}} = \min_{1 \leq s < t \leq k} N_{st}, \quad N_{st} := |\{(i,j) | (i,j) \in \mathcal{E}, i \in c_s, j \in c_t\}|. \quad (50)$$

**Lemma 4.2.** Given an absolute constant  $c_{\text{inter}} > 0$ . Suppose the intra-cluster rate  $q$  and  $N_{\text{inter}}$  satisfy the following constraint:

$$q \geq \sqrt{\frac{c_{\text{inter}} \log(n)}{N_{\text{inter}}}}.$$

Then we have with probability at least  $1 - \frac{m^2 k^2}{n \frac{c_{\text{inter}}}{8}}$ ,

$$\min_{1 \leq a \leq m} C_{st}(a, a) > \max_{1 \leq a \neq b \leq m} C_{st}(a, b), \quad 1 \leq s \neq t \leq k. \quad (51)$$

*Proof:* First of all, it is easy to check that

$$E[C_{st}(a, b)] = \begin{cases} \frac{1}{m}(1-q) + q & a = b \\ \frac{1}{m}(1-q) & a \neq b \end{cases}$$

We apply union bounds by showing that with probability at least  $1 - \frac{m^2 k^2}{n \frac{c_{\text{inter}}}{8}}$ , we have

$$\min_{1 \leq s \neq t \leq k, 1 \leq a \leq m} C_{st}(a, a) > \frac{1}{m}(1-q) + \frac{q}{2}, \quad (52)$$

$$\max_{1 \leq s \neq t \leq k, 1 \leq a \neq b \leq m} C_{st}(a, b) < \frac{1}{m}(1-q) + \frac{q}{2}. \quad (53)$$

Note that each diagonal element  $X_{ji} X_{ij}^{\text{in}} X_{is} i(a, a)$  is a random Bernoulli random variable with probability  $\frac{1-q}{m} + q$ , we can apply lower Chernoff bound to obtain a lower tail bound on  $C_{st}(a, a)$ , which is

$$\Pr[C_{st}(a, a) \leq N_{st}(\frac{1-q}{m} + q - \frac{q}{2})] \leq \exp(-\frac{N_{st} q^2}{8(q + \frac{1-q}{m})}) \leq \exp(-\frac{N_{\text{inter}} q^2}{8(q + \frac{1-q}{m})}). \quad (54)$$

Similarly, each off-diagonal element  $X_{jt} X_{ij}^{in} X_{isi}(a, b)$  is a random Bernoulli random variable with probability  $\frac{1-q}{m}$ , we can apply upper Chernoff bound to obtain a upper tail bound on  $C_{st}(a, b)$ , which is

$$Pr[C_{st}(a, b) \leq N_{st}(\frac{1-q}{m} + \frac{q}{2})] \leq \exp(-\frac{N_{st}q^2}{8(\frac{q}{6} + \frac{1-q}{m})}) \leq \exp(-\frac{N_{inter}q^2}{8(\frac{q}{6} + \frac{1-q}{m})}). \quad (55)$$

Since  $q = \sqrt{c_{inter} \frac{\log(n)}{N_{inter}}}$ . It follows that combining (54) and (55) lead to

$$Pr[C_{st}(a, a) \leq N_{st}(\frac{1-q}{m} + q - \frac{q}{2})] \leq \exp(-\frac{c_{inter} \log(n)}{8}) \leq \frac{1}{n^{\frac{c_{inter}}{8}}} \quad (56)$$

$$Pr[C_{st}(a, b) \geq N_{st}(\frac{1-q}{m} + \frac{q}{2})] \leq \frac{1}{n^{\frac{c_{inter}}{8}}} \quad (57)$$

Applying union bounds (56) and (57), we have that the inter-cluster maps can be recovered with probability at least  $1 - \frac{m^2 k^2}{n^{\frac{c_{inter}}{8}}}$ .  $\square$

Since the observation graph is generated from the Erdős-Rényi model  $G(n, t)$ . It is easy to check that the number of inter-cluster edges between a pair of clusters concentrates at  $[\frac{n^2 t}{2k^2}, \frac{2n^2 t}{k^2}]$  with overwhelming probability (for example using Chernoff bound), which ends the proof.  $\square$

## References

[Eldridge et al., 2018] Eldridge, J., Belkin, M., and Wang, Y. (2018). Unperturbed: spectral analysis beyond davis-kahan. volume abs/1706.06516. [1](#)