Submitted to the Annals of Statistics

ESTIMATING NETWORK MEMBERSHIPS BY SIMPLEX VERTEX HUNTING

BY JIASHUN JIN[‡], ZHENG TRACY KE[§], AND SHENGMING LUO[‡]

Carnegie Mellon University[‡] *and Harvard University* [§]

Consider an undirected mixed membership network with n nodes and K communities. For each node i, $1 \leq i \leq n$, we model the membership by a Probability Mass Function (PMF) $\pi_i = (\pi_i(1), \pi_i(2), \ldots, \pi_i(K))'$, where $\pi_i(k)$ is the probability that node i belongs to community k, $1 \leq k \leq K$. We call node i "pure" if π_i is degenerate and "mixed" otherwise. The primary interest is to estimate π_i, $1 \leq i \leq n$.

We model the adjacency matrix A with a Degree Corrected Mixed Membership (DCMM) model, and let $\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_K$ be the first K eigenvectors. Define a matrix $\hat{R} \in \mathbb{R}^{n,K-1}$ by $\hat{R}(i,k) = \hat{\xi}_{k+1}(i)/\hat{\xi}_1(i)$, $1 \leq k \leq K-1$, $1 \leq i \leq n$. The oracle counterpart of $\hat{R}$ (denoted by R) under the DCMM model contains all information we need for the memberships. In fact, we have an interesting insight: there is a simplex $\mathcal{S}$ in $\mathbb{R}^{K-1}$ such that row i of R corresponds to a vertex of $\mathcal{S}$ if node i is pure, and corresponds to an interior point of $\mathcal{S}$ otherwise. Vertex Hunting (i.e., estimating the vertices of $\mathcal{S}$) is therefore the key to our problem.

We propose a new approach *Mixed-SCORE* to membership estimation, with an easy-to-use Vertex Hunting step. We derive the convergence rate of Mixed-SCORE using delicate spectral analysis, especially tight row-wise deviation bounds for $\hat{R}$. We have also applied it to 4 network data sets with encouraging results.

1. Introduction. In the study of social networks, the problem of estimating the mixed memberships has received a lot of attention [3, 4, 27, 60]. Consider an undirected network $\mathcal{N} = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the set of nodes and E is the set of edges. We assume that the network consists of K perceivable communities

$$\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_K, \tag{1.1}$$

and that for each node i, there is a Probability Mass Function (PMF) $\pi_i = (\pi_i(1), \pi_i(2), \ldots, \pi_i(K))' \in \mathbb{R}^K$ such that

$$\pi_i(k) \text{ is the "weight" of node } i \text{ on } \mathcal{C}_k,\ 1 \leq k \leq K. \tag{1.2}$$

Primary 62H30, 91C20; secondary 62P25.

Keywords and phrases: citation network, coauthor network, communities, complexity reduction, Ideal Simplex, k-means, mixed membership, SCORE, spectral clustering

Table 1
Four network data sets (K is determined by prior knowledge; see Section 4).

Data name	Source	Nodes	Edges	n	K
Polbook	Krebs (unpublished)	books	frequently co-purchased	105	2
Football	Girvan & Newman (2002)	teams	games played	115	4
Coauthor	Ji & Jin (2016)	authors	coauthorship	236	2
Citee	Ji & Jin (2016)	authors	cited by the same authors	1790	3

We call node i "pure" if π_i is degenerate (i.e., one entry is 1, all others are 0) and "mixed" otherwise. The primary interest is to estimate π_i, $1 \le i \le n$.

Table 1 lists several data sets we study in this paper. Take the Polbook for example: each node is a book on politics of USA sold by Amazon.com, and there is an edge between two nodes if they are frequently co-purchased; the nodes are manually labeled as either "Conservative", "Liberal", or "Neutral". We model the data with a two-community ("Conservative" and "Liberal") mixed membership model, where a "Neutral" node is thought of as having mixed memberships. The goal is to use the network information to estimate how much weight each book puts on "Conservative" and "Liberal".

Alternatively, one could use a non-mixing non-overlapping model (e.g., [39]) where we have three communities: "Conservative", "Liberal", and "Neutral". However, we prefer to use a mixed membership model for

- A non-mixing model usually assumes more communities than necessary, and some of them may be hard to interpret or not meaningful.
- A mixed membership model allows us to assess the weight each node puts on each community, while a non-mixing model does not.

Such a viewpoint is valid for many data sets including the Polbook. See Section 4 for more discussion and also details for all the data sets.

Existing works on mixed membership estimation include the LDA method [3], the tensor method [4], and the OCCAM method [60]. Also, [46, 50] proposed some spectral methods (note these works appear more than 15 months later than the first submission of our paper to the Annals of Statistics). Note that most works (e.g., [3, 4, 46, 50]) focus on the case where we do not model degree heterogeneity. As the primary interest of this paper is on the case where we have severe degree heterogeneity, how to extend these works to our setting is unclear. Zhang *et al* [60] allows for severe degree heterogeneity, but their method requires that the fraction of mixed nodes is relatively small, which is relatively restrictive and thus leaves room for improvement.

We propose *Mixed-SCORE* as a new approach to membership estimation, which accommodates the much broader settings where almost all nodes may be mixed and where we allow severe degree heterogeneity. We analyze Mixed-

SCORE using delicate spectral analysis, especially tight entry-wise deviation bounds for the leading eigenvectors of A. Mixed-SCORE is applied to all data sets in Table 1 with interesting results.

In theory, we derive the rate of convergence for the maximum ℓ^1-error on estimating π_i's. Our theoretical framework allows the number of communities, the average degree, and the relative similarity across communities (to be defined) to depend on n. Compared with existing theoretical results, our settings are broader and our error rate is sharper (see Section 1.5).

1.1. *Degree Corrected Mixed Membership (DCMM) model.* To facilitate the analysis, we use the DCMM model. DCMM can be viewed as an extension of the *Mixed Membership Stochastic Block (MMSB)* model by Airoldi *et al.* [3], to accommodate degree heterogeneity, and can also be viewed as an extension of the *Degree Corrected Block Model* (DCBM) by Karrer and Newman [39], to accommodate mixed memberships. DCMM can also be viewed as a reparametrization of the OCCAM model by Zhang *et al.* [60].

Let $A \in \mathbb{R}^{n,n}$ be the (symmetric) adjacency matrix of $\mathcal{N}$; by convention, the diagonals of A are 0, since self-edges are not counted. Let $P \in \mathbb{R}^{K,K}$ be a symmetric non-negative matrix

$$\text{that is non-singular, and has unit diagonals.} \tag{1.3}$$

For positive degree heterogeneity parameters $\theta(1), \theta(2), \ldots, \theta(n)$, DCMM models the upper triangular entries of A (excluding the diagonal entries) as independent Bernoulli random variables satisfying

$$P(A(i,j) = 1) = \theta(i)\theta(j) \sum_{k=1}^{K} \sum_{\ell=1}^{K} \pi_i(k)\pi_j(\ell)P(k,\ell), \qquad 1 \leq i,j \leq n, \tag{1.4}$$

Introduce the degree heterogeneity matrix $\Theta \in \mathbb{R}^{n,n}$ and the membership matrix $\Pi \in \mathbb{R}^{n,K}$:

$$\Theta = \mathrm{diag}\big(\theta(1), \theta(2), \ldots, \theta(n)\big),^{1} \qquad \Pi = [\pi_1, \pi_2, \ldots, \pi_n]'. \tag{1.5}$$

DEFINITION 1.1. *We call model (1.1)-(1.4) the Degree Corrected Mixed Membership (DCMM) model, and denote it by $DCMM_n(K, P, \Theta, \Pi)$.*

Remark. DCMM includes several other well-known network models as special cases. When all π_i's are degenerate, (1.4) implies $P(A(i,j) = 1) =$

[1] For a vector v, $\mathrm{diag}(v)$ is the diagonal matrix with entries of v on its diagonal. For a matrix M, $\mathrm{diag}(M)$ is the diagonal matrix with diagonal elements of M on its diagonal.

$\theta(i)\theta(j)P(k,\ell)$ for $i \in \mathcal{C}_k$ and $j \in \mathcal{C}_\ell$, and DCMM reduces to DCBM [39]. When $\theta(i) \equiv \sqrt{\alpha_n}$, (1.4) implies $P(A(i,j)=1) = \alpha_n \sum_{\ell=1}^K \pi_i(k)\pi_j(\ell)P(k,\ell)$, and DCMM reduces to (a simplified version of) MMSB [3]. The DCMM is equivalent to the OCCAM [60], but the difference is: OCCAM normalizes π_i to have unit-ℓ^2-norm, so it is not a PMF, while DCMM normalizes π_i to have unit ℓ^1-norm, so it is a PMF. Viewing π_i as a PMF is more consistent with the conventional definition of "mixed-memberships" [3], and helps interpret the membership estimation in practice.

We now decompose A into the sum of a "signal" part and a "noise" part:

$$A = [\Omega - \mathrm{diag}(\Omega)] + W, \qquad \text{where } \Omega = \Theta\Pi P\Pi'\Theta. \tag{1.6}$$

Note that $\Omega(i,j) = P(A(i,j)=1)$ when $i \neq j$ and W is a generalized Wigner matrix [56]. Here, P is the matrix that directly models the difference across communities. Our primary interest is the membership matrix Π.

The proposition below shows that the DCMM model is identifiable.

PROPOSITION 1.1 (Identifiability). *When each community has at least one pure node, the DCMM model is identifiable: For eligible* (Θ, Π, P) *and* $(\tilde{\Theta}, \tilde{\Pi}, \tilde{P})$, *if* $\Theta\Pi P\Pi'\Theta = \tilde{\Theta}\tilde{\Pi}\tilde{P}\tilde{\Pi}'\tilde{\Theta}$, *then* $\Theta = \tilde{\Theta}$, $\Pi = \tilde{\Pi}$, *and* $P = \tilde{P}$.

Remark. Compared to other models (e.g., MMSB, DCBM), DCMM has many more parameters (for degree heterogeneity and for mixed memberships). These parameters have more degrees of freedom than those in MMSB or DCBM, and so DCMM requires stronger conditions to be identifiable. For example, the identifiability of MMSB does not need P to have unit diagonals, and the identifiability of DCBM does not require that P has full rank. See the appendix for a detailed discussion of parameter identifiability.

1.2. *The Ideal Simplex (IS) and the Ideal Mixed-SCORE.* We start by considering the *oracle* case where Ω is given. In Section 1.3, we extend what we have learned in the oracle case to the real case.

Noting $\mathrm{rank}(\Omega) = K$, we let $\lambda_1, \lambda_2, \cdots, \lambda_K$ be all the nonzero eigenvalues of Ω (arranged in the descending order in magnitude), and let $\xi_1, \cdots, \xi_K$ be the corresponding eigenvectors. Write $\Xi = [\xi_1, \cdots, \xi_K]$. By elementary linear algebra, Ξ and Ω have the same column space; at the same time, since $\Omega = \Theta\Pi P\Pi'\Theta$, the column space of Ω is equal to the column space of $\Theta\Pi$. It follows that there is a unique non-singular matrix $B = [b_1, b_2, \ldots, b_K] \in \mathbb{R}^{K,K}$ such that

$$\Xi = \Theta\Pi \cdot B. \tag{1.7}$$

By Perron's theorem [31], without loss of generality, we can assume all entries of ξ_1 and all entries of b_1 are strictly positive (Lemma B.2 and Lemma B.4).

The goal of the oracle approach is to use Ξ to exactly recover Π. [2] One of the major nuisance is the diagonal matrix Θ; see (1.7). However, if we divide each column of Ξ by its first column entry-wise, then the matrix Θ is cancelled out; this normalization on eigenvectors was introduced in [35] and is called the SCORE normalization. In light of this, we define the *Matrix of Entry-wise Ratios* $R \in \mathbb{R}^{n,K-1}$ by

$$R(i,k) = \xi_{k+1}(i)/\xi_1(i), \qquad 1 \le i \le n, \;\; 1 \le k \le K-1. \tag{1.8}$$

We shall reveal a low-dimensional geometrical structure associated with the rows of R, which is the key to estimating Π.

Let $V = [v_1, v_2, \ldots, v_K] \in \mathbb{R}^{K-1,K}$ be the matrix

$$v_k(\ell) = b_{\ell+1}(k)/b_1(k), \qquad 1 \le \ell \le K-1, \;\; 1 \le k \le K. \tag{1.9}$$

The vector b_1 plays a pivotal role. The next lemma is proved in the appendix.

Lemma 1.1. *For $1 \le k \le K$, $b_1(k) = [\lambda_1 + v_k' \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_K) v_k]^{-1/2}$.*

Introduce

$$w_i = (b_1 \circ \pi_i)/\|b_1 \circ \pi_i\|_1, \qquad 1 \le i \le n,$$

where '$\circ$' is the Hadamard product (i.e., $b_1 \circ \pi_i$ denotes the vector in $\mathbb{R}^K$ where the k-th entry is $b_1(k)\pi_i(k)$). Since all entries of b_1 are positive, w_i is a weight vector.[3] Write $R = [r_1, r_2, \ldots, r_n]'$. By (1.7) and direct calculations,

$$R = \begin{pmatrix} r_1' \\ r_2' \\ \cdots \\ r_n' \end{pmatrix} = \begin{pmatrix} (b_1 \circ \pi_1)/\|b_1 \circ \pi_1\|_1 \\ (b_1 \circ \pi_2)/\|b_1 \circ \pi_2\|_1 \\ \cdots \\ (b_1 \circ \pi_n)/\|b_1 \circ \pi_n\|_1 \end{pmatrix} \begin{pmatrix} v_1' \\ v_2' \\ \cdots \\ v_K' \end{pmatrix} \equiv \begin{pmatrix} w_1' \\ w_2' \\ \cdots \\ w_n' \end{pmatrix} V'.$$

It turns out that, the rows of R form a $(K-1)$-simplex [4] in $\mathbb{R}^{K-1}$ which we call the *Ideal Simplex (IS)*, with $v_1, v_2, \ldots, v_K$ being the vertices. Denoting the simplex by $\mathcal{S}^{ideal}(v_1, v_2, \ldots, v_K)$, we have

- A pure row (row i of R is pure if node i is pure and is mixed otherwise) falls on one of the K vertices, and a mixed row falls in the interior.

[2] The choice of Ξ is not unique, and different choices are different by an orthogonal column transformation; still, they give exactly the same oracle reconstruction of Π.

[3] We call a vector a weight vector if all its entries are nonnegative with a sum of 1.

[4] A k-simplex is the k-dimensional polytope that is the convex hull of its $(k+1)$ vertices.

- Each r_i is a convex linear combination of $v_1, \ldots, v_K$ with weights w_i:

$$(1.10) \qquad r_i = V w_i = \sum_{k=1}^{K} w_i(k) v_k, \qquad \text{where} \qquad w_i \propto (b_1 \circ \pi_i).$$

This low-dimensional geometrical structure allows us to conveniently retrieve the vertices $v_1, v_2, \ldots, v_K$ using the point cloud: When each community has at least one pure node, the convex hull of $\{r_i\}_{i=1}^n$ is exactly equal to the Ideal Simplex. Therefore, to retrieve $v_1, v_2, \ldots, v_K$, we only need to compute the convex hull of the point cloud, which is a well-studied problem in computational geometry. There are many available convex hull algorithms, such as beneath-beyond, Gift-wrapping, and Seidel's shelling; see [52] for a review. Since our target convex hull is a polytope with $O(K)$ facets, the complexity of finding the convex hull is only $O(nK)$ [52].

The above gives rise to the following three-stage algorithm which we call *Ideal Mixed-SCORE*. Input: Ω. Output: π_i, $1 \le i \le n$.

- *SCORE step*. Obtain $(\lambda_1, \xi_1), \ldots, (\lambda_K, \xi_K)$ and the matrix R by (1.8).
- *Vertex Hunting (VH) step*. Run a convex hull algorithm on rows of R. Denote vertices of the obtained convex hull by $v_1, v_2, \ldots, v_K$.
- *Membership Reconstruction (MR) step*. Obtain b_1 by Lemma 1.1. For each $1 \le i \le n$, first, use r_i and $v_1, v_2, \ldots, v_K$ to obtain w_i, according to (1.10); next, use w_i and b_1 to obtain π_i, according to $w_i \propto (b_1 \circ \pi_i)$.

The following theorem is proved in the appendix:

THEOREM 1.1 (Ideal Mixed-SCORE). *Fix $K > 1$ and $n > 1$. Consider a $DCMM_n(K, P, \Theta, \Pi)$ where $P(\Pi'\Theta^2\Pi)$ is an irreducible matrix and where each community has at least one pure node (i.e., the set $\{1 \le i \le n : \pi_i(k) = 1\}$ is non-empty for all $1 \le k \le K$). Despite that Ξ may not be uniquely defined, b_1 and $\{w_i\}_{i=1}^n$ are uniquely defined and the Ideal Mixed-SCORE exactly recovers the membership matrix Π.*

Remark *(Why the simplex structure is non-trivial)*. In the special case where $\theta(1) = \ldots = \theta(n) = \sqrt{\alpha_n}$ (i.e., MMSB), (1.7) gives rise to a simplex structure directly: Let $\mathcal{S}_0 \subset \mathbb{R}^K$ be the simplex whose vertices are the K rows of the matrix $\sqrt{\alpha_n}B$. (i) For a pure node i, row i of Ξ falls on one vertex of $\mathcal{S}_0$, and for a mixed node i, row i of Ξ falls in the interior of $\mathcal{S}_0$. (ii) Each row of Ξ is a convex combination of the K vertices with weights π_i. This low-dimensional geometrical structure can be used to estimate Π under MMSB. However, for general DCMM, the situation is much more complicated:

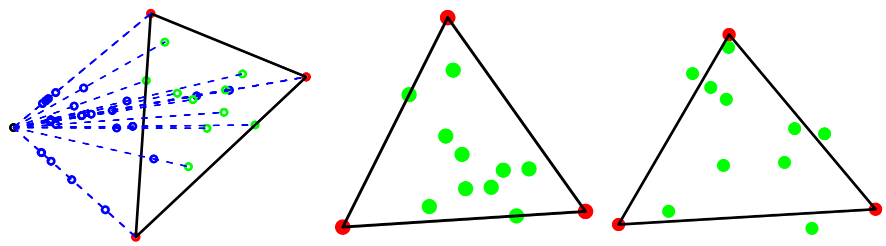

Fig 1. *Why obtaining the simplex is non-trivial ($K = 3$, simulated data). Left: rows of Ξ (blue points). The point cloud are contained in a simplicial cone, and it desires to normalize the points to have a simplex. Middle: rows of R (red: pure nodes; green: mixed nodes). It suggests that the SCORE normalization successfully produces a simplex. Middle: rows of Ξ normalized by row-wise ℓ^1-norm (for visualization, we have projected these points to $\mathbb{R}^2$). It suggests that this normalization fails to produce a simplex.*

- The rows of Ξ are no longer contained in a simplex. Instead, they are contained in a simplicial cone, where a pure row falls in one supporting ray and a mixed row falls in the interior. Note that it is much harder to identify the supporting rays of a simplicial cone from the point cloud than to retrieve the vertices of a simplex, especially in the presence of strong noise. It is therefore desirable to use a proper point-wise normalization to turn the simplicial cone into a simplex.
- In community detection literature, there are multiple ways to normalize Ξ to remove the effect of degree heterogeneity (see the appendix of [35] for the family of scaling invariant mappings). Unfortunately, most of them fail to produce a simplex structure when mixed memberships present. For example, a popular approach is to normalize each row of Ξ by its ℓ^q-norm, but this does not yield a simplex for any $q > 0$. The reason is that Ξ is *not* a non-negative matrix: Unless restricted to the positive orthant, the sphere of an ℓ^1-ball is non-convex. Fortunately, the SCORE normalization in (1.8) works well for our cases.
- Given a simplex structure, the estimates of π_i are obtained by expressing each point as a *weighted* linear combination of the simplex vertices. In MMSB, this weight vector equals to π_i. However, in DCMM, the weight vector usually is not π_i, and the relationship between two vectors are non-obvious, which we need to figure out carefully.

1.3. *Mixed-SCORE and the Vertex Hunting problem.* We now extend the idea to the real case. Let $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_K$ be the K largest (in magnitude) eigenvalues of A, and let $\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_K$ be the corresponding eigenvectors. Fixing a threshold $T > 0$, let $\hat{R} = [\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n]'$ be the empirical counterpart of R such that for $1 \leq i \leq n$ and $1 \leq k \leq K-1$,

$$\hat{R}(i,k) = \mathrm{sign}(\hat{\xi}_{k+1}(i)/\hat{\xi}_1(i)) \cdot \min\{|\hat{\xi}_{k+1}(i)/\hat{\xi}_1(i)|,\ T\}. \tag{1.11}$$

The following algorithm, which we call *Mixed-SCORE*, is a natural extension of the Ideal Mixed-SCORE. Input: A, K. Output: $\hat{\pi}_i$, $1 \leq i \leq n$.

- *SCORE step.* Obtain $(\hat{\lambda}_1, \hat{\xi}_1), \ldots, (\hat{\lambda}_K, \hat{\xi}_K)$ and $\hat{R} = [\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n]'$.
- *Vertex Hunting (VH) step.* Use $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n$ to estimate the vertices of Ideal Simplex. Denote the estimated vertices by $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_K$.
- *Membership Reconstruction (MR) step.* Obtain an estimate of b_1 by

  $$\hat{b}_1(k) = [\hat{\lambda}_1 + \hat{v}_k' \mathrm{diag}(\hat{\lambda}_2, \ldots, \hat{\lambda}_K)\hat{v}_k]^{-1/2}, \qquad 1 \leq k \leq K. \tag{1.12}$$

  For each $1 \leq i \leq n$, let $\hat{w}_i \in \mathbb{R}^K$ be the unique solution of the linear equations: $\hat{r}_i = \sum_{k=1}^K \hat{w}_i(k)\hat{v}_k$, $\sum_{k=1}^K \hat{w}_i(k) = 1$. Define a vector $\hat{\pi}_i^* \in \mathbb{R}^K$ by $\hat{\pi}_i^*(k) = \max\{0, \hat{w}_i(k)/\hat{b}_1(k)\}$, $1 \leq k \leq K$. Estimate π_i by $\hat{\pi}_i = \hat{\pi}_i^*/\|\hat{\pi}_i^*\|_1$, $1 \leq i \leq n$.

Here, Steps 1 and 3 are straightforward extensions of Steps 1 and 3 in Ideal Mixed-SCORE, respectively. The main challenge is how to extend Step 2 (i.e., Vertex Hunting) of Ideal Mixed-SCORE: in the point cloud formed by $\{\hat{r}_i\}_{i=1}^n$ in $\mathbb{R}^{K-1}$, the Ideal Simplex is blurred and is not directly observable, so we can not directly use a convex hull algorithm as before.

The point is illustrated in Figure 2, where the data is generated according to a DCMM with $(n, K) = (500, 3)$, $P \in \mathbb{R}^{3,3}$ has unit diagonals and 0.3 on all off-diagonals, $\{\theta^{-1}(i)\} \overset{iid}{\sim} \mathrm{Unif}[1,5]$. Among all nodes, 300 are pure nodes, with 100 in each community, and 200 are mixed nodes evenly distributed in 4 groups, where the PMFs equal to $(0.8, 0.2, 0.0)$, $(0.0, 0.2, 0.8)$, $(0.2, 0.4, 0.4)$ and $(1/3, 1/3, 1/3)$, in each of the four groups, respectively.

The VH step can be viewed as a "plug-in" step: for any VH approach that produces a good estimate of the vertices, we can plug it into our procedure and result in a different version of Mixed-SCORE. Of course, as long as the vertices estimation are sufficiently accurate, different versions of Mixed-SCORE give more or less similar membership estimations.

DEFINITION 1.2. *We call an algorithm a Vertex Hunting (VH) algorithm if it does the following job: Given input $X_1, X_2, \ldots, X_n \in \mathbb{R}^p$, where each X_i is a perturbation of $X_i^* \in \mathbb{R}^p$ and $X_1^*, X_2^*, \ldots, X_n^*$ satisfy that (i) they are contained in a simplex $\mathcal{S}(V_1, V_2, \ldots, V_K)$ and (ii) there exists an X_i^* located at each vertex V_k, the algorithm outputs estimated vertices $\hat{V}_1, \hat{V}_2, \ldots, \hat{V}_K$.*

For optimality, it is desirable to have VH algorithms that are efficient.

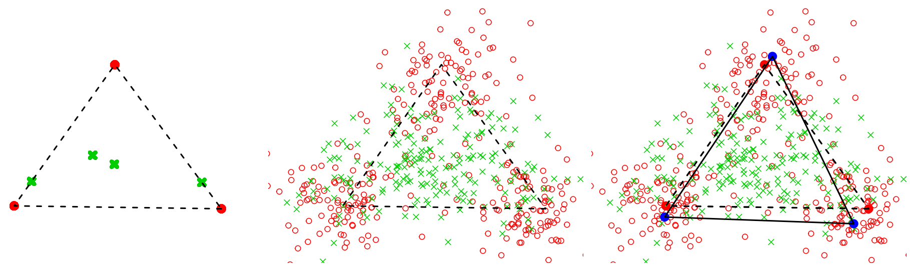

FIG 2. *Left: rows of R (many rows are equal so a point may represent many rows). Middle: each point is a row of $\hat{R}$. Right: same as the middle panel except that a triangle (solid blue) estimated by the Vertex Hunting algorithm is added. In all panels, dashed triangle is the Ideal Simplex, and red/green points correspond to pure/mixed nodes respectively.*

DEFINITION 1.3 (Efficient Vertex Hunting). *A Vertex Hunting algorithm is efficient if it satisfies* $\max_{1\leq k\leq K} \|\hat{V}_k - V_k\| \leq C \max_{1\leq i\leq n} \|X_i - X_i^*\|$.

In Section 3, we discuss several VH approaches, some are adapted from existing literature and some are new. In particular, we show three of these approaches are efficient.

The computing cost of Mixed-SCORE mainly comes from PCA and Vertex Hunting. PCA is quite manageable even for large matrices. The VH in Mixed-SCORE is operated in low-dimension ($p = K - 1$) and so is fast. As for complexity, PCA has a complexity of $O(n^2K)$. The complexity of VH depends on which algorithm we plug in; for the four VH algorithms to be discussed in Section 3, their complexity is summarized in Table 3. Among these algorithms, SP has the lowest complexity, which is $O(nK^2)$. As a result, the total complexity of Mixed-SCORE-SP is $O(n^2K)$.

1.4. *Optimality and a preview of the rate of convergence.* The formal statement of the rate is relatively long, and so is deferred to Section 2. In this section, we consider the special case where we only have a *mild* degree heterogeneity in the sense that $\theta_{max} \leq C\theta_{min}$, with θ_{min} and θ_{max} being the minimum and maximum values of $\{\theta_1, \theta_2, \ldots, \theta_n\}$, and use it for a preview of our results on the convergence rate.

Let β_n be the smallest eigenvalue of P in magnitude. Since our main interest is to estimate $\pi_1, \pi_2, \ldots, \pi_n$. which are PMFs, it is natural to measure the errors by the ℓ^1-loss. In Section 2, provided with some mild conditions, especially the VH step is efficient, then

$$\max_{1\leq i\leq n} \|\hat{\pi}_i - \pi_i\|_1 \leq \frac{K\sqrt{K\log(n)}}{\sqrt{n\bar{\theta}^2\beta_n^2}}.$$

If K is fixed or grows with n slowly enough (e.g., $K = \log(n)$), we can ignore the factor $K\sqrt{K\log(n)}$, so the rate of the Mixed-SCORE is $O((n\bar{\theta}^2\beta_n^2)^{-1/2})$, up to a multi-$\log(n)$ factor. This is in fact the optimal rate compared to the lower bound for MMSB, as proved by [36]. As a result, the Mixed-SCORE is rate optimal when $\theta_{max} \leq C\theta_{min}$ and K is either fixed or grows with n slowly enough. When K grows relatively fast with n (e.g., $K = n^a$ for some constant $a > 0$), it is unclear whether the Mixed-SCORE continues to be rate optimal. However, the rate of the Mixed-SCORE is much faster than those of existing approaches; see Table 2 for details.

Note that our general results (see Section 2) allow for *severe degree heterogeneity*, where the order of $\theta_{max}/\theta_{min}$ can get very close to that of $\sqrt{n}$.

1.5. *Comparison with other membership estimation approaches.* Several existing methods for mixed-membership estimation are designed for the MMSB [3], such as the Bayesian approach by Airoldi *et al.* [3] and the tensor approach by Anandkumar *et al.* [4].

- These methods only apply to the cases where θ_i's are equal, but we are interested in the realistic settings with severe degree heterogeneity.
- They need to assume that π_i's are *iid* drawn from a Dirichlet distribution, but how to validate this assumption in real networks is unclear.

Additionally, these methods are computationally much more expensive.

The OCCAM [60] is an interesting spectral approach. We note that both Mixed-SCORE and OCCAM use the idea of Jin's SCORE [35] by applying a post-PCA row-wise normalization. However, Mixed-SCORE has a major innovation that was not seen in OCCAM: Mixed-SCORE discovers a simplex structure and uses it for vertex hunting.

- Not aware of the simplex structure, OCCAM in fact uses the k-median algorithm for vertex hunting (in our language). An obvious *challenge* it faces is that, when there are too many mixed-nodes, then the estimated vertices (in our language) will be heavily biased towards the interior of the simplex. Hence, OCCAM only works when the fraction of mixed modes is relatively small. Such a condition is implicitly contained in their Assumption-B. In comparison, Mixed-SCORE allows for as many as $(n - K)$ pure nodes (see Theorem 2.2).
- The Mixed-SCORE also has a better rate than OCCAM; see Table 2.

After the first manuscript of our work was submitted in 2016, a few other methods were proposed for mixed membership estimation, including SPACL [46] and SPOC [50]. They are designed for MMSB and cannot accommodate degree heterogeneity. These works recognized a simplex structure associated

TABLE 2

Comparison of error rates. (First, in the rate for OCCAM, α_0 can be arbitrarily close to 0*. Second, in presenting the rate of SPACL, we replace their parameter ν by its lower bound K to give this method more favor. Third, the rates are based on different loss functions, which sometimes requires an adjustment of $\sqrt{K}$ in the comparison.)*

	θ_i	π_i	allowing $\beta_n{=}o(1)$	allowing $K{\gg}\log(n)$	Dependence on $(n, \bar{\theta}, \beta_n)$	Dependence on K
Ours	non-random, arbitrary	non-random, arbitrary	yes	yes	$(n\bar{\theta}^2\beta_n^2)^{-1/2}$	$K\sqrt{K}$
OCCAM [60]	*iid* from a distribution	*iid* from a distribution	no	no	$(n^{1-\alpha_0}\bar{\theta}^2)^{-1/5}$	–
SPACL [46]	all equal to $\bar{\theta}$	*iid* from Dirichlet	yes	yes	$(n\bar{\theta}^2\beta_n^2)^{-1/2}$	$K^6\sqrt{K}$
SPOC [50]	all equal to $\bar{\theta}$	*iid* from a distribution	no	yes	$(n\bar{\theta}^4)^{-1/2}$	K^2

with the unnormalized eigenvectors. Their simplex structure is less sophisticated than ours (see the remark in the end of Section 1.2).

We compare the rate of Mixed-SCORE with those of [46, 50, 60]; see Table 2. Since [46, 50] allow no degree heterogeneity, for a fair comparison, we state our error rate in the special case of $\theta_{\max} \leq C\theta_{\min}$ (but our theory applies to much more general cases; see Section 2). When K is fixed, Mixed-SCORE is rate optimal, as implied by the lower bound argument in [36]. In comparison, the approaches by Zhang et al. [60] and Panov et al. [50] are not rate optimal and do not cover the case of $\beta_n = o(1)$. Mao et al. [46] is in fact rate optimal when K is fixed, but the rate is slower than that of Mixed-SCORE when K grows with n; recall that β_n is the smallest eigenvalue (in magnitude) of P. In summary,

- Our theoretical results allow K, θ_i's, and β_n to vary with n.
- When K is finite, our error rate is already minimax optimal.
- When K grows with n, yet the optimal rate is unknown, our error rate is faster than that of all other approaches.

Most importantly, Mixed-SCORE allows severe degree heterogeneity where the order $\theta_{max}/\theta_{min}$ can get very close to that of $\sqrt{n}$, but [46, 50] have been focused on the case of $\theta_{max}/\theta_{min} = 1$.

1.6. *Comparison with the problem of community detection.* Mixed membership estimation is connected to community detection [17, 35, 53, 54, 61] and overlapping community detection [29, 42, 6, 43], but is different in important ways. Statistically, the community detection is a clustering problem, where π_i's take only finitely many possible values. Hence, it is less challenging than membership estimation. Indeed, community detection can achieve exponential error rates [23] while membership estimation can only achieve polynomial rates [36]. The two problems also have different practical meanings. The membership estimation provides richer information of nodes; for example, it allows us to rank the relevance of nodes to a particular community, while the (overlapping) community detection cannot.

We illustrate with a simple example. Consider an MMSB model, where $K = 2$, $\pi_i \overset{iid}{\sim} \text{Dirichlet}(\alpha_0)$, and for $1 \leq i \neq j \leq n$,

$$P(A(i,j) = 1) = n^{-1} \cdot \pi_i' \begin{bmatrix} a & b \\ b & a \end{bmatrix} \pi_j.$$

As $n \to \infty$, α_0 is fixed but (a, b) can change with n. The parametrization here is similar to those in many literatures of stochastic block model, so that it facilitates a convenient comparison. Translating it to our parametrization yields $\bar{\theta} \asymp n^{-1/2}\sqrt{a}$ and $\beta_n \asymp (a - b)/a$. Hence, the rate of convergence of Mixed-SCORE is $O(T^{-1/2}\sqrt{\log(n)})$, where $T = (a - b)^2/a$. In comparison, when π_i's are all degenerate and the goal is community detection, the error rate is $\exp(-O(T))$ [59]. Compared with community detection, membership estimation is a more difficult task, and so the rate is slower.

1.7. *Content and notations.* The remaining of this paper is organized as follows. Section 2 presents the main theoretical results, including a row-wise large-deviation bound for eigenvectors and the rate of convergence of Mixed-SCORE. Section 3 studies different Vertex Hunting methods, especially the new VH algorithm in Section 3.5. In Section 4, we apply Mixed-SCORE to all data sets in Table 1 and interpret the results. In Section 6, we prove main theorems. Section 5 contains simulations and Section 7 contains discussions. Proofs of secondary results are relegated to the appendix.

For any vector x, $\|x\|_q$ denotes the ℓ_q-norm, $q > 0$. The subscript is dropped for simplicity if $q = 2$. For any matrix M, $\|M\|$ denotes the spectral norm, $\|M\|_F$ denotes the Frobenius norm, and $\|M\|_1$ denotes the matrix ℓ_1-norm. We use C to denote a generic positive constant that may vary from occurrence to occurrence. For two positive sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$, we say $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$, and we say $a_n \asymp b_n$ if there is a constant $C > 1$ such that $b_n/C \leq a_n \leq Cb_n$ for sufficiently large n.

2. Main results. Consider a sequence of models $DCMM_n(K, P, \Theta, \Pi)$ indexed by n, where (K, P, Θ, Π) may change with n. Recall that $\{\theta(i)\}_{i=1}^n$ are the degree heterogeneity parameters. Let $\theta_{\max} = \max_{1 \leq i \leq n}\{\theta(i)\}$, $\theta_{\min} =$

$\min_{1\le i\le n}\{\theta(i)\}$, $\bar\theta = (1/n)\sum_{i=1}^n \theta(i)$, and $\bar\theta_* = [(1/n)\sum_{i=1}^n \theta^2(i)]^{1/2}$. The following quantity is closely related to our error rate:

$$err_n = err_n(\Theta) = [(\theta_{\max}^{3/2}\bar\theta^{3/2})/(\theta_{\min}\bar\theta_*^2)]\cdot\sqrt{\log(n)/(n\bar\theta^2)}.$$

We assume, as $n\to\infty$,

$$\theta_{\max}\le C, \qquad \text{and} \qquad err_n\to 0. \tag{2.13}$$

Introduce the matrix $G = K\|\theta\|^{-2}(\Pi'\Theta^2\Pi)\in\mathbb{R}^{K\times K}$. We assume

$$\|P\|_{\max}\le C, \qquad \|G\|\le C, \qquad \text{and} \qquad \|G^{-1}\|\le C. \tag{2.14}$$

Denote by $\lambda_k(PG)$ the k-th largest right eigenvalue of PG, and by $\eta_k\in\mathbb{R}^K$ the associated right eigenvector, $1\le k\le K$. For a constant $c_1>0$ and a sequence $\{\beta_n\}_{n=1}^\infty$ such that $\beta_n\le 1$, we assume

$$|\lambda_2(PG)|\le(1-c_1)\lambda_1(PG), \quad \text{and} \quad c_1\beta_n\le|\lambda_K(PG)|\le|\lambda_2(PG)|\le c_1^{-1}\beta_n. \tag{2.15}$$

We also assume η_1, the first (unit-norm) right singular vector of PG, satisfies that

$$\min_{1\le k\le K}\eta_1(k)>0, \qquad \text{and} \qquad \frac{\max_{1\le k\le K}\eta_1(k)}{\min_{1\le k\le K}\eta_1(k)}\le C. \tag{2.16}$$

We explain why all these conditions are mild. Consider Condition (2.13) first. Note that when $\theta_{\max}\le C\theta_{\min}$, $err_n\asymp\sqrt{\log(n)/(n\bar\theta^2)}$, where $n\bar\theta^2$ is the order of the expected average node degree. Hence, $err_n\to 0$ only requires that the average node degree grows to infinity at a rate faster than $\log(n)$, allowing for sparse networks. Consider Condition (2.14) next. The matrix G captures the distribution of degrees across communities. It is illuminating to consider a case where all nodes are pure so that $G = \|\theta\|^{-2}\cdot\mathrm{diag}(K\|\theta^{(1)}\|^2,\ldots,K\|\theta^{(K)}\|^2)$, where $\|\theta^{(k)}\|^2=\sum_{i\in\mathcal{C}_k}\theta_i^2$; then, the above condition on G reduces to that $\max_k\|\theta^{(k)}\|^2\le C\min_k\|\theta^{(k)}\|^2$. This is also a mild condition. Consider Condition (2.15) next. The first inequality is a mild eigen-gap condition. In the second inequality, β_n captures the 'distinction' between communities, and it will enter the error bound. Here, we assume $\lambda_2,\ldots,\lambda_K$ are at the same order. This is only for convenience and can be relaxed to the case where $\lambda_2,\ldots,\lambda_K$ form several groups such that eigenvalues in the same group are at the same order. Last, consider Condition (2.16). The condition is automatically satisfied in either of the following cases: As $n\to\infty$, (a) all entries of P are lower bounded by a constant, (b) K is fixed and P tends to a fixed irreducible matrix P_0, (c) K is fixed and G tends to a fixed irreducible matrix G_0. See Section A.2 in the appendix.

Below, first, we present a key technical result about the row-wise deviation of the matrix $\hat R$, then we present the rate of convergence of Mixed-SCORE.

2.1. *A row-wise large-deviation bound for $\hat{R}$.* At the heart of the analysis of Mixed-SCORE is the study of the matrix $\hat{R}$ defined in (1.11). Its population counterpart R is defined in (1.8). We have the following theorem:

THEOREM 2.1. *Consider a sequence of $DCMM_n(K, P, \Theta, \Pi)$, where (2.13)-(2.16) hold and there is at least one pure node for each of the K communities. With probability $1-o(n^{-3})$, there exists an orthogonal matrix $H \in \mathbb{R}^{K-1,K-1}$ such that*

$$\max_{1\leq i\leq n} \|H\hat{r}_i - r_i\| \leq CK^{3/2}\beta_n^{-1} err_n.$$

If, additionally, $\theta_{\max} \leq C\theta_{\min}$, then with probability $1 - o(n^{-3})$,

$$\max_{1\leq i\leq n} \|H\hat{r}_i - r_i\| \leq \frac{CK\sqrt{K\log(n)}}{\sqrt{n\bar{\theta}^2\beta_n^2}}.$$

The proof of Theorem 2.1 hinges on a row-wise large deviation bound for the eigenvectors of the adjacency matrix, which is of independent interest:

LEMMA 2.1. *Under conditions of Theorem 2.1, with probability $1-o(n^{-3})$, there exist $\omega \in \{\pm 1\}$ and an orthogonal matrix $X \in \mathbb{R}^{K-1,K-1}$ (both ω and X depend on A and are stochastic) such that:*

- $\|\omega\hat{\xi}_1 - \xi_1\| \leq C\|\theta\|^{-2}K\sqrt{\theta_{\max}\|\theta\|_1}$.
- $\|\hat{\Xi}_0 X - \Xi_0\|_F \leq C\beta_n^{-1}\|\theta\|^{-2}K^{3/2}\sqrt{\theta_{\max}\|\theta\|_1}$.
- $\|\omega\hat{\xi}_1 - \xi_1\|_\infty \leq C\|\theta\|^{-3}\theta_{\max}^{3/2}K\sqrt{\|\theta\|_1\log(n)}$.
- $\max_{1\leq i\leq n} \|X'\hat{\Xi}_{0,i} - \Xi_{0,i}\| \leq C\beta_n^{-1}\|\theta\|^{-3}\theta_{\max}^{3/2}K^{3/2}\sqrt{\|\theta\|_1\log(n)}$.

If $\beta_n = o(1)$, then the factor K in the bounds for $\|\omega\hat{\xi}_1 - \xi_1\|$ and $\|\omega\hat{\xi}_1 - \xi_1\|_\infty$ can be removed.

Remark. It is well-known that obtaining row-wise deviation bounds for eigenvectors of a random matrix is much more challenging than obtaining ℓ^2-norm deviation bounds. In network analysis, there are only few row-wise deviation results [1, 22, 46, 50], and these works only study models without degree heterogeneity. However, our results allow for severe degree heterogeneity. With degree heterogeneity, not only the conclusion changes, but also the proofs are very different. In particular, some techniques for models without degree heterogeneity are hard to extend to the case with degree heterogeneity; they either fail to give a sharp rate (e.g., the approach in [50]) or become extremely tedious (e..g, the approach in [46]).

Remark. Our proof of Theorem 2.1 heavily uses the "leave-one-out" technique developed in Abbe *et al.* [1] for entry-wise eigenvector analysis.

2.2. *Rate of convergence of Mixed-SCORE.* Depending on which VH algorithm is used, Mixed-SCORE has many variants. In order for our results to be general, we introduce a notion of "efficiency" for VH algorithms.

Our main theorem applies to any VH algorithm that is efficient (recall that a Vertex Hunting algorithm is efficient if it satisfies $\max_{1\leq k\leq K}\|\hat{V}_k - V_k\| \leq C\max_{1\leq i\leq n}\|X_i - X_i^*\|$). Among the existing VH algorithms, the successive projection [5] and archetypal analysis [19] can be proved to be efficient. In Section 3, we show that the new VH algorithm in Section 3.5 also satisfies this requirement.

We are ready to present the main theorem:

THEOREM 2.2 (Estimation error of Mixed-SCORE). *Consider a sequence of $DCMM_n(K, P, \Theta, \Pi)$, where (2.13)-(2.16) hold and there is at least one pure node for each of the K communities. Suppose there is an efficient Vertex Hunting algorithm available. We apply Mixed-SCORE with $T = \sqrt{\log(n)}$ in (1.8) and plug in this Vertex Hunting algorithm. As $n \to \infty$, with probability $1 - o(n^{-3})$, the Mixed-SCORE estimates $\hat{\pi}_i$ of π_i, $i = 1, 2, \ldots, n$, satisfy*

$$\max_{1\leq i\leq n}\|\hat{\pi}_i - \pi_i\|_1 \leq CK^{3/2}\beta_n^{-1}err_n.$$

If, additionally, $\theta_{\max} \leq C\theta_{\min}$, then, with probability $1 - o(n^{-3})$,

$$\max_{1\leq i\leq n}\|\hat{\pi}_i - \pi_i\|_1 \leq \frac{CK\sqrt{K\log(n)}}{\sqrt{n\bar{\theta}^2\beta_n^2}}.$$

3. Efficient Vertex Hunting. In Theorem 2.2, we require the VH algorithm employed in the Mixed-SCORE to be efficient. We now show that several VH algorithms, including the new ones to be introduced in Sections 3.4-3.5, are efficient. Below, first, in Section 3.1, we review existing methods in the literature and explain why we need to propose new VH algorithms. Next, in Sections 3.2-3.5, we study four different VH algorithms, two of which are existing and two of which are new, and show that they are all efficient under mild conditions. In Section 3.6, we compare these VH approaches and the resultant Mixed-SCORE of these approaches.

3.1. *Overview.* The VH problem is equivalent to the constrained linear unmixing (also called endmember extraction) problem in the literature of hyperspectral image unmixing [10]. Below, we review three classes of linear unmixing methods: (1)-(3).

(1). Optimization methods. Such methods treat $V_1, \ldots, V_K$ as parameters to optimize and solve a constrained optimization. (a) N-FINDER [58]. It

maximizes the volume of the simplex, subject to the constraint that each vertex of the simplex is placed on a data point. Several algorithms were proposed to solve N-FINDER, such as the SGA algorithm [16]. (b) Minimum-volume transformation (MVT) [18]. It minimizes the volume of the simplex, subject to that the simplex contains every data point. MVT is a popular approach in hyperspectral unmixing, with many algorithms and variants, such as MVSA[45], SISAL [9], NMF-MNT [55], and MVES [15]. (c) Archetypal analysis [19]. It minimizes the sum of squared Euclidean distance from data points to the simplex, subject to that each vertex is placed in the convex hull of data cloud. This method also has a few variants, such as [33].

(2). Greedy algorithms. Unlike optimization methods which optimize over all vertices together, these algorithms decide one vertex at a time in a greedy fashion. (a) Successive projection [5]: Given k vertices that are already determined, this algorithm sets the $(k+1)$-th vertex as the data point with the largest Euclidean distance to the linear span of previous k vertices. In linear unmixing, this method is also commonly called the VGA algorithm [48]. (b) Iterative error analysis (IEA) [49]. At each iteration, this algorithm computes the "error" of approximating each data point by a convex combination of previously determined vertices, and sets the next vertex as the data point with the largest error.

(3). Purity calculation. These methods compute a "purity score" for each data point, which roughly measures how close it is to the nearest vertex of the convex hull of data cloud. Then, data points with highest scores are output as vertices or used as candidates of vertices for other algorithms. The most commonly used purity score is the pixel purity index (PPI) [13], which calculates the frequency of a data point to become the extreme point under a large number of random projections.

Classes (1)-(2) are more popular, and in comparison, approaches in (2) typically have low computational complexities, but their practical performances are less satisfactory than those in (1) (possible reason: once a highly biased vertex is determined in a greedy algorithm, it can never be corrected). This trade-off motivates us to consider methods both in (1) and (2). In Section 3.2, we study a greedy algorithm, and in Section 3.3, we study an optimization approach. We show both are efficient under mild conditions.

While these approaches may be impressive in theory, their numerical performance is found to be relatively unsatisfactory. In practice, it is desirable to have a VH algorithm that performs well under the presence of strong noise and is also robust to outliers, but unfortunately, existing VH algorithms perform unsatisfactorily in such situations. For example, in optimization approaches, a constraint usually involves *every* data point, and a single outlier can significantly affect the solution. In greedy algorithm approaches, we select the "extreme" points at each iteration and the "extreme" points are sensitive to outliers.

These motivate a new VH approach to be introduced in Section 3.4-3.5. Section 3.4 discusses a special case of the algorithm which is easier to digest, and Section 3.5 introduce the algorithm in general forms. We shall see that the new algorithm not only has much better numerical performance but also leads to a faster rate of convergence in some settings.

Recall that in Mixed-SCORE we apply a VH algorithm to rows of $\hat{R}$. Comparing Theorem 2.1 with Definitions 1.2-1.3, we see that it corresponds to a VH setting where $p = K-1$, $X_i^* = H^{-1}r_i$, and $X_i = \hat{r}_i$. Then, the efficiency condition translates to $\max_{1\leq k\leq K}\|H\hat{v}_k - v_k\| \leq C\max_{1\leq i\leq n}\|H\hat{r}_i - r_i\|$.

3.2. *VH algorithm 1: Successive projection (SP).* The successive projection [5, 48] is a greedy algorithm. Its main idea is to successively project data into the orthogonal space of previously determined vertices and to decide the next vertex by identifying the extreme point after projection.

Successive projection (input: K, $\hat{r}_1, \ldots, \hat{r}_n$).

- Initialize $Y_i = (1, \hat{r}_i')' \in \mathbb{R}^K$, for $1 \leq i \leq n$.
- At iteration $k = 1, 2, \ldots, K$: Find $i_k = \text{argmax}_{1\leq i\leq n}\|Y_i\|$ and let $u_k = Y_{i_k}/\|Y_{i_k}\|$. Set the k-th estimated vertex as $\hat{v}_k = \hat{r}_{i_k}$. Project all data points by updating Y_i to $(1 - u_k u_k')Y_i$, for $1 \leq i \leq n$.
- Output $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_K$.

The error bound of the SP algorithm has been studied in [24, 25]. To apply their results to our setting, we only need to verify that the Ideal Simplex is not "ill-conditioned." The following lemma is proved in the appendix.

Lemma 3.1 (Efficiency of SP). *Suppose conditions of Theorem 2.2 hold. Let H be the same matrix as in Theorem 2.1. We apply the SP algorithm to rows of $\hat{R}$. With probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy that*

$$\max_{1\leq k\leq K}\|H\hat{v}_k - v_k\| \leq C\max_{1\leq i\leq n}\|H\hat{r}_i - r_i\|.$$

As a result, the SP algorithm is efficient.

For complexity, the main cost of each iteration is computing the projection of Y_i to $(1 - u_k u_k')Y_i$, for all i. It is seen that computing one such projection has a complexity of $O(K-1)$. Hence, the total complexity is $O(nK^2)$.

3.3. *VH algorithm 2: Archetypal analysis (AA).* The archetypal analysis has a few variants [19, 47, 33]. Here, we present the version in [33] for it has explicit theory. Let $\mathcal{D}(x, E)$ denotes the Euclidean distance from a point x to a set E (note: for any $x \in E$, $\mathcal{D}(x, E) = 0$), let $\mathcal{S}(v_1, \ldots, v_K)$ denote the simplex with $v_1, \ldots, v_K$ as vertices, and let $\mathcal{H}(\hat{r}_1, \ldots, \hat{r}_n)$ denote the convex hull of $\{\hat{r}_i\}_{i=1}^n$. For a tuning parameter $\delta > 0$, AA solves an optimization:

$$\begin{aligned} \min_{v_1,\ldots,v_K \in \mathbb{R}^{K-1}} \quad & \sum_{k=1}^{K} \mathcal{D}^2\big(v_k, \mathcal{H}(\hat{r}_1, \ldots, \hat{r}_n)\big), \\ \text{subject to} \quad & \mathcal{D}^2\big(\hat{r}_i, \mathcal{S}(v_1, \ldots, v_K)\big) \le \delta^2, \quad 1 \le i \le n. \end{aligned} \tag{3.17}$$

The constraint ensures that the simplex contains the majority of data points, and minimizing the objective prevents the simplex from being too "large."

Lemma 3.2 is adapted from [33, Theorem 1] and is proved in the appendix.

LEMMA 3.2 (Efficiency of AA). *Suppose conditions of Theorem 2.2 hold. Let H be the same matrix as in Theorem 2.1. Consider (3.17) with $\delta = C_* K^{3/2} \beta_n^{-1} err_n$, where $C_* > 0$ is a properly large constant. With probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy that*

$$\max_{1\le k\le K} \|H\hat{v}_k - v_k\| \le CK^3\sqrt{K} \max_{1 \le i \le n} \|H\hat{r}_i - r_i\|.$$

As a result, AA is efficient, provided that K is bounded as $n \to \infty$.

For computation, it is more convenient to consider the penalized form:

$$\min_{v_1,\ldots,v_K \in \mathbb{R}^{K-1}} \Big\{ \sum_{i=1}^{n} \mathcal{D}^2\big(\hat{r}_i, \mathcal{S}(v_1, \ldots, v_K)\big) + \lambda \sum_{k=1}^{K} \mathcal{D}^2\big(v_k, \mathcal{H}(\hat{r}_1, \ldots, \hat{r}_n)\big) \Big\}.$$

This is a nonconvex optimization, and [33] proposed a proximal alternating linearized minimization algorithm that guarantees to converge to a stationary point. The exact complexity of this algorithm is unknown.

3.4. *VH algorithm 3 (new): Combinatorial vertex search (CVS).* We introduce a new VH algorithm, which is also an optimization approach:

$$\begin{aligned} \min_{v_1,\ldots,v_K} \quad & \max_{1\le i\le n} \mathcal{D}\big(\hat{r}_i, \mathcal{S}(v_1, \ldots, v_K)\big), \\ \text{subject to} \quad & v_k \in \{\hat{r}_1, \ldots, \hat{r}_n\}, \qquad 1 \le i \le n, \\ & \mathcal{S}(v_1, \ldots, v_K) \text{ has a nonzero volume.} \end{aligned} \tag{3.18}$$

Here, $\mathcal{D}\big(\hat{r}_i, \mathcal{S}(v_1, \ldots, v_K)\big)$ is the Euclidean distance from $\hat{r}_i$ to the simplex $\mathcal{S}(v_1, \ldots, v_K)$. For $\hat{r}_i$ is in the interior of the simplex, this distance is zero by definition. As a result, this method searches for all non-degenerate simplexes each vertex of which is placed in a data point, and finds the simplex such that the maximum distance from any outside point to this simplex is minimized.

Compared with the optimization approach in Section 3.3, our method has no tuning parameter, and it is efficient even for a growing K:

LEMMA 3.3 (Efficiency of CVS). *Under conditions of Theorem 2.2, consider* (3.18)*. Let H be the same matrix as in Theorem 2.1. With probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy that*

$$\max_{1 \le k \le K} \|H\hat{v}_k - v_k\| \le C \max_{1 \le i \le n} \|H\hat{r}_i - r_i\|.$$

As a result, the CVS is efficient.

We solve (3.18) by a combinatorial search: The requirement $v_k \in \{\hat{r}_1, \ldots, \hat{r}_n\}$ implies that there are at most $\binom{n}{K} = O(n^K)$ possible simplexes. For each candidate simplex, we compute $\mathcal{D}\big(\hat{r}_i, \mathcal{S}(v_1, \ldots, v_K)\big)$ via a standard quadratic programming: minimize $\|\hat{r}_i - \sum_{1 \le k \le K} x_k v_k\|^2$, subject to that $0 \le x_k \le 1$ and $\sum_{k=1}^K x_k = 1$. The complexity of solving this quadratic programming is $O(K^3)$. It follows that the overall complexity of CVS is $O(K^3 n^{K+1})$. This is a polynomial-time algorithm when K is bounded.

Remark *(Smart CVS)*. We propose a better way to conduct the combinatorial search. The idea is to use a "purity" score to order the data points, so that high-purity points will be investigated first. We also progressively eliminate simplexes that are impossible to be the solution.

- Rank indices according to the descending order of $p_i \equiv \|\hat{r}_i - n^{-1}\sum_{j=1}^n \hat{r}_j\|$. Without loss of generality, we assume $p_1 \ge p_2 \ge \ldots \ge p_n$.
- Obtain an ordered collection of all K-out-of-n index tuples, where they are arranged in the natural order.[5] Denote this collection by $\mathcal{H}$.
- Initialize $d = \infty$ and $\hat{v}_1, \ldots, \hat{v}_K$ to be empty vectors. Run the following iteration until $\mathcal{H}$ is empty:
  - Let $(i_1, \ldots, i_K)$ be the first element in $\mathcal{H}$. Set $u_k = \hat{r}_{i_k}$, $1 \le k \le K$. Remove this element from $\mathcal{H}$. If $u_1, \ldots, u_K$ are affinely dependent, continue to next iteration.

[5]For example, for $(n, K) = (5, 3)$, the natural order is $(1, 2, 3)$, $(1, 2, 4)$, $(1, 2, 5)$, $(1, 3, 4)$, $(1, 3, 5)$, $(1, 4, 5)$, $(2, 3, 4)$, $(2, 3, 5)$,$(2, 4, 5)$, $(3, 4, 5)$. This order guarantees that the K-tuples formed by smaller indices appear earlier in the list. In *Matlab*, such an ordered collection can be conveniently obtained using the function `nchoosek`.

- Compute $y_i \equiv \mathcal{D}(\hat{r}_i, \mathcal{S}(u_1, \ldots, u_K))$, for $1 \leq i \leq n$. If $\max_i y_i < d$, update $d = \max_i y_i$ and $\hat{v}_k = u_k$, $1 \leq k \leq K$.
- For each element in $\mathcal{H}$, obtain the sum of y_i's of K indices in this element. If the sum equals to zero, remove this element from $\mathcal{H}$.[6]

• Output $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_K$.

SmartCVS gives the same output as CVS but is much faster, as the first a few iterations almost always remove a large number of candidate simplexes.

3.5. *VH algorithm 4 (new): Sketched vertex search (SVS).* Near the end of Section 3.1, we explain the challenges faced by existing VH approaches and why it is desirable to have a new algorithm. We now propose a two-stage new algorithm, which includes VH algorithm 3 as a special case.

The main idea is to first apply classical k-means to the point cloud and identify a few (but more than K) "local centers". Each "local center" is the average of many nearby points and is relatively robust to outliers. We then apply the CVS algorithm in Section 3.4 by treating these "local centers" as the data points. The set of "local centers" serves as a *sketch* of the original data cloud, so we name our method the *sketched vertex search*.

Sketched vertex search (input: K, a tuning integer $L \geq K$, $\hat{r}_1, \ldots, \hat{r}_n$).

- *Local clustering.* Apply the classical k-means algorithm to $\hat{r}_1, \cdots, \hat{r}_n$ assuming there are L clusters. Denote the centers of the clusters by $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L \in \mathbb{R}^{K-1}$.
- *Combinatorial vertex search.* For any K distinct indices $1 \leq j_1 < \ldots < j_K \leq L$, let $\mathcal{H}(\hat{m}_{j_1}, \ldots, \hat{m}_{j_K})$ be the convex hull of $\hat{m}_{j_1}, \ldots, \hat{m}_{j_K}$, and denote the maximal Euclidean distance between the convex hull and those cluster centers outside the convex hull by

$$d_L(j_1, \cdots, j_K) = \max_{1 \leq j \leq L} \text{distance}\big(\hat{m}_j,\ \mathcal{H}\{\hat{m}_{j_1}, \cdots, \hat{m}_{j_K}\}\big). \tag{3.19}$$

  Let $\hat{j}_1 < \ldots < \hat{j}_K$ be the indices such that

$$(\hat{j}_1, \hat{j}_2, \ldots, \hat{j}_K) = \text{argmin}_{\{1 \leq j_1 < j_2 < \ldots < j_K \leq L\}} d_L(j_1, j_2, \ldots, j_K).$$

  Output $\hat{v}_k = \hat{m}_{\hat{j}_k}$, $1 \leq k \leq K$.

In the special case of $L = n$, the local clustering is skipped, and it reduces to the CVS algorithm in Section 3.4.

[6]For an element $\mathcal{H}$, if the sum of y_i's equals to zero, it means all the K points are in the interior of $\mathcal{S}(u_1, \ldots, u_K)$. As a result, the whole simplex with these K points as vertices must be in the interior of $\mathcal{S}(u_1, \ldots, u_K)$, and it cannot be the solution

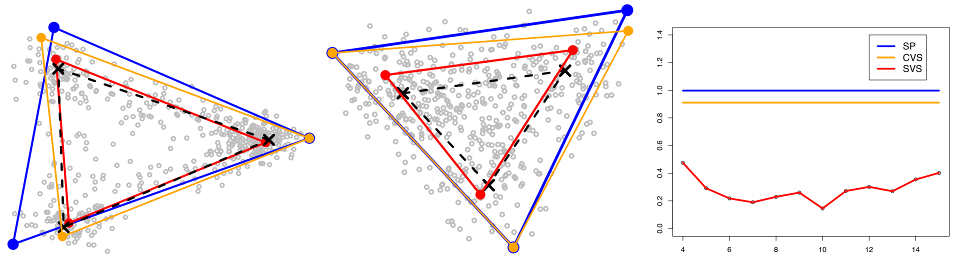

FIG 3. *Comparison of VH methods (black: truth; blue: SP; yellow: CVS; red: SVS). Left: The case of weak noise. CVS and SVS perform well, but SP performs less satisfactorily (possible reason: SP is a greedy algorithm). Middle: The case of strong noise. SVS performs well, but SP and CVS perform unsatisfactorily. This is because SVS is much less sensitive to outliers. Right: Robustness of SVS to the choice of L (y-axis is* $\max_k \|H\hat{v}_k - v_k\|$*).*

As for the rationale of this algorithm, we recognize that, under mild conditions, each vertex of the Ideal Simplex is surrounded by a cluster of points, where each point represents a row of $\hat{R}$ corresponding to a pure node; as a result, each vertex falls close to one of the "local centers". At the same time, the remaining "local centers" lie in the interior of the Ideal Simplex. Therefore, if we perform vertex hunting on these "local centers", the algorithm will pick up those "local centers" close to true vertices.

Figure 3 illustrates the advantage of SVS, especially under strong noise. The data are generate from a DCMM, where $(n, K) = (500, 3)$, P is a matrix whose diagonals are 1 and off-diagonals are 0.3. Each community has 50 pure nodes. For π_i's of the remaining 350 nodes, half of them are *iid* drawn from Dirichlet$(0.6, 0.2, 0.2)$, and half are *iid* drawn from Dirichlet$(0.3, 0.4, 0.3)$. We consider two cases: (a) Weak noise ($\theta_i \equiv 0.7$, and the network is denser) (b) Strong noise ($\theta_i \equiv 0.4$, and the network is sparser). In both cases, SVS has the best performance. Especially, in the strong noise case, SP and CVS perform unsatisfactorily due to that they are sensitive to outliers. We also vary the tuning parameter L in SVS and find that its performance is robust. All these observations can be justified in theory; see below.

While this algorithm has appealing practical performance, its theoretical analysis is complicated, due to the local clustering step. Although this step uses a clustering algorithm, the purpose is not clustering but noise reduction. The key of the analysis is to understand behavior of the "local centers", even when the data cloud has no clustering structure. Due to technical challenges, we only consider three settings:

- Setting 1: π_i's are *iid* sampled from a continuous distribution on the standard simplex.

- Setting 2: π_i's are fixed, but they form a few clusters (π_i's in the same cluster are close to each other but do not necessarily overlap).
- Setting 3: π_i's are fixed, but most of them are degenerate (i.e., most nodes are pure nodes).

Setting 1 is the seemingly difficult case for SVS, since the true π_i's have no clustering structure. We show that the algorithm indeed satisfies the efficiency condition, where the proof uses the Borel-Lebesgue covering theorem. Settings 2-3 are more favorable to SVS. We use them to demonstrate: the advantage of SVS under strong noise is not only a practical observation but also justifiable in theory.

We first discuss Setting 1. Let $e_1, e_2, \ldots, e_K$ be the standard basis vectors of $\mathbb{R}^K$, and let $\mathcal{S}_0 = \mathcal{S}_0(e_1, e_2, \ldots, e_K)$ be the simplex spanned by them. Fix a density g defined over $\mathcal{S}_0$ and let $\mathcal{R} = \{\pi \in \mathcal{S}_0 : g(\pi) > 0\}$ be the support of g. We suppose there is a constant $c_0 > 0$ such that

$$\mathcal{R} \text{ is an open subset of } \mathcal{S}_0, \text{ and distance}(e_k, \mathcal{R}) \geq c_0,\ 1 \leq k \leq K. \tag{3.20}$$

Let $\delta_v(\pi)$ denote the point mass at $\pi = v$. We invoke a random design model where π_i's are *iid* drawn from a mixture

$$f(\pi) = \sum_{k=1}^{K} \epsilon_k \cdot \delta_{e_k}(\pi) + \Big(1 - \sum_{k=1}^{K} \epsilon_k\Big) \cdot g(\pi), \tag{3.21}$$

where $\epsilon_k > 0$ are constants such that $\sum_{k=1}^K \epsilon_k < 1$.

Lemma 3.4 (Efficiency of SVS, Setting 1). *Suppose conditions of Theorem 2.2 hold. Additionally, suppose K is fixed and rows of Π are iid generated from (3.20)-(3.21). We apply the SVS algorithm to rows of $\hat{R}$ with an L that does not change with n. Then, there exists $L_0 = L_0(g, \epsilon_1, \ldots, \epsilon_K)$ such that, as long as $L \geq L_0$, with probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy*

$$\max_{1 \leq k \leq K} \|H\hat{v}_k - v_k\| \leq C \max_{1 \leq i \leq n} \|H\hat{r}_i - r_i\|.$$

As a result, the SVS algorithm is efficient.

Remark *(Why overshooting of L is not an issue)*. Lemma 3.4 suggests that the performance of SVS is robust to the choice of L in the sense that an overshooting of L has negligible effects. This has an intuitive explanation. As L increases, more local centers emerge. There are two most likely cases. First, a new local center appears in the interior of the simplex, while local centers close to true vertices are (almost) not affected. Since interior local centers will never be selected in the second step, the output of the algorithm is (almost) not changed. Second, a local center close to a true vertex splits into multiple local centers. The vertex search step may pick up one of these new local centers as a vertex. However, since all of these local centers are close to the original local center, the change on the solution is also small.

Next, we introduce Setting 2. Let $\mathcal{N}_k = \{1 \leq i \leq n : \pi_i(k) = 1\}$ be the set of pure nodes of community k, $1 \leq k \leq K$, and let $\mathcal{M} = \{1 \leq i \leq n : \max_{1\leq k\leq K} \pi_i(k) < 1\}$ be the set of all mixed nodes. We assume there are constants $c_1, c_2 \in (0,1)$ such that

$$\min_{1\leq k\leq K} |\mathcal{N}_k| \geq c_1 n, \qquad \min_{1\leq k\leq K} \sum_{i\in\mathcal{N}_k} \theta^2(i) \geq c_2\|\theta\|^2. \tag{3.22}$$

Furthermore, for a fixed integer $L_0 \geq 1$, we assume there is a partition of $\mathcal{M}$, $\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_{L_0}$, a set of PMF's $\gamma_1, \cdots, \gamma_{L_0}$, and constants $c_3, c_4 > 0$ such that (e_k: k-th standard basis vector of $\mathbb{R}^K$)

$$\Big\{ \min_{1\leq j\neq \ell\leq L_0} \|\gamma_j - \gamma_\ell\|, \min_{1\leq \ell\leq L_0, 1\leq k\leq K} \|\gamma_\ell - e_k\| \Big\} \geq c_3, \tag{3.23}$$

and for each $1 \leq \ell \leq L_0$ (note: err_n is the same as that in Section 2),

$$|\mathcal{M}_\ell| \geq c_4|\mathcal{M}| \geq n\beta_n^{-2} err_n^2, \qquad \max_{i\in\mathcal{M}_\ell} \|\pi_i - \gamma_\ell\| \leq 1/\log(n). \tag{3.24}$$

LEMMA 3.5 (Efficiency of SVS, Setting 2). *Suppose conditions of Theorem 2.2 hold. Additionally, suppose K is fixed and (Θ, Π) satisfy (3.22)-(3.24). For any integer $L \geq 1$, denote by $\epsilon_L(\hat{R})$ the sum of squared residuals of applying k-means to rows of $\hat{R}$ to get L clusters. We apply the SVS algorithm to rows of $\hat{R}$, with a data-drive choice of L:*

$$\hat{L}_n(A) = \min\{L \geq K+1 : \epsilon_L(\hat{R}) < \epsilon_{L-1}(\hat{R})/\log(\log(n))\}. \tag{3.25}$$

With probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy

$$\max_{1\leq k\leq K} \|H\hat{v}_k - v_k\| \leq C\Big(n^{-1}\sum_{i=1}^n \|H\hat{r}_i - r_i\|^2\Big)^{1/2} \leq C \max_{1\leq i\leq n} \|H\hat{r}_i - r_i\|. \tag{3.26}$$

As a result, the SVS algorithm is efficient.

Remark. In other applications, we may encounter VH settings where the noise $(X_i - X_i^*)$ may be *iid* subGaussian (e.g., [10]). In such settings, the rate of (3.26) can be improved by a $\sqrt{n}$ factor: $\max_{1\leq k\leq K} \|H\hat{v}_k - v_k\| \leq$

$C \cdot n^{-1/2} \cdot \max_{1\leq i\leq n} \|H\hat{r}_i - r_i\|$. Such a sharp rate is not available for the SP and AA approaches, and the reason is that, SVS uses a local clustering step which serves as noise reduction, which other VH approaches do not use. This shows SVS does have significant theoretical advantages.

Remark. The conditions (3.22)-(3.24) can be further relaxed. For example, in (3.24), we assume mixed π_i's in each true cluster are within a distance of $1/\log(n)$ to the center γ_ℓ. Here, $1/\log(n)$ can be relaxed to a constant $\tilde{c}_3$ that depends on c_3 in an implicit way.

Last, we consider Setting 3. Recall that $\mathcal{M}$ is the set of mixed nodes. We assume, for a sequence $\zeta_n \to 0$,

$$|\mathcal{M}| \leq n\zeta_n \qquad \min_{i\in\mathcal{M}, 1\leq k\leq K} \|\pi_i - e_k\| \geq c_5. \tag{3.27}$$

LEMMA 3.6 (Efficiency of SVS, Setting 3). *Suppose conditions of Theorem 2.2 hold. Additionally, suppose K is fixed and (Θ, Π) satisfy (3.22) and (3.27), for $\zeta_n \leq C\beta_n^{-1} err_n$. We apply the SVS algorithm to rows of $\hat{R}$, with $L = K$. With probability $1 - o(n^{-3})$, the estimated $\hat{v}_1, \ldots, \hat{v}_K$ satisfy*

$$\max_{1\leq k\leq K} \|H\hat{v}_k - v_k\| \leq C\Big(n^{-1}\sum_{i=1}^n \|H\hat{r}_i - r_i\|^2\Big)^{1/2} + C\beta_n^{-1} err_n.$$

Let F_n be the event that $\max_{1\leq i\leq n} \|H\hat{r}_i - r_i\| > \beta_n^{-1} err_n$. On the event F_n, the SVS algorithm is efficient.

We note that the event F_n^c has no impact on the error bound of Mixed-SCORE. In a remark in the proof of Lemma 3.6, we show that the conclusion of Theorem 2.2 still holds on the event F_n^c.

The conclusions in Lemmas 3.5-3.6 are stronger than the efficiency condition: The vertex hunting error is bounded by the root mean square error (RMSE) on $\hat{r}_i$'s, which is smaller than the maximum error. In fact, the large-deviation bound for the RMSE is smaller than the bound for the maximum error by at least a factor of $\sqrt{\log(n)}$. This advantage translates to a faster rate of convergence for the ℓ^2-error on estimating memberships. As a direct corollary of Theorem 2.2, by plugging in an efficient VH algorithm, the estimates $\hat{\pi}_i$ from Mixed-SCORE satisfy that

$$\frac{1}{n}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|^2 \leq CK^3\beta_n^{-2} err_n^2, \tag{3.28}$$

where err_n is the same as in (2.13). In contrast, if we plug in the SVS algorithm, we will obtain a faster rate of convergence for the ℓ^2-error. Introduce

$$err_n^* = err_n^*(\Theta) = [(\theta_{\max}^{1/2}\bar{\theta}^{3/2})/(\theta_{\min}\bar{\theta}_*)] \cdot (n\bar{\theta}^2)^{-1/2}.$$

We note that $err_n^* = err_n[\bar{\theta}_*/(\theta_{\max}\sqrt{\log(n)})] \leq err_n/\sqrt{\log(n)}$. The following theorem shows that the rate in (3.28) can be improved by at least a factor of $\log(n)$. For large-scale networks, the improvement is significant.

Theorem 3.1 (Faster rate of Mixed-SCORE in Settings 2 & 3). *Consider the setting of Lemma 3.5 or the setting of Lemma 3.6 for $\zeta_n \leq C\beta_n^{-1}err_n^*$. We apply Mixed-SCORE with $T = \sqrt{\log(n)}$ in (1.8) and plug in the SVS algorithm for vertex hunting, where L is chosen as in Lemmas 3.5-3.6. As $n \to \infty$, with probability $1 - o(n^{-3})$, the Mixed-SCORE estimates $\hat{\pi}_i$ of π_i, $i = 1, 2, \ldots, n$, satisfy*

$$\frac{1}{n}\sum_{i=1}^{n}\|\hat{\pi}_i - \pi_i\|^2 \leq CK^3\beta_n^{-2}(err_n^*)^2 \leq \frac{CK^3\beta_n^{-2}err_n^2}{\log(n)}.$$

Remark *(Variants of SVS)*. The idea of SVS can be combined with other VH algorithms. The local clustering step reduces the original VH problem to a new VH problem with L data points. This inspires us to replace the second step by any existing VH algorithms, such as the SP and AA in Sections 3.2-3.3 or the SmartCVS in Section 3.4. These are all variants of SVS.

Remark. Among the four VH approaches, AA has one tuning parameter δ_n but how to set it in practice (as in [33] where it was proposed) remains unknown. SVS has an tuning parameter L, and we have carefully discussed how to set it in Lemma 3.4-3.6 from a theoretical perspective. For practice, we find the following approach to setting L is numerically appealing. For each L, we compute $d_L(\hat{R}) = d_L(\hat{j}_1, \cdots, \hat{j}_K)$ as in (3.19). Let

$$\delta_L(\hat{R}) = \min_{\{j_1,\ldots,j_K\}:\text{a permutation of }\{1,\ldots,K\}} \max_{1\leq k\leq K}\{\|\hat{v}_{j_k}^{(L)} - \hat{v}_k^{(L-1)}\|\}.$$

This quantity tracks the change of estimated vertices when we increase the tuning parameter from $(L-1)$ to L. We select the L such that this change is relatively small:

$$\hat{L}_n^*(A) = \text{argmin}_{K+1\leq L\leq 3K}\{\delta_L(\hat{R})/(1 + d_L(\hat{R}))\}; \tag{3.29}$$

if there is a tie, pick the largest integer. Our numerical results suggest that the final result of Mixed-SOCRE is relatively insensitive to the choice of L. See for example Figure 3 and the discussion therein.

3.6. *Summary.* The VH is a problem of its own interest, and Mixed-SCORE is just one example where VH can be useful. In this section we briefly review the existing VH approaches, and analyze two of them: SP and AA. We also propose two new approaches: CVS, and SVS. Table 3 summarizes the pros and cons of different methods. Unfortunately, there is no available VH method that is uniformly good in every aspect. The advantages of the SVS approaches we propose including (a) the rate can be faster by at least a $\log(n)$ factor in the current setting, and can be faster by a $\sqrt{n}$ factor in some other VH settings, and (b) numerically it is fast in real time and provide satisfactory performance both for the weak noise and strong noise settings. The SVS also has some room to improve, especially in how to further relax the regularity conditions. We leave this to the future.

TABLE 3

Comparison of four VH approaches (the last two are new; CVS is a special case of SVS). For the vertex estimation error, the rate of SVS is faster than those of others by a $\log(n)$ factor, and is faster by a $\sqrt{n}$ factor when $(X_i - X_i^*)$ are iid (as in many applications where VH is desired; see the remark below Lemma 3.5). The SVS has a high complexity but fast in real time, as it uses the Llyod's algorithm.

Name	Efficient	Rate	Weak noise	Strong noise	Complexity	Real time
SP	yes	standard	good	rather unsatisfactory	$O(K^2n)$	fast
AA	yes	standard	good	unsatisfactory	unknown	relatively slow
CVS	yes	standard	good	unsatisfactory	$O(K^3n^{K+1})$	relatively slow
SVS	yes	**faster**	good	**satisfactory**	$O(n^{KL}+K^3L^{K+1})$	**fast**

4. Application to all network data sets in Table 1.

Let $\hat{\pi}_i$ be the estimated PMF for node i, $1 \leq i \leq n$. We need the following definition.

DEFINITION 4.1. *Fix $1 \leq i \leq n$. We call $\max_{1\leq k\leq K}\{\hat{\pi}_i(k)\}$ the (estimated) purity of node i and call community k the (estimated) home base of node i if $k = \mathrm{argmax}_{1\leq \ell\leq K}\{\hat{\pi}_i(\ell)\}$.*

When applying Mixed-SCORE to all data sets, we set $T = \log(n)$ in obtaining $\hat{R}$ and use the data-driven choice of L in (3.29).

4.1. *The two networks for statisticians.*

In a recent paper, Ji and Jin [34] has collected a network data set for statisticians, based on all published papers in Annals of Statistics, Biometrika, JASA, and JRSS-B, 2003 to the first half of 2012. The data set allows us to construct many networks. For reasons of space, we focus our study on a coauthorship network and a citee network, where each node is an author, and edges are defined as follows.

- *Coauthorship network.* There is an edge between two authors if they have coauthored at least two papers in the range of the data set. Our study focuses on the giant component of the network (236 nodes).

- *Citee network*. There is an edge between two authors if they have been cited at least once by the same author (other than themselves). We also focus on the giant component (1790 nodes) for our study.

Consider the Coauthorship network first. The network was suggested by [34] as the "High Dimensional Data Analysis" group which has a "Carroll-Hall" sub-group (including researchers in nonparametric and semi-parametric statistics, functional estimation, etc.) and a "North Carolina" sub-group (including researchers from Duke, North Carolina, and NCSU, etc.). In light of this, we consider a DCMM model assuming (a) there are two communities called "Carroll-Hall" and "North Carolina" respectively, and (b) some of the nodes have mixed memberships in two communities. We have applied Mixed-SCORE to the network, and the results are in Table 4.

TABLE 4
Left and Middle: high-degree pure nodes in the "Carroll-Hall" community and the "North Carolina" community. Right: highly mixed nodes (data: Coauthorship network).

Name	Deg.	Name	Deg.	Name	Deg.	Estimated PMF
Peter Hall	21	Joseph G Ibrahim	14	Jianqing Fan	16	54% of Carroll-Hall
Raymond J Carroll	18	David Dunson	8	Jason P Fine	5	54% of Carroll-Hall
T Tony Cai	10	Donglin Zeng	7	Michael R Kosorok	5	57% of Carroll-Hall
Hans-Georg Muller	7	Hongtu Zhu	7	J S Marron	4	55% of North Carolina
Enno Mammen	6	Alan E Gelfand	5	Hao Helen Zhang	4	51% of North Carolina
Jian Huang	6	Ming-Hui Chen	5	Yufeng Liu	4	52% of North Carolina
Yanyuan Ma	5	Bing-Yi Jing	4	Xiaotong Shen	4	55% of North Carolina
Bani Mallick	4	Dan Yu Lin	4	Kung-Sik Chan	4	55% of North Carolina
Jens Perch Nielsen	4	Guosheng Yin	4	Yichao Wu	3	51% of Carroll-Hall
Marc G Genton	4	Heping Zhang	4	Yacine Ait-Sahalia	3	51% of Carroll-Hall
Xihong Lin	4	Qi-Man Shao	4	Wenyang Zhang	3	51% of Carroll-Hall
Aurore Delaigle	3	Sudipto Banerjee	4	Howell Tong	2	52% of North Carolina
Bin Nan	3	Amy H Herring	3	Chunming Zhang	2	51% of Carroll-Hall
Bo Li	3	Bradley S Peterson	3	Yingying Fan	2	52% of North Carolina
Fang Yao	3	Debajyoti Sinha	3	Rui Song	2	52% of Carroll-Hall
Jane-Ling Wang	3	Kani Chen	3	Per Aslak Mykland	2	52% of North Carolina
Jiashun Jin	3	Weili Lin	3	Bee Leng Lee	2	54% of Carroll-Hall

In particular, it was found in [34] that the "Fan" group (Jianqing Fan and collaborators) has strong ties to both communities. Our results confirm such a finding but shed new light on the "Fan" group: many of the nodes (e.g., Yingying Fan, Rui Song, Yichao Wu, Chunming Zhang, Wenyang Zhang) have highly mixed memberships, and for each mixed node, we are able to quantify its weights in two communities. For example, both Runze Li (former graduate of UNC-Chapel Hill) and Jiancheng Jiang (former post-doc at UNC-Chapel Hill and current faculty member at UNC-Charlotte) have mixed memberships, but Runze Li is more on the "Carroll-Hall" community (weight: 73%) and Jiancheng Jiang is more on the "North Carolina" community (weight: 62%).

We now move to the Citee network. Ji and Jin [34] suggested that the

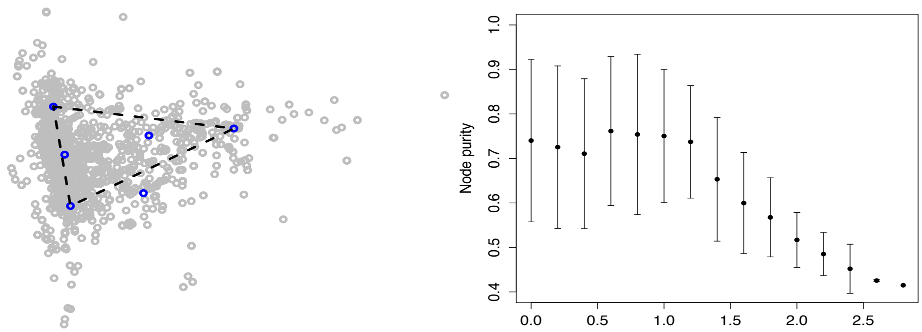

FIG 4. *Left: rows of $\hat{R}$; the dashed line traces the estimated* 2*-simplex by Mixed-SCORE. Right: node purity versus degree; x-axis represents the estimated degree parameters $\hat{\theta}(i)$ which are grouped together with an interval of .2; we plot the mean and standard deviation of $\hat{\theta}(i)$ in each group (data: Citee network).*

network has three meaningful communities: "Large Scale Multiple Testing" (MulTest), "Spatial and Nonparametric Statistics" (SpatNon) and "Variable Selection" (VarSelect). In light of this, we use a DCMM model with $K = 3$, and apply the Mixed-SCORE to the data. Figure 4 (left) presents the rows of $\hat{R} \in \mathbb{R}^{n,2}$, where a 2-simplex (i.e., triangle) is clearly visible in the cloud.

Tables 5-6 present the estimated PMF of high degree nodes. The results confirm those in [34] (especially on the existence of three communities aforementioned), but also shed new light on the network. First, it seems that high degree nodes in VarSelect are frequently observed to have an interest in MulTest, and this is not true the other way around (e.g., compare *Jianqing Fan*, *Hui Zou* with *Yoav Benjamini*, *Joseph Romano*). Second, the citations from SpatNon to either MulTest or VarSelect are comparably lower than those between MulTest and VarSelect. This fits well with our impression.

Conceivably, a node with higher degree tends to be more senior and so tends to be more mixed. This is confirmed by our results. Figure 4 (right) presents the plot of the node purity (see Definition 4.1) versus the estimated degree heterogeneity parameter $\hat{\theta}(i)$.[7] The results show a clear negative correlation between two quantities (especially on the right end, which corresponds to nodes with high degrees), which indicates that nodes with higher degrees tend to be more mixed.

4.2. *The Polbook network.* The network has 105 nodes, each represents a book on US politics published around the time of the 2004 presidential

[7] Letting $\hat{\xi}_1$, $\hat{b}_1$ and $\{\hat{w}_i\}_{i=1}^n$ be the same as those in Mixed-SCORE, we estimate $\theta(i)$ by $\hat{\theta}(i) = \hat{\xi}_1(i) \sum_{k=1}^K [\hat{w}_i(k)/\hat{b}_1(k)]$. In the oracle setting, the right hand side equals to $\theta(i)$.

TABLE 5
Estimated PMF of the 100 *nodes with the highest degrees in the Citee network, among which only the* 12 *purist nodes in each community are reported.*

Name	Deg.	MulTest	SpatNon	VarSelect	Name	Deg.	MulTest	SpatNon	VarSelect	Name	Deg.	MulTest	SpatNon	VarSelect
Felix Abramovich	366	0.943	0	0.057	Peter Muller	429	0.326	0.613	0.061	Lixing Zhu	432	0.121	0	0.879
Joseph Romano	377	0.868	0	0.132	Jeffrey Morris	452	0.146	0.519	0.335	Zhiliang Ying	382	0.107	0.027	0.866
Sara van de Geer	372	0.834	0	0.166	Michael Jordan	383	0.321	0.495	0.184	Zhezhen Jin	361	0.134	0	0.866
Yoav Benjamini	478	0.821	0	0.179	Mahlet Tadesse	383	0.373	0.493	0.134	Dennis Cook	424	0.253	0	0.747
David Donoho	484	0.819	0	0.181	Naijun Sha	383	0.373	0.493	0.134	Wenbin Lu	405	0.255	0	0.745
Christopher Genovese	521	0.810	0	0.190	Michael Stein	379	0.093	0.449	0.458	Dan Yu Lin	527	0.257	0	0.743
Larry Wasserman	535	0.800	0	0.200	Adrian Raftery	413	0.175	0.446	0.379	Donglin Zeng	489	0.270	0	0.730
Jon Wellner	387	0.798	0.05	0.152	Robert Kohn	429	0.310	0.428	0.262	Gerda Claeskens	404	0.247	0.033	0.720
Alexandre Tsybakov	521	0.784	0	0.216	George Casella	430	0.303	0.425	0.271	Yingcun Xia	358	0.302	0	0.698
Jiashun Jin	441	0.780	0	0.220	Marina Vannucci	571	0.304	0.418	0.278	Naisyin Wang	586	0.283	0.043	0.674
Yingying Fan	410	0.741	0	0.259	Bernard Silverman	577	0.514	0.395	0.091	Hua Liang	509	0.334	0	0.666
John Storey	544	0.737	0	0.263	Catherine Sugar	501	0.450	0.360	0.190	Wolfgang Karl Hardle	456	0.343	0	0.657

TABLE 6
Estimated PMF of the 12 *nodes with the highest degrees in the Citee network.*

Name	Deg.	MulTest	SpatNon	VarSelect	Name	Deg.	MulTest	SpatNon	VarSelect
Jianqing Fan	977	0.365	0.220	0.415	Peter Buhlmann	742	0.527	0.121	0.352
Raymond Carroll	850	0.282	0.294	0.424	Hans-Georg Muller	714	0.413	0.237	0.350
Hui Zou	824	0.348	0.225	0.427	Yi Lin	693	0.417	0.137	0.446
Peter Hall	780	0.501	0.032	0.467	Nocolai Meinshausen	692	0.462	0.125	0.413
Runze Li	778	0.282	0.226	0.491	Peter Bickel	692	0.529	0.216	0.255
Ming Yuan	748	0.391	0.166	0.444	Jian Huang	677	0.572	0	0.428

election and sold by the online bookseller Amazon.com. The edges are assigned by Amazon, where two books have an edge if they are frequently co-purchased by the same buyers, as indicated by the "*customers who bought this book also bought these other books*" feature on Amazon. By reading the descriptions and reviews of the books posted on Amazon, Mark Newman (see [41]) labeled each book as *liberal*, *neutral*, or *conservative*. Such labels are not exactly accurate but can be used as a reference.

We view the network as having two communities (liberal and conservative) and view neutral nodes as having mixed memberships in two communities, and so a DCMM model with $K = 2$ is appropriate. We applied Mixed-SCORE to the data with $K = 2$ and Figure 5 presents the estimated PMF for all nodes; note that for each node, the two entries of the estimated PMF are the estimated weights in liberal and conservative respectively. Since two weights sum to 1, Figure 5 only reports the weights in liberal.

For all except 9 books listed in Table 7, our results are nicely consistent with the community labels assigned by Newman: for a book that is labeled as liberal or conservative by Newman, our estimated PMF has a weight of approximately 1 in liberal or conservative, respectively; for a book that is labeled as neutral by Newman, our estimated PMF has significant weights in both liberal and conservative.

For the 9 books in Table 7, our results do not agree well with the labels assigned by Newman, and we have checked the background information of these books using multiple online resources (e.g., reader's comments, news

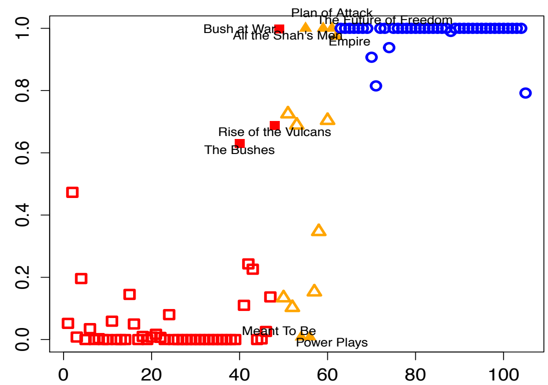

FIG 5. *Estimated PMF for all nodes (Polbook data, $K = 2$; y-axis: the first entry of the estimated PMF). Blue, red, and yellow represent conservative, liberal, and neutral nodes respectively, based on the labels assigned manually by Newman.*

TABLE 7
Books the memberships of which disagree with the labels (data: Polbook network).

Title	Author	Estimated PMF	Newman's label	Reasons for discrepancy
Empire	Michael Hardt	91.1% liberal	neutral	liberal book
The Future of Freedom	Fareed Zakaria	98.1% liberal	neutral	liberal book
Rise of the Vulcans	Michael Hardt	65.6% liberal	conservative	liberal book
All the Shah's Men	Stephen Kinzer	98.2% liberal	neutral	liberal author
Bush at War	Bob woodward	93.2% liberal	conservative	liberal author
Plan of Attack	Bob woodward	96.8% liberal	neutral	liberal author
Power Plays	Dick Morris	98.6% conservative	neutral	conservative author
Meant To Be	Lauren Morrill	98.7% conservative	neutral	not a political book
The Bushes	Peter Schweizer	60.3% liberal	conservative	our estimation is inaccurate

pages). For books #1-#6, we find either the book or the author is liberal. Note that we estimate these books as highly liberal, while Newman labeled them as either neutral or conservative. Similarly, for book #7, we find the author is conservative. Note that we estimate this book as conservative while Newman labeled it as neutral. For these reasons, we believe our estimates for these 7 books are more accurate. See [32, 51] where the authors also found Newman's labels could be incorrect for some of the nodes.

Book #8 is not a political book; this may explain the discrepancy between our result and Newman's label. For book #9, Newman's label seems to be right and our estimate may not be accurate enough.

4.3. *The Football network.* This is a network for American football games between Division I-A college teams during the regular football season of Fall 2000 (Girvan and Newman [26]). Each node represents a team and there is an edge between two teams if they have played one or more games. There

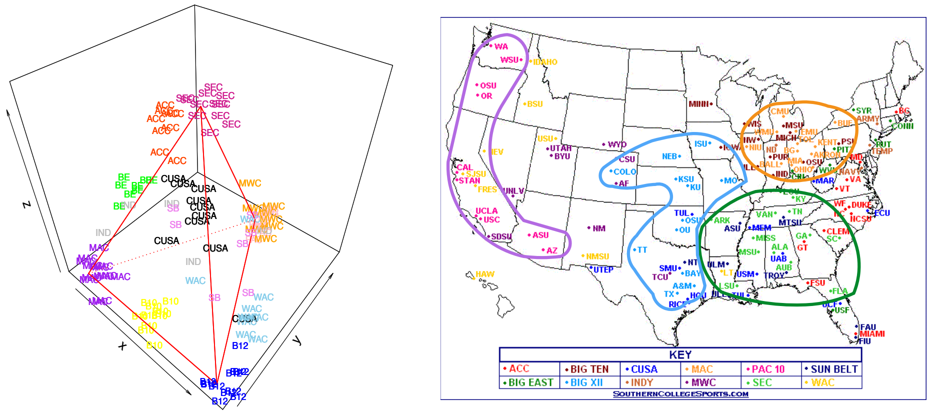

FIG 6. *Left: each point represents a row of $\hat{R}$ (and each row corresponds to a team where the label is the conference it belongs to); a 3-simplex is clearly visible in the cloud. Right: locations of all teams. Four contours are manually added to highlight the connection between the geographical locations and community structures. In each panel, teams in the same conference are in the same color (data: Football network).*

are a total of 115 nodes, where 5 of them are called "Independents". For administration purpose, the remaining 110 nodes are *manually* divided into 11 conferences, each with a size from 7 to 13; see Table 8.

We note that a conference is not necessarily a community, and vice versa. We hypothesize a DCMM model holds with fewer than 11 communities; such a viewpoint is different from [26] which assumes a non-mixing non-overlapping network model where each conference is interpreted as a community. In this spirit, we have applied Mixed-SCORE to the network assuming there are K communities for $2 \leq K \leq 6$, and it seems $K = 4$ gives the most interpretable results. Below, we report the result for $K = 4$.

In particular, for $K = 4$, if we let $\hat{R} \in \mathbb{R}^{n,3}$ be the matrix of entry-wise ratios and view each row of $\hat{R}$ as a point in $\mathbb{R}^3$ as before, then in the cloud of points, a nice 3-simplex is clearly visible; see Figure 6 (left panel).

In Table 8, for each of the 11 conferences, we tabulate the average of the estimated PMF (i.e., $\hat{\pi}_i$) across different teams. The results suggest that geographical locations play an important role in the community structures:

- The four communities can be interpreted as "North East",[8] "South East", "South Central", and "West Coast", respectively.
- The four conferences MAC, SEC, Big 12, and PAC 10 consist of most of the pure nodes in "North East", "South East", "South Central",

[8] Most of the teams are located in north east or in middle west.

TABLE 8

The average of estimated PMF across different teams in each conference (the 4 entries of the PMF are in Columns 3-6, respectively; numbers in the brackets: standard deviations).

Conference (abbreviation)	size	"North East"	"South East"	"South Central"	"West Coast"
Mid-American (MAC)	13	**.93** (.06)	.03 (.05)	.03 (.04)	.01 (.03)
Southeastern (SEC)	12	.03 (.04)	**.94** (.04)	.01 (.02)	.02 (.03)
Big Twelve (Big 12)	12	.03 (.04)	.02 (.02)	**.92** (.06)	.03 (.06)
Pacific Ten (PAC 10)	10	.02 (.02)	0 (.0)	.02 (.03)	**.96** (.05)
Atlantic Coast (ACC)	9	.24 (.04)	.73 (.04)	0 (.0)	.03 (.02)
Big East (Big East)	8	.54 (.06)	.33 (.04)	0 (.0)	.13 (.04)
Big Ten (Big 10)	11	.56 (.05)	0 (.0)	.25 (.06)	.19 (.06)
Conference USA (CUSA)	10	.10 (.11)	.61 (.18)	.26 (.15)	.03 (.08)
Mountain West (MWC)	8	0(.0)	.23 (0.10)	.12 (.09)	.65 (.12)
Sun Belt (Sun Belt)	7	.06 (.11)	.40 (.16)	.33 (.20)	.21 (.25)
Western Athletic (WAC)	10	.02 (.07)	.16 (.09)	.53 (.15)	.29 (.13)

and "West Coast", respectively.[9]

- The other seven conferences contain mostly mixed nodes (the 5 independent teams are also mixed nodes).

Figure 6 presents the geographical locations for all 115 teams (teams in the same conference are in the same color).[10] For illustration, we have grouped the teams in MAC, SEC, Big 12, and PAC 10 with a contour in orange, green, blue, and purple, respectively, to highlight the connection between the community partition and the geographical locations (e.g., the purple contour circumvents all teams in PAC 10; note that some other teams also fall within the contour).

For most of the mixed nodes, our estimated PMF is consistent with the geographical distance of the node to each of the four communities. One example is MWC (i.e., Mountain West Conference), where for most teams in this conference, the estimated PMF has a high weight in "West Coast". This is consistent with the fact that these teams are close to West Coast geographically. Another example is WAC (i.e., Western Athletic Conference), where a similar claim can be drawn. Especially, for each team in WAC, the estimated PMF has very little weight in "North East".

Compared to Girvan and Newman [26], our results (especially that on the connection between geographical locations and community structures) shed new light on the data set and provide very different perspectives.

5. Simulations. We compare Mixed-SCORE with OCCAM [60] (in Experiments 2-6) and LPC [30] (in Experiment 7). The reason for choos-

[9] Due to estimation errors, we rarely see an estimated PMF has exactly 1 nonzero entry; we think a node as pure if the estimated purity (Definition 4.1) is very close to 1.

[10] The figure was downloaded from SouthernCollegeSports.com. For very few teams, the figure does not match the data set because conferences change occasionally.

ing these two competitors is that they are both model-based methods that output node "memberships", and they both account for degree heterogeneity (explicitly or implicitly). OCCAM assigns to each node a non-negative "membership" vector with unit ℓ_2-norm; we renormalize these vectors by their ℓ_1-norms and use them as the estimated PMF. LPC outputs a posterior PMF for each node (describing its posterior probabilities of being drawn from different components of a mixture), which we use as the estimated PMF of that node; to implement LPC, we use the R package *latentnet* and the default algorithm parameters. To implement Mixed-SCORE, we plug in the SVS algorithm for vertex hunting, where L is the data-driven one in (3.29).

For most experiments below, we set $n = 500$ and $K = 3$. For $0 \leq n_0 \leq 160$, let each community have n_0 number of pure nodes. Fixing $x \in (0, 1/2)$, let the mixed nodes have four different memberships $(x, x, 1-2x)$, $(x, 1-2x, x)$, $(1-2x, x, x)$ and $(1/3, 1/3, 1/3)$, each with $(500 - 3n_0)/4$ number of nodes. Fixing $\rho \in (0, 1)$, the matrix P has diagonals 1 and off-diagonals ρ. Fixing $z \geq 1$, we generate the degree parameters such that $1/\theta(i) \overset{iid}{\sim} U(1, z)$, where $U(1, z)$ denotes the uniform distribution on $[1, z]$. The tuning parameter L is selected as in (3.29). For each parameter setting, we report $n^{-1}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|^2$ averaged over 100 repetitions.

Experiment 1: Tuning parameter selection. We first study the choice of the tuning parameter L in Mixed-SCORE. We aim to see (i) how the estimation errors change for a range of L, and (ii) how the adaptive choice $\hat{L}_n^*(A)$ in (3.29) performs. Fix $(x, \rho, z) = (0.4, 0.2, 5)$ and let n_0 range in $\{60, 80, 100\}$. For each setting, we run Mixed-SCORE with $L \in \{4, 5, \cdots, 9\}$ and $\hat{L}_n^*(A)$. The results are displayed in Figure 7. First, when there are relatively few mixed nodes (e.g., $n_0 = 100$), small values of L yield good performance; but as the number of mixed nodes going up, we favor larger values of L; these match our theoretical results (Lemmas 3.4-3.6). Second, under the circumstances of a moderate number of mixed nodes (e.g., $n_0 = 60, 80$), for a range of L (e.g., $L \in \{7, 8, 9\}$), the statistical errors of Mixed-SCORE are similar, and $\hat{L}_n^*(A)$ falls in this range with high probability. Figure 8 shows the estimated 2-simplex in one repetition ($n_0 = 80$), and the simplex changes very little when L falls in a range.

Experiment 2: Fraction of pure nodes. Fix $(x, \rho, z) = (0.4, 0.1, 5)$ and let n_0 range in $\{40, 60, 80, 100, 120, 160\}$. As n_0 increases, the fraction of pure nodes increases from around 25% to around 95%. The results are displayed in top left panel of Figure 9. It suggests that when the fraction of pure nodes is $<$ 70%, Mixed-SCORE significantly outperforms OCCAM; when the fraction of pure nodes is $>$ 70%, the two methods have similar performance.

Experiment 3: Connectivity across communities. Fix $(x, n_0, z) = (0.4, 80, 5)$

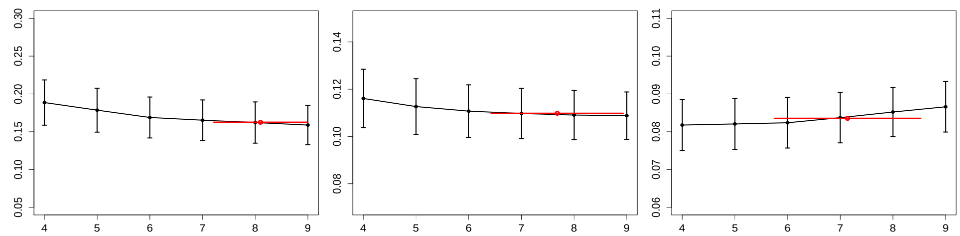

FIG 7. *Performance of Mixed-SCORE as the tuning parameter L varies (y-axis: estimation errors; $\hat{L}_n^*(A)$ is plotted in red; both mean and standard deviation are displayed). From left to right, there are* $60, 80, 100$ *pure nodes in each community, respectively.*

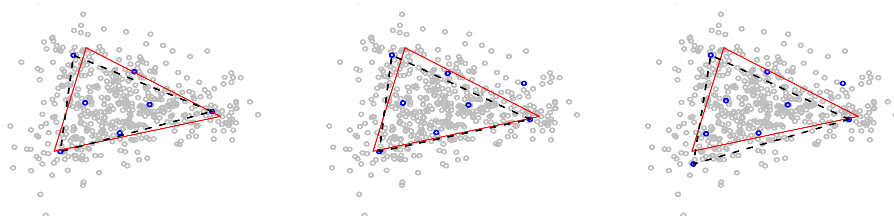

FIG 8. *Illustration of the Vertex Hunting step. From left to right,* $L = 7, 8, 9$. *Although the local cluster centers (blue points) are different, the estimated* 2*-simplex (dashed black) changes very little, and it approximates the IS (solid red) well.*

and let ρ range in $\{0.05, 0.1, 0.15, \cdots, 0.5\}$. The larger ρ, the more edges across different communities. The results are presented in top right panel of Figure 9. We see that the performance of Mixed-SCORE improves as ρ decreases. One possible reason is that, for ρ large, it is relatively more difficult to identify the vertices of the Ideal Simplex. Furthermore, Mixed-SCORE is better than OCCAM in all settings.

Experiment 4: Purity of mixed nodes. Fix $(n_0, \rho, z) = (80, 0.1, 5)$ and let x range in $\{0.05, 0.1, 0.15, \cdots, 0.5\}$. We recall Definition 4.1 for the "purity" of a node. In our settings, there are four types of mixed nodes, and the purity of the first three types of mixed nodes is $(1-2x)1\{x \leq 1/3\} + x1\{x > 1/3\}$. Therefore, as x increases to $1/3$, these nodes become less pure; then, as x further increases, these nodes become more pure. The results are in bottom left panel of Figure 9. It suggests that estimating the memberships becomes harder as the purity of mixed nodes decreases, and Mixed-SCORE outperforms OCCAM in almost all settings. Especially in the highly "mixing" case (say, x is close to $1/3$), Mixed-SCORE is much better than OCCAM.

Experiment 5: Degree heterogeneity. Fix $(x, n_0, \rho) = (0.4, 80, 0.1)$ and let z range in $\{1, 2, \cdots, 8\}$. Since $1/\theta(i) \stackrel{iid}{\sim} U(1, z)$, a larger z implies that the

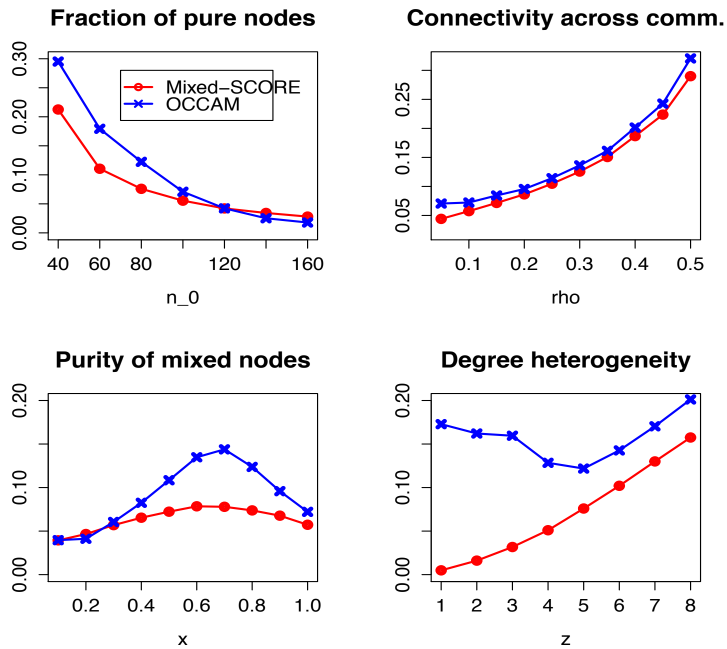

FIG 9. *Estimation errors of Mixed-SCORE and OCCAM (y-axis: $n^{-1}\sum_{i=1}^{n}\|\hat{\pi}_i - \pi_i\|^2$).*

nodes have lower degrees and are more heterogeneous (hence, the problem becomes more difficult). The results are presented in bottom right panel of Figure 9. It suggests that Mixed-SCORE uniformly outperforms OCCAM. Interestingly, when z is small (so the problem is "easy"), Mixed-SCORE is very accurate, but the performance of OCCAM is unsatisfactory.

Experiment 6: Mixed memberships taking continuous values. In this experiment, we generate the mixed memberships from a continuous distribution. Set $(n, K) = (500, 3)$ and let P have diagonals 1 and off-diagonals 0.3. Each community has $n_0 = 25$ pure nodes. The π_i of remaining nodes are *iid* drawn as follows: We generate $\pi_i(1)$ and $\pi_i(2)$ independently from $U(1/6, 1/2)$ and set $\pi_i(3) = 1 - \pi_i(1) - \pi_i(2)$. The degree parameters $\theta(i)$ are *iid* drawn from $\alpha_n \cdot U(1, 2)$, where $\alpha_n > 0$ controls the sparsity of the network. Let α_n range in $\{0.02, 0.04, 0.06, \cdots, 0.20\}$. The results are presented in Table 9. This setting does not satisfy the regularity conditions (3.23)-(3.24) on π_i's, however, Mixed-SCORE still has a good performance and outperforms OCCAM. It suggests that the regularity conditions on π_i's are only for theoretical convenience, and our method indeed works for broader settings.

Experiment 7: Comparison with latent space approach. We compare Mixed-SCORE with the Bayesian method based on LPC [30] (we use the R package

TABLE 9
Estimation errors in Experiment 6, where π_i's take continuous values.

α_n	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
Mixed-SCORE	.38	.35	.36	.32	.30	.28	.23	.18	.15	.12
OCCAM	.44	.42	.41	.41	.38	.36	.32	.28	.26	.23

latentnet). In this experiment, we fix $n = 120$, $K = 3$, $(x, \rho, z) = (0.4, 0.3, 5)$, and let n_0 range in $\{12, 16, 20, \cdots, 32, 36\}$ (so the number of mixed nodes in each group decreases from 21 to 3). The results are displayed in Figure 10. We find that, when the fraction of mixed nodes is comparably small, LPC has a perfect performance; however, as the fraction of mixed nodes increases to more than 40%, the performance of LPC deteriorates rapidly; one reason is that, when n_0 is not very large, LPC often estimates the PMF of all the nodes as the same. In contrast, the performance of Mixed-SCORE is quite stable. In terms of computing time, Mixed-SCORE takes only seconds for one repetition while LPC takes > 20 minutes (both measured in R).

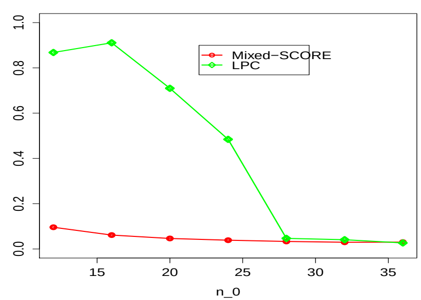

FIG 10. *Estimation errors of Mixed-SCORE and LPC (y-axis: $n^{-1}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|^2$).*

6. Proof of Theorem 2.2. Let H be the orthogonal matrix in Theorem 2.1. We aim to show that, with probability $1 - o(n^{-3})$, for all $1 \leq i \leq n$,

$$\|\hat{\pi}_i - \pi_i\|_1 \leq C\|H\hat{r}_i - r_i\| + C \max_{1\leq k\leq K} \|H\hat{v}_k - v_k\| + CKerr_n. \tag{6.30}$$

Once (6.30) is true, the claim follows immediately from the efficiency of VH algorithm (see Definition 1.3) and the bound in Theorem 2.1.

Below, we show (6.30). In the Membership Reconstruction (MR) step, we compute $\hat{w}_i$ and $\hat{b}_1$, then use them to construct

$$\hat{\pi}_i^*(k) = \max\{0, \hat{w}_i(k)/\hat{b}_1(k)\}, \qquad 1 \leq k \leq K, \tag{6.31}$$

and then estimates π_i by $\hat{\pi}_i = \hat{\pi}_i^*/\|\hat{\pi}_i^*\|_1$. We shall study $\hat{w}_i$ and $\hat{b}_1$ separately and then combine their error bounds to get (6.30).

First, we study $\hat{w}_i$. By definition,

$$\underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_K \end{pmatrix}}_{\equiv Q} w_i = \begin{pmatrix} 1 \\ r_i \end{pmatrix}, \qquad \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ H\hat{v}_1 & \cdots & H\hat{v}_K \end{pmatrix}}_{\equiv \hat{Q}} \hat{w}_i = \begin{pmatrix} 1 \\ H\hat{r}_i \end{pmatrix}.$$

We thus write

$$\begin{aligned} \hat{w}_i - w_i &= \hat{Q}^{-1}\begin{pmatrix} 1 \\ H\hat{r}_i \end{pmatrix} - Q^{-1}\begin{pmatrix} 1 \\ r_i \end{pmatrix} \\ &= \hat{Q}^{-1}\left[\begin{pmatrix} 1 \\ H\hat{r}_i \end{pmatrix} - \begin{pmatrix} 1 \\ r_i \end{pmatrix}\right] - (Q^{-1} - \hat{Q}^{-1})\begin{pmatrix} 1 \\ r_i \end{pmatrix} \\ &= \hat{Q}^{-1}\begin{pmatrix} 0 \\ H\hat{r}_i - r_i \end{pmatrix} - \hat{Q}^{-1}(\hat{Q} - Q)Q^{-1}\begin{pmatrix} 1 \\ r_i \end{pmatrix} \\ &= \hat{Q}^{-1}\begin{pmatrix} 0 \\ H\hat{r}_i - r_i \end{pmatrix} - \hat{Q}^{-1}(\hat{Q} - Q)w_i. \end{aligned}$$

It follows that

$$\|\hat{w}_i - w_i\| \le \|\hat{Q}^{-1}\| \cdot \big(\|H\hat{r}_i - r_i\| + \|(\hat{Q} - Q)w_i\|\big). \tag{6.32}$$

This matrix Q is studied in the proof of Lemma B.3, where we prove $\|Q^{-1}\| = O(1/\sqrt{K})$; see (B.52). This means the minimum singular value of Q is lower bounded by $C\sqrt{K}$. Moreover, $\|\hat{Q}-Q\| \le \|\hat{Q}-Q\|_F \le \sqrt{K}\max_{1\le k\le K}\|H\hat{v}_k - v_k\| = o(\sqrt{K})$. As a result, the minimum singular value of $\hat{Q}$ is also lower bounded by $C\sqrt{K}$. It leads to

$$\|\hat{Q}^{-1}\| \le C/\sqrt{K}.$$

We note that $(\hat{Q} - Q)w_i \in \mathbb{R}^K$ is a vector whose first entry is 0 and whose remaining entries are equal to $\sum_{k=2}^K w_i(k)(\hat{v}_k - v_k) \in \mathbb{R}^{K-1}$. Since w_i contains the coefficients of writing r_i as a convex combination of $v_1, \ldots, v_K$, we have $\|w_i\|_1 = 1$. Therefore,

$$\|(\hat{Q}-Q)w_i\| = \Big\|\sum_{k=1}^K w_i(k)(H\hat{v}_k - v_k)\Big\| \le \sum_{k=1}^K w_i(k)\|H\hat{v}_k - v_k\| \le \max_{1\le k\le K}\|H\hat{v}_k - v_k\|.$$

Plugging in the above results into (6.32) gives

$$\|\hat{w}_i - w_i\| \le CK^{-1/2}\big(\|H\hat{r}_i - r_i\| + \max_{1\le k\le K}\|H\hat{v}_k - v_k\|\big). \tag{6.33}$$

Next, we study $\hat{b}_1$. Recall that

$$\hat{b}_1(k) = [\hat{\lambda}_1 + \hat{v}_k' \mathrm{diag}(\hat{\lambda}_2, \cdots, \hat{\lambda}_K)\hat{v}_k]^{-1/2}.$$

By Lemma 1.1, $b_1(k)$ has the same form except that $\hat{\lambda}_k$ and $\hat{v}_k$ are replaced by their population counterparts. Letting $\Lambda_0 = \mathrm{diag}(\lambda_2, \cdots, \lambda_K)$ and $\hat{\Lambda}_0 = \mathrm{diag}(\hat{\lambda}_2, \cdots, \hat{\lambda}_K)$, we write

$$\frac{1}{\hat{b}_1^2(k)} = \hat{\lambda}_1 + \hat{v}_k' \hat{\Lambda}_0 \hat{v}_k, \qquad \frac{1}{b_1^2(k)} = \lambda_1 + v_k' \Lambda_0 v_k.$$

By direct calculations,

$$\begin{aligned}
&|\frac{1}{\hat{b}_1^2(k)} - \frac{1}{b_1^2(k)}| \leq |\hat{\lambda}_1 - \lambda_1| + |\hat{v}_k'\hat{\Lambda}_0\hat{v}_k - v_k'\Lambda_0 v_k| \\
&= |\hat{\lambda}_1 - \lambda_1| + |\hat{v}_k' H' H \hat{\Lambda}_0 \hat{v}_k - v_k' \Lambda_0 v_k| \\
&\leq |\hat{\lambda}_1 - \lambda_1| + |\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| + |\hat{v}_k' H'(H\hat{\Lambda}_0 - \hat{\Lambda}_0 H)\hat{v}_k| + |v_k'(\hat{\Lambda}_0 - \Lambda_0)v_k| \\
&\leq |\hat{\lambda}_1 - \lambda_1| + |\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| + \|\hat{v}_k\|^2 \|H\hat{\Lambda}_0 - \hat{\Lambda}_0 H\| + \|v_k\|^2 \|\hat{\Lambda}_0 - \Lambda_0\| \\
&\leq (1 + \|v_k\|^2) | \max_\ell |\hat{\lambda}_\ell - \lambda_\ell| + |\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| + \|\hat{v}_k\|^2 \|H\hat{\Lambda}_0 - \hat{\Lambda}_0 H\|.
\end{aligned}$$

First, by Lemma C.1, $\max_\ell |\hat{\lambda}_\ell - \lambda_\ell| \leq C\sqrt{\theta_{\max}\|\theta\|_1}$. Second, by Lemma C.4, $\|H\hat{\Lambda}_0 - \hat{\Lambda}_0 H\| \leq C\sqrt{\theta_{\max}\|\theta\|_1}$. Third, by Lemma B.3, $\|v_k\| \leq C\sqrt{K}$; since $\max_\ell \|\hat{v}_\ell - v_\ell\| = o(\sqrt{K})$, it follows that $\|\hat{v}_k\| \leq C\sqrt{K}$. Combining the above gives

$$|\frac{1}{\hat{b}_1^2(k)} - \frac{1}{b_1^2(k)}| \leq |\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| + CK\sqrt{\theta_{\max}\|\theta\|_1}. \tag{6.34}$$

Since $\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k = v_k'\hat{\Lambda}_0 v_k + 2v_k'\hat{\Lambda}_0(H\hat{v}_k - v_k) + (H\hat{v}_k - v_k)'\hat{\Lambda}_0(H\hat{v}_k - v_k)$, we have

$$|\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| \leq 2\|v_k\|\|\hat{\Lambda}_0\|\|H\hat{v}_k - v_k\| + \|\hat{\Lambda}_0\|\|H\hat{v}_k - v_k\|^2.$$

By Lemma B.1 and Lemma C.1, $\|\Lambda_0\| \leq C\beta_n K^{-1}\|\theta\|^2$ and $\|\hat{\Lambda}_0 - \Lambda_0\| \leq C\sqrt{\theta_{\max}\|\theta\|_1} = o(K\beta_n^{-1}\|\theta\|^2)$. It follows that $\|\hat{\Lambda}_0\| \leq C\beta_n K^{-1}\|\theta\|^2$. Also, as we have argued before, $\|v_k\| \leq C\sqrt{K}$ and $\|H\hat{v}_k - v_k\| = o(\sqrt{K})$. Plugging these results into the above inequality gives

$$|\hat{v}_k' H' \hat{\Lambda}_0 H \hat{v}_k - v_k'\hat{\Lambda}_0 v_k| \leq CK^{-1/2}\beta_n\|\theta\|^2\|H\hat{v}_k - v_k\|.$$

We then plug it into (6.34) to get

$$|\frac{1}{\hat{b}_1^2(k)} - \frac{1}{b_1^2(k)}| \leq CK^{-1/2}\beta_n\|\theta\|^2\|H\hat{v}_k - v_k\| + CK\sqrt{\theta_{\max}\|\theta\|_1}. \tag{6.35}$$

In the proof of Lemma B.2, we have shown $b_1(k) \asymp \|\theta\|^{-1}$; see (B.48). Then, $\frac{1}{b_1^2(k)} \asymp \|\theta\|^2$. Combining it with (6.35), we have $\frac{1}{\hat{b}_1^2(k)} = \frac{1}{b_1^2(k)}[1+o(1)] \asymp \|\theta\|^2$. It follows that

$$
\begin{aligned}
|\frac{1}{\hat{b}_1(k)} - \frac{1}{b_1(k)}| &= |\frac{1}{\hat{b}_1(k)} + \frac{1}{b_1(k)}|^{-1} \cdot |\frac{1}{\hat{b}_1^2(k)} - \frac{1}{b_1^2(k)}| \\
&\leq C\|\theta\|^{-1} \cdot |\frac{1}{\hat{b}_1^2(k)} - \frac{1}{b_1^2(k)}| \\
&\leq CK^{-1/2}\beta_n\|\theta\|\|H\hat{v}_k - v_k\| + C\|\theta\|^{-1}K\sqrt{\theta_{\max}\|\theta\|_1} \\
&\leq CK^{-1/2}\beta_n\|\theta\|\|H\hat{v}_k - v_k\| + CK\|\theta\|err_n, \qquad (6.36)
\end{aligned}
$$

where the last line is because $err_n = (\theta_{\max}/\theta_{\min}) \cdot \|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1 \log(n)} \gg \|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1}$.

Last, we combine the results for $(\hat{w}_i, \hat{b}_1)$ to prove (6.30). Recall that $\hat{\pi}_i^*$ is as defined in (6.31). Introduce its non-stochastic counterpart π_i^* by

$$
\pi_i^*(k) = w_i(k)/b_1(k), \qquad 1 \leq k \leq K. \qquad (6.37)
$$

Since $\pi_i^*(k) \geq 0$, in (6.31), the operation of truncating at zero can only make it closer to $\pi_i^*(k)$. It follows that

$$
\begin{aligned}
|\hat{\pi}_i^*(k) - \pi_i^*(k)| &\leq |\hat{w}_i(k)/\hat{b}_1(k) - \pi_i^*(k)| \\
&= |\hat{w}_i(k)/\hat{b}_1(k) - w_i(k)/b_1(k)| \\
&\leq \frac{1}{\hat{b}_1(k)}|\hat{w}_i(k) - w_i(k)| + w_i(k)|\frac{1}{\hat{b}_1(k)} - \frac{1}{b_1(k)}|.
\end{aligned}
$$

We sum over k on both sides and note that $\hat{b}_1(k) \asymp \|\theta\|^{-1}$ (see the paragraph above (6.36)) and $\|w_i\|_1 = 1$. It yields

$$
\begin{aligned}
\|\hat{\pi}_i^* - \pi_i^*\|_1 &\leq C\|\theta\|\|\hat{w}_i - w_i\|_1 + |\frac{1}{b_1(k)} - \frac{1}{\hat{b}_1(k)}| \\
&\leq C\|\theta\|\sqrt{K}\|\hat{w}_i - w_i\| + \max_{1\leq k\leq K}|\frac{1}{b_1(k)} - \frac{1}{\hat{b}_1(k)}| \\
&\leq C\|\theta\|\big(\|H\hat{r}_i - r_i\| + \max_{1\leq k\leq K}\|H\hat{v}_k - v_k\| + Kerr_n\big), \qquad (6.38)
\end{aligned}
$$

where in the second line we have used Cauchy-Schwarz inequality and in the last line we have plugged in (6.33) and (6.36). By definition, $\hat{\pi}_i = \hat{\pi}_i^*/\|\hat{\pi}_i^*\|_1$. By the triangular inequality,

$$
|\hat{\pi}_i(k) - \pi_i(k)| \leq \frac{1}{\|\pi_i^*\|_1}|\hat{\pi}_i^*(k) - \pi_i^*(k)| + \hat{\pi}_i^*(k)|\frac{1}{\|\hat{\pi}_i^*\|_1} - \frac{1}{\|\pi_i^*\|_1}|
$$

$$
\begin{aligned}
&= \frac{1}{\|\pi_i^*\|_1}|\hat{\pi}_i^*(k) - \pi_i^*(k)| + \frac{\hat{\pi}_i(k)}{\|\pi_i^*\|_1}|\|\hat{\pi}_i^*\|_1 - \|\pi_i^*\|_1| \\
&\leq \frac{1}{\|\pi_i^*\|_1}\big(|\hat{\pi}_i^*(k) - \pi_i^*(k)| + \hat{\pi}_i(k)\|\hat{\pi}_i^* - \hat{\pi}_i^*\|_1\big),
\end{aligned}
$$

where the last inequality is because $|\|\hat{\pi}_i^*\|_1 - \|\pi_i^*\|_1| \leq \|\hat{\pi}_i^* - \pi_i^*\|_1$. We sum over k on both sides and note that $\sum_k \hat{\pi}_i(k) = 1$ by definition. It follows that

$$
\|\hat{\pi}_i - \pi_i\|_1 \leq \frac{1}{\|\pi_i^*\|_1} \cdot 2\|\hat{\pi}_i^* - \pi_i^*\|_1.
$$

By (6.37), $\|\pi^*\|_1 \geq \|w_i\|_1 \cdot \min_k \frac{1}{b_1(k)}$. In the paragraph above (6.36), we have seen that $b_1(k) \asymp \|\theta\|^{-1}$. This suggests that $\|\pi_i^*\|_1 \geq C\|\theta\|$. As a result,

$$
\begin{aligned}
\|\hat{\pi}_i - \pi_i\|_1 &\leq C\|\theta\|^{-1} \cdot \|\hat{\pi}_i^* - \pi_i^*\|_1 \\
&\leq C\big(\|H\hat{r}_i - r_i\| + \max_{1\leq k\leq K} \|H\hat{v}_k - v_k\| + Kerr_n\big).
\end{aligned}
\tag{6.39}
$$

This gives (6.30). The proof is now complete. □

7. Discussion. The paper is related to community detection [17, 35, 53, 54, 61, 35] but deals with a more challenging problem of membership estimation, where the simplex structure and the Vertex Hunting algorithm are new. The theory is also much more demanding technically, which requires sharp row-wise deviation bounds for empirical eigenvectors and careful analysis of the Vertex Hunting error.

Most works on membership estimation focus on the MMSB model [3, 4, 28], which does not model degree heterogeneity. In comparison, our method allows severe degree heterogeneity and provides encouraging results in analyzing real networks (see Section 4).

A key component of our method is the SCORE normalization [35] which normalizes the leading eigenvectors by dividing it with the first eigenvector $\hat{\xi}_1$ entry by entry. Alternatively, we may normalize with a different eigenvector (e.g., $\hat{\xi}_2$) or the row-wise ℓ^q-norms of $\hat{\Xi} = [\hat{\xi}_1, \ldots, \hat{\xi}_K]$. Unfortunately, other normalizations do not produce a simplex structure (see our remarks in Section 1.2). Another key component is the Vertex Hunting (VH). In our method, VH is treated as a "plug-in" step, where different plugged-in VH algorithms lead to different variants of Mixed-SCORE. We have investigated the use of existing VH algorithms, and also proposed two new VH algorithms that significantly improve existing ones. The new VH algorithms can be applied to other applications (e.g., hyperspectral unmixing [18, 58] or archetypal analysis [47]) and are of independent interest.

Our work is related to Nonnegative Matrix Factorization [44]. The Ideal Simplex we discover here is reminiscent of the simplicial cone [21] in NMF problems, but is different: the Ideal Simplex is associated with eigenvectors, while the simplicial cone in NMF is associated with the data matrix. Mixed-SCORE is a PCA approach and is thus related to the recent work on IF-PCA [38], sparse PCA [8, 14], and high-dimensional clustering [57, 37, 38].

Our idea can be extended in many ways. In a forthcoming manuscript, we extend the idea to directed or bi-partite networks. In another forthcoming manuscript, we investigate the optimality of Mixed-SCORE using a minimax framework. The idea can also be extended to topic estimation in text mining [2, 11, 12], a problem that has received a lot of attention recently. See [40]. It is also of interest to extend the idea to study dynamic networks and network data where some covariates (i.e., ages and affiliations of the authors in the statistician's network) are available.

REFERENCES

[1] Abbe, E., Fan, J., Wang, K. and Zhong, Y. (2017). Entrywise eigenvector analysis of random matrices with low expected rank. *arXiv:1709.09565*.

[2] Airoldi, E. and Bischof, J. (2017). A regularization scheme on word occurrence rates that improves estimation and interpretation of topical content (with discussion). *J. Amer. Statist. Assoc.* **111** 1382–1403.

[3] Airoldi, E., Blei, D., Fienberg, S. and Xing, E. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.

[4] Anandkumar, A., Ge, R., Hsu, D. and Kakade, S. (2014). A tensor approach to learning mixed membership community models. *J. Mach. Learn. Res.* **15** 2239–2312.

[5] Araújo, M. C. U., Saldanha, T. C. B., Galvao, R. K. H., Yoneyama, T., Chame, H. C. and Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems* **57** 65–73.

[6] Ball, B., Karrer, B. and Newman, M. (2011). An efficient and principled method for detecting communities in networks. *Phys. Rev. E* **84** 036103.

[7] Bandeira, A. S. and Van Handel, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.* **44** 2479–2506.

[8] Berthet, Q. and Rigollet, P. (2013). Optimal detection of sparse principal components in high dimension. *Ann. Statist.* **41** 1780–1815.

[9] Bioucas-Dias, J. M. (2009). A variable splitting augmented Lagrangian approach to linear spectral unmixing. In *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS'09. First Workshop on* 1–4. IEEE.

[10] Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P. and Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing* **5** 354–379.

[11] Blei, D. and Lafferty, J. (2007). A correlated topic model of science. *Ann. Appl. Statist.* **1** 17–35.

[12] Blei, D., Ng, A. and Jordan, M. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* **3** 993–1022.

[13] Boardman, J. W., Kruse, F. A. and Green, R. O. (1995). Mapping target signatures via partial unmixing of AVIRIS data.

[14] Cai, T., Ma, Z. and Wu, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist.* **41** 3074–3110.

[15] Chan, T.-H., Chi, C.-Y., Huang, Y.-M. and Ma, W.-K. (2009). A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Transactions on Signal Processing* **57** 4418–4432.

[16] Chang, C.-I., Wu, C.-C., Liu, W. and Ouyang, Y.-C. (2006). A new growing method for simplex-based endmember extraction algorithm. *IEEE Transactions on Geoscience and Remote Sensing* **44** 2804–2819.

[17] Chen, Y., Li, X. and Xu, J. (2018). Convexified modularity maximization for degree-corrected stochastic block models. *Ann. Statist.* **46** 1573–1602.

[18] Craig, M. D. (1994). Minimum-volume transforms for remotely sensed data. *IEEE Trans. on Geoscience and Remote Sens.* **32** 542–552.

[19] Cutler, A. and Breiman, L. (1994). Archetypal analysis. *Technometrics* **36** 338–347.

[20] Davis, C. and Kahan, W. M. (1970). The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* **7** 1–46.

[21] Donoho, D. and Stodden, V. (2004). When does non-negative matrix factorization give a correct decomposition into parts? In *Adv. Neural Inf. Process. Syst.* 1141–1148.

[22] Erdős, L., Schlein, B. and Yau, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *The Annals of Probability* **37** 815–852.

[23] Gao, C., Ma, Z., Zhang, A. Y. and Zhou, H. H. (2018). Community detection in degree-corrected block models. *Ann. Statist.* **46** 2153–2185.

[24] Gillis, N. and Vavasis, S. A. (2013). Fast and robust recursive algorithms for separable nonnegative matrix factorization. *IEEE transactions on pattern analysis and machine intelligence* **36** 698–714.

[25] Gillis, N. and Vavasis, S. A. (2015). Semidefinite programming based preconditioning for more robust near-separable nonnegative matrix factorization. *SIAM Journal on Optimization* **25** 677–698.

[26] Girvan, M. and Newman, M. (2002). Community structure in social and biological networks. *Proc. Nat. Acad. Sci.* **99** 7821–7826.

[27] Goldenberg, A., Zheng, A., Fienberg, S. and Airoldi, E. (2010). A survey of statistical network models. *Foundations and Trends in Machine Learning* **2** 129–233.

[28] Gopalan, P., Gerrish, S., Freedman, M., Blei, D. and Mimno, D. (2012). Scalable inference of overlapping communities. In *Adv. Neural Inf. Process. Syst.* 2249–2257.

[29] Gregory, S. (2010). Finding overlapping communities in networks by label propagation. *New J. Phys.* **12** 103018.

[30] Handcock, M., Raftery, A. and Tantrum, J. (2007). Model-based clustering for social networks. *J. Roy. Statist. Soc. A* **170** 301–354.

[31] Horn, R. and Johnson, C. (1985). *Matrix Analysis*. Cambridge University Press.

[32] Hric, D., Darst, R. and Fortunato, S. (2014). Community detection in networks: Structural communities versus ground truth. *Phys. Rev. E* **90** 062805.

[33] Javadi, H. and Montanari, A. (2019). Nonnegative Matrix Factorization Via Archetypal Analysis. *Journal of the American Statistical Association* 1–22.

[34] Ji, P. and Jin, J. (2016). Coauthorship and citation networks for statisticians (with discussion). *Ann. Appl. Statist.* **10** 1779–1812.

[35] Jin, J. (2015). Fast community detection by SCORE. *Ann. Statist.* **43** 57–89.

[36] Jin, J. and Ke, Z. T. (2017). A sharp lower bound for mixed-membership estimation. *arXiv:1709.05603*.

[37] Jin, J., Ke, Z. T. and Wang, W. (2017). Phase transitions for high dimensional clustering and related problems. *Ann. Statist.* **45** 2151–2189.

[38] Jin, J. and Wang, W. (2016). Influential Features PCA for high dimensional clustering (with discussion). *Ann. Statist.* **44** 2323–2359.

[39] Karrer, B. and Newman, M. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* **83** 016107.

[40] Ke, Z. T. and Wang, M. (2017). A new SVD approach to optimal topic estimation. *arXiv:1704.07016*.

[41] Krebs, V. (2004). *Unpublished*.

[42] Lancichinetti, A., Radicchi, F., Ramasco, J. and Fortunato, S. (2011). Finding statistically significant communities in networks. *PloS One* **6** e18961.

[43] Latouche, P., Birmelé, E. and Ambroise, C. (2011). Overlapping stochastic block models with application to the french political blogosphere. *Ann. Appl. Statist.* **5** 309–336.

[44] Lee, D. and Seung, S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* **401** 788–791.

[45] Li, J. and Bioucas-Dias, J. M. (2008). Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International* **3** III–250. IEEE.

[46] Mao, X., Sarkar, P. and Chakrabarti, D. (2017). Estimating mixed memberships with sharp eigenvector deviations. *arXiv:1709.00407*.

[47] Mørup, M. and Hansen, L. K. (2012). Archetypal analysis for machine learning and data mining. *Neurocomputing* **80** 54–63.

[48] Nascimento, J. M. and Dias, J. M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing* **43** 898–910.

[49] Neville, R. (1999). Automatic endmember extraction from hyperspectral data for mineral exploration. In *International Airborne Remote Sensing Conference and Exhibition, 4 th/21 st Canadian Symposium on Remote Sensing, Ottawa, Canada*.

[50] Panov, M., Slavnov, K. and Ushakov, R. (2017). Consistent estimation of mixed memberships with successive projections. In *International Conference on Complex Networks and their Applications* 53–64. Springer.

[51] Peng, L. and Carvalho, L. (2016). Bayesian degree-corrected stochastic blockmodels for community detection. *Electron. J. Stat.* **10** 2746–2779.

[52] Preparata, F. and Shamos, M. (2012). *Computational geometry: an introduction*. Springer Science & Business Media.

[53] Qin, T. and Rohe, K. (2013). Regularized spectral clustering under the degree-corrected stochastic blockmodel. In *Adv. Neural Inf Process. Syst.* 3120–3128.

[54] Saldana, D. F., Yu, Y. and Feng, Y. (2017). How many communities are there? *J. Comp. Graph. Stat.* **26** 171–181.

[55] Tao, X., Wang, B., Zhang, L. and Zhang, J. Q. (2007). A new scheme for decomposition of mixed pixels based on nonnegative matrix factorization. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International* 1759–1762. IEEE.

[56] Tulino, A. and Verdu, S. (2004). *Radom matrix theory and wireless communications* **1**. Now Publisher Inc.

[57] Verzelen, N. and Arias-Castro, E. (2017). Detection and feature selection in sparse mixture models. *Ann. Statist.* **45** 1920–1950.

[58] Winter, M. E. (1999). N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation* 266–275. International Society for Optics and Photonics.

[59] Zhang, A. Y. and Zhou, H. H. (2016). Minimax rates of community detection in stochastic block models. *Ann. Statist.* **44** 2252–2280.

[60] Zhang, Y., Levina, E. and Zhu, J. (2014). Detecting overlapping communities in networks with spectral methods. *arXiv:1412.3432*.

[61] Zhao, Y., Levina, E. and Zhu, J. (2012). Consistency of community detection in networks under degree-corrected stochastic block models. *Ann. Statist.* **40** 2266–2292.

Submitted to the Annals of Statistics

SUPPLEMENT OF "ESTIMATING NETWORK MEMBERSHIPS BY SIMPLEX VERTEX HUNTING"

BY JIASHUN JIN[‡], ZHENG TRACY KE[§], AND SHENGMING LUO[‡]

Carnegie Mellon University[‡] *and Harvard University* [§]

CONTENTS

APPENDIX A: IDENTIFIABILITY AND REGULARITY CONDITIONS

We prove Proposition 1.1, which is about the identifiability of the DCMM model. We then give sufficient conditions for the regularity condition (2.16).

A.1. Proof of Proposition 1.1. Let $G = K\|\theta\|^{-2}\Pi'\Theta^2\Pi$ be the same as in Section 2. We consider two cases: (1) PG is an irreducible matrix. (2) PG is a reducible matrix.

First, we study Case (1). When PG is irreducible, the matrix R is well-defined (see the proof of Theorem 1.1). Additionally, by (1.8)-(1.9) in the paper and derivation therein, there exists the Ideal Simplex, which is uniquely determined by the eigenvectors $\xi_1, \xi_2, \ldots, \xi_K$ of Ω. Note that for either (Θ, Π, P) or $(\tilde{\Theta}, \tilde{\Pi}, \tilde{P})$, we have an Ideal Simplex. The two Ideal Simplexes can be different only when there are multiple choices of $\xi_1, \xi_2, \ldots, \xi_K$. By Lemma B.4, the first eigenvalue of Ω has a multiplicity 1, so by basic linear algebra, $[\xi_1, \xi_2, \ldots, \xi_K]$ are uniquely defined up to a rotation matrix of the form

$$\begin{bmatrix} a & 0 \\ 0 & S \end{bmatrix}, \quad \text{where } a \in \{-1, 1\} \text{ and } S \in \mathbb{R}^{K-1,K-1} \text{ is an orthogonal matrix.}$$

Recalling $R = [\text{diag}(\xi_1)]^{-1}[\xi_2, \xi_3, \ldots, \xi_K]$, it is seen that the property of "a row of R falls on one of the vertices of the Ideal Simplex" is invariant to the above rotation. Therefore, a row of Π equals to the corresponding row of $\tilde{\Pi}$, as long as one of them is pure.

We now proceed to showing $(\Theta, \Pi, P) = (\tilde{\Theta}, \tilde{\Pi}, \tilde{P})$. By the above argument and that each community has at least one pure node, we assume without loss of generality that for $1 \leq k \leq K$, the k-th node is a pure node in community k. Comparing the first K rows and the first K columns of $\Theta\Pi P\Pi'\Theta$ with those of $\tilde{\Theta}\tilde{\Pi}\tilde{P}\tilde{\Pi}'\tilde{\Theta}'$, it follows that

$$\text{diag}(\theta_1, \ldots, \theta_K) \cdot P \cdot \text{diag}(\theta_1, \ldots, \theta_K) = \text{diag}(\tilde{\theta}_1, \ldots, \tilde{\theta}_K) \cdot \tilde{P} \cdot \text{diag}(\tilde{\theta}_1, \ldots, \tilde{\theta}_K).$$

As both P and $\tilde{P}$ have unit diagonal entries, $P = \tilde{P}$ and $\theta_k = \tilde{\theta}_k$, $1 \leq k \leq K$.

Moreover, note that $P\Pi'\Theta$ has a full row-rank. Since $\Theta\Pi P\Pi'\Theta = \tilde{\Theta}\tilde{\Pi}\tilde{P}\tilde{\Pi}'\tilde{\Theta}$, it is seen that $\Theta\Pi = \tilde{\Theta}\tilde{\Pi}\Delta$, where $\Delta = \tilde{P}\tilde{\Pi}'\tilde{\Theta}X'(XX')^{-1}$, with $X = P\Pi'\Theta$ for short. We compare the first K rows of $\Theta\Pi$ and $\tilde{\Theta}\tilde{\Pi}\Delta$, recalling that the first K rows are pure and that $\theta_k = \tilde{\theta}_k$ for $1 \leq k \leq K$. It follows that Δ equals to the $K \times K$ identity matrix. Therefore,

$$\Theta\Pi = \tilde{\Theta}\tilde{\Pi}.$$

Since each row of Π or $\tilde{\Pi}$ is a PMF, $\Theta = \tilde{\Theta}$, $\Pi = \tilde{\Pi}$, and the claim follows.

Next, we study Case (2). By (1.7) in the paper,

$$\Xi = \Theta\Pi B, \qquad \text{for a non-singular matrix } B.$$

Row i of Ξ equals to θ_i times a convex combination of rows of B. It follows that *all rows of Ξ are contained in a simplicial cone with K supporting rays, where a pure row falls on one supporting ray, and a mixed row falls in the interior of the simplicial cone*. Note that Ξ is uniquely defined up to a $K \times K$ orthogonal matrix. The effect of this orthogonal matrix is to simultaneously rotate all rows of Ξ. Such a rotation does not change the property that "a pure row falls on one supporting ray". Therefore, a row of Π equals to the corresponding row of $\tilde{\Pi}$, provided that one of them is pure. The remaining of the proof is similar to Case (1).

Remark. Compared to other models (e.g., MMSB, DCBM), DCMM has many more parameters (for degree heterogeneity and for mixed memberships). These parameters have more degrees of freedom than those in MMSB or DCBM, and so DCMM requires stronger conditions to be identifiable.

- The assumption that P has unit diagonals is not needed for the identifiability of MMSB, but it is necessary for the identifiability of DCMM. Consider a DCMM with parameters (Θ, Π, P). Given any $K \times K$ diagonal matrix D with positive diagonals, let

  $$\tilde{P} = DPD, \quad \tilde{\pi}_i = (D^{-1}\pi_i)/\|D^{-1}\pi_i\|_1, \quad \text{and} \quad \tilde{\theta}_i = \|D^{-1}\pi_i\|_1 \cdot \theta_i.$$

  It is seen that $\Theta\Pi P\Pi'\Theta = \tilde{\Theta}\tilde{\Pi}\tilde{P}\tilde{\Pi}'\tilde{\Theta}$. This case will be eliminated by requiring P to have unit diagonals.
- The assumption that P has a full rank is not needed for the identifiability of DCBM, but is necessary for the identifiability of DCMM. If the rank of P is $< K$, there exists a nonzero vector $\beta \in \mathbb{R}^K$ such that

  $$P\beta = 0.$$

  As long as there is a π_i such that $\pi_i(k) > 0$ for all k, we can change (π_i, θ_i) to $(\tilde{\pi}_i, \tilde{\theta}_i)$ but keep Ω unchanged. This is done by letting

  $$\tilde{\pi}_i = (\pi_i + \epsilon\beta)/\|\pi_i + \epsilon\beta\|_1, \quad \text{and} \quad \tilde{\theta}_i = \|\pi_i + \epsilon\beta\|_1 \cdot \theta_i,$$

  for a sufficiently small $\epsilon > 0$. Since the two vectors, $\theta_i \cdot P\pi_i$ and $\tilde{\theta}_i \cdot P\tilde{\pi}_i$, are equal, Ω remains unchanged.

A.2. The condition (2.16). The following proposition gives three cases where (2.16) is satisfied. Below, for a matrix M, let $\lambda_k(M)$ denote the k-th largest eigenvalue in magnitude.

PROPOSITION A.1. *Consider a DCMM model where* $\Omega = \Theta\Pi P\Pi'\Theta$ *and* $\|P\|_{\max} \leq C$. *Write* $G = K\|\theta\|^{-2}(\Pi'\Theta^2\Pi)$. *Let* η_1 *be the first (unit-norm) right singular vector of* PG. *As* $n \to \infty$, *suppose at least one of the following conditions hold, where* $c > 0$ *is a constant:*

- $\min_{1\leq k,\ell\leq K} P(k,\ell) \geq c$, *and* $\min_k\{\sum_{i=1}^n \theta_i^2\pi_i(k)\} \geq c\max_k\{\sum_{i=1}^n \theta_i^2\pi_i(k)\}$.
- K *is fixed,* $\min_k G(k,k) \geq c$, *and* $|\lambda_1(PG)| \geq c+|\lambda_2(PG)|$. *For a fixed irreducible matrix* P_0, $\|P - P_0\| \to 0$.
- K *is fixed, and* $|\lambda_1(PG)| \geq c+|\lambda_2(PG)|$. *For a fixed irreducible matrix* G_0, $\|G - G_0\| \to 0$.

Then, we can select the sign of η_1 *such that all its entries are strictly positive. Furthermore,* $[\max_{1\leq k\leq K}\eta_1(k)]/[\min_{1\leq k\leq K}\eta_1(k)] \leq C$.

Proof of Proposition A.1: Consider the first case. Let $x_k = K\|\theta\|^{-2}\sum_{i=1}^n \theta_i^2\pi_i(k)$. It is seen that $\sum_{k=1}^K x_k = K$. The condition says that $\min_k x_k \geq c\max_k x_k$. Therefore, $x_k \asymp 1$ for all k. Moreover, by definition and direct calculations, $\sum_{\ell=1}^K G(\ell,k) = K\|\theta\|^{-2}\sum_{\ell=1}^K\sum_{i=1}^n \theta_i^2\pi_i(\ell)\pi_i(k) = x_k$. It follows that

$$\max_k\Big\{\sum_\ell G(\ell,k)\Big\} \asymp \min_k\Big\{\sum_\ell G(\ell,k)\Big\} \asymp 1.$$

For any $1 \leq m,k \leq K$, the (m,k)-th entry of PG equals to $\sum_\ell P(m,\ell)G(\ell,k)$, which is between $c\sum_\ell G(\ell,k)$ and $C\sum_\ell G(\ell,k)$ by the assumption on P. It follows that

$$\max_{k,\ell}\big\{(PG)(k,\ell)\big\} \asymp \min_{k,\ell}\big\{(PG)(k,\ell)\big\} \asymp 1. \tag{A.40}$$

In particular, PG is a positive matrix. By Perron's theorem [31, Theorem 8.2.8], the first right singular value $\lambda_1(PG)$ is positive and has a multiplicity of 1, and the first eigenvector η_1 is a positive vector. Write $\lambda = \lambda_1(PG)$ for short. By definition,

$$\lambda\eta_1 = (PG)\eta_1.$$

It follows that

$$\max_k \eta_1(k) \leq \frac{\|\eta\|_1}{\lambda}\max_{k,\ell}\{(PG)(k,\ell)\}, \quad \min_k \eta_1(k) \geq \frac{\|\eta\|_1}{\lambda}\min_{k,\ell}\{(PG)(k,\ell)\}. \tag{A.41}$$

Combining (A.40)-(A.41) gives $\max_k \eta_1(k) \asymp \min_k \eta_1(k)$. The claim follows.

Consider the second case. We first state and prove a useful result:

(A.42) Let A and B be two nonnegative matrices with strictly positive diagonals. If A is irreducible, then AB is irreducible.

The proof uses the definition of primitive matrices (a subclass of irreducible matrices; see [31, Section 8.5]). We aim to show AB is a primitive matrix. By [31, Theorem 8.5.2], it suffices to show that there exists $m \geq 1$, such that $(AB)^m$ is a strictly positive matrix. By the assumption, A is an irreducible matrix with positive diagonals; it follows from [31, Theorem 8.5.4] that A is a primitive matrix. By [31, Theorem 8.5.2] again, there exists $m \geq 1$ such that A^m is a strictly positive matrix. Let $\alpha > 0$ be the minimum diagonal entry of B. Since A and B are nonnegative matrices, each entry of $(AB)^m$ is lower bounded by α^m times the corresponding entry of A^m; hence, $(AB)^m$ is also a strictly positive matrix. It follows that AB is a primitive matrix, which is also an irreducible matrix.

We then show the claim. Note that P and G are both nonnegative matrices with positive entries. Since $\|P - P_0\| \to 0$, the support of P has to be a superset of the support of P_0; as a result, when P_0 is an irreducible matrix, P has to be an irreducible matrix. We apply (A.42) to obtain that PG is an irreducible matrix. It follows that $\lambda_1(PG) > 0$ and it has a multiplicity 1; additionally, the first right eigenvector η_1 is a positive vector.

It remains to show $\max_k \eta_1(k) \asymp \min_k \eta_1(k)$. We prove by contradiction. Write $\eta_1 = \eta_1^{(n)}$, $P = P^{(n)}$ and $G = G^{(n)}$ to emphasize the dependence on n. If the claim is not true, then there is a subsequence $\{n_s\}_{s=1}^{\infty}$ such that

$$\lim_{s\to\infty} \left\{ \frac{\min_k \eta_1^{(n_s)}(k)}{\max_k \eta_1^{(n_s)}(k)} \right\} \to 0. \tag{A.43}$$

Since K is fixed, all the entries of $G^{(n_s)}$ are bounded. It follows that there exists a subsequence of $\{n_s\}_{s=1}^{\infty}$, which we still denote by $\{n_s\}_{s=1}^{\infty}$ for notation convenience, such that $G^{(n_s)} \to G^*$ for a fixed matrix G^*. Therefore,

$$\left\| (PG)^{(n_s)} - P_0 G^* \right\| \to 0, \qquad \text{as } s \to \infty. \tag{A.44}$$

Let η_1^* be the first right eigenvector of $P_0 G^*$. Since $|\lambda_1(PG)| \geq c + |\lambda_2(PG)|$, by the sin-theta theorem (e.g., see Lemma C.2), it follows from (A.44) that

$$\|\eta_1^{(n_s)} - \eta_1^*\| \to 0, \qquad \text{as } s \to \infty. \tag{A.45}$$

We now derive a contradiction from (A.43)-(A.45). On the one hand, combining (A.44)-(A.45) and noting that η_1^* is a fixed vector, we conclude that the minim entry of η_1^* is zero. On the other hand, the assumption of $\min_k G(k,k) \geq c$ ensures that G^* has strictly positive diagonals. We apply (A.42) to conclude that $P_0 G^*$ is a fixed irreducible matrix. By Perron's theorem, η_1^* should be a strictly positive vector. This yields a contradiction.

Consider the third case. The proof is similar to that of the second case, except that we switch the roles of P and G. Note that we do not need additional conditions on the diagonals of P, since P always has unit diagonals. $\square$

APPENDIX B: THE ORACLE CASE AND IDEAL MIXED-SCORE

We consider the oracle case where Ω is observed. In Section B.1, we study eigenvalues and eigenvectors of Ω, as well as the matrix R. In Sections B.2-B.3, we prove the main results about Ideal Mixed-SCORE.

B.1. Spectral analysis of Ω. First, we study the leading eigenvalues of Ω. Let $\lambda_1, \ldots, \lambda_K$ be all the nonzero eigenvalues of Ω, sorted descendingly in magnitude. Let $a_1, a_2, \ldots, a_K$ be all eigenvalues of $PG \in \mathbb{R}^{K,K}$, sorted descendingly in magnitude. It turns out that (see Lemma B.4) a_k's are real and

$$\lambda_k = a_k \cdot K^{-1}\|\theta\|^2, \qquad k = 1, 2, \ldots, K.$$

As a result, we can prove the following lemmas:

LEMMA B.1. *Under conditions of Theorem 2.1, the following statements are true:*

- $C^{-1}K^{-1}\|\theta\|^2 \leq \lambda_1 \leq C\|\theta\|^2$. *If* $\beta_n = o(1)$, *then* $\lambda_1 \asymp \|\theta\|^2$.
- $\lambda_1 - |\lambda_2| \asymp \lambda_1$.
- $|\lambda_k| \asymp \beta_n K^{-1}\|\theta\|^2$, *for* $2 \leq k \leq K$.

Next, we study the leading eigenvectors of Ω. Let $\xi_1, \ldots, \xi_K$ be the eigenvectors of Ω associated with $\lambda_1, \ldots, \lambda_K$, respectively. Write $\Xi_0 = [\xi_2, \cdots, \xi_K] \in \mathbb{R}^{n,K-1}$, and let $\Xi_{0,i}'$ be its i-th row, $1 \leq i \leq n$. According to (1.7), there exists a non-singular matrix $B = [b_1, b_2, \ldots, b_K] \in \mathbb{R}^{K,K}$ such that

$$\xi_k = \Theta\Pi b_k, \qquad k = 1, 2, \ldots, K.$$

Moreover, it can be shown that the entries of B are $O(\|\theta\|^{-1})$ (see the proof of Lemma B.2). We have the following lemma:

LEMMA B.2. *Under conditions of Theorem 2.1, the following statements are true:*

- *If we choose the sign of* ξ_1 *such that* $\sum_{i=1}^n \xi_1(i) > 0$, *then the entries of* ξ_1 *are positive satisfying* $C^{-1}\theta(i)/\|\theta\| \leq \xi_1(i) \leq C\theta(i)/\|\theta\|$, $1 \leq i \leq n$.
- $\|\Xi_{0,i}\| \leq C\sqrt{K}\theta(i)/\|\theta\|$, $1 \leq i \leq n$.

Last, we study the entry-wise ratio matrix R. Recall that $w_i \in \mathbb{R}^K$ is the vector of coefficients when r_i is expressed as a convex combination of the vertices of the Ideal Simplex; see (1.10).

LEMMA B.3. *Under conditions of Theorem 2.1, the following statements are true:*

- *The vertices of the Ideal Simplex satisfy that* $\max_{1\leq k\leq K}\|v_k\| \leq C\sqrt{K}$ *and* $\min_{k\neq \ell}\|v_k - v_\ell\| \geq C\sqrt{K}$.
- $C^{-1}\|\pi_i - \pi_j\|_1 \leq \|w_i - w_j\|_1 \leq C\|\pi_i - \pi_j\|_1$, *for all* $1 \leq i, j \leq n$.
- $C^{-1}\sqrt{K}\|w_i - w_j\| \leq \|r_i - r_j\| \leq C\sqrt{K}\|w_i - w_j\|$, *for all* $1 \leq i, j \leq n$.

Below, we prove Lemmas B.1-B.3.

B.1.1. *A useful lemma and its proof.* Recall that B is the matrix in (1.7) and $G = K\|\theta\|^{-2}(\Pi'\Theta^2\Pi)$ is as in Section 2. Let (λ_k, ξ_k) be the k-th eigenpair of Ω. We state and prove a useful lemma:

LEMMA B.4. *Consider* $DCMM_n(K, P, \Theta, \Pi)$*, where* PG *is an irreducible matrix and there is at least one pure node for each community. The following statements are true:*

- *There is a non-singular matrix* $B \in \mathbb{R}^{K,K}$ *such that* $\Theta\Pi B = \Xi$*, and* B *is unique once* Ξ *is chosen.*
- *For* $1 \leq k \leq K$*, denote by* a_k *the* k*-th largest (in magnitude) eigenvalue of* PG*. Then,* a_k*'s are real, and the nonzero eigenvalues of* Ω *are* $\lambda_k = (K^{-1}\|\theta\|^2)a_k$, $1 \leq k \leq K$.
- *For* $1 \leq k \leq K$, b_k *is a (right) eigenvector of* PG *associated with* a_k.
- $\lambda_1 > 0$ *and it has a multiplicity* 1 *(so* ξ_1 *is uniquely determined up to a factor of* ± 1*).*
- ξ_1 *can be chosen such that all of its entries are positive. For this choice of* ξ_1*, all the entries of the associated* b_1 *are also positive.*

Proof of Lemma B.4: Consider the first claim. Denote by $Span(\cdot)$ the column space of a matrix. It suffices to show that $Span(\Theta\Pi) = Span(\Xi)$. Then, since $\xi_1, \cdots, \xi_K$ form an orthonormal basis of this subspace, there is a unique, non-singular matrix $\tilde{B}$ such that $\Theta\Pi = \Xi\tilde{B}$. We then take $B = \tilde{B}^{-1}$.

We now show $Span(\Theta\Pi) = Span(\Xi)$. By the assumption that there is at least one pure node in each community, we can find K rows of Π such that they form a $K \times K$ identity matrix. So Π has a rank K. Since Θ and P are both non-singular matrices, Ω also has a rank K. By definition, $\Omega\xi_k = \lambda_k\xi_k$, for $1 \leq k \leq K$. It follows that

$$\Theta\Pi(P\Pi'\Theta\xi_k) = \lambda_k\xi_k.$$

Hence, each ξ_k is in the column space of $\Theta\Pi$. This means the column space of Ξ is contained in the column space of $\Theta\Pi$. Since both matrices have a rank K, the two column spaces are the same.

Consider the second claim. Note that P is symmetric and G is positive definite. Let $G^{1/2}$ be the unique square root of G. For any matrices $A \in \mathbb{R}^{m,n}$ and $B \in \mathbb{R}^{n,m}$, if $m \geq n$, then the nonzero eigenvalues of AB are the same as the nonzero eigenvalues of BA [31, Theorem 1.3.22]. As a result, eigenvalues of PG are the same as eigenvalues of the symmetric matrix $G^{1/2}PG^{1/2}$. It implies that $a_1, a_2, \ldots, a_K$ are real.

Furthermore, the nonzero eigenvalues of $\Omega = (\Theta\Pi)(P\Pi'\Theta)$ are the same as the nonzero eigenvalues of $(P\Pi'\Theta)(\Theta\Pi) = (K^{-1}\|\theta\|^2)(PG)$. Hence, the nonzero eigenvalues of Ω are $(K^{-1}\|\theta\|^2)a_1, (K^{-1}\|\theta\|^2)a_2, \ldots, (K^{-1}\|\theta\|^2)a_K$.

Consider the third claim. Write $\tilde{G} \equiv K^{-1}\|\theta\|^2 G = \Pi'\Theta^2\Pi$. We note that $\Omega\xi_k = \lambda_k\xi_k$ and $\xi_k = \Theta\Pi b_k$. Therefore, $(\Theta\Pi P\Pi'\Theta)(\Theta\Pi b_k) = \lambda_k(\Theta\Pi b_k)$. Multiplying both sides by $\Pi'\Theta$ from the left, we have

$$\tilde{G}P\tilde{G}b_k = \lambda_k\tilde{G}b_k$$

Since $\tilde{G}$ is non-singular, $P\tilde{G}b_k = \lambda_k b_k$. Plugging in $\tilde{G} = (K^{-1}\|\theta\|^2)G$ and $\lambda_k = (K^{-1}\|\theta\|^2)a_k$, we obtain $PGb_k = a_k b_k$. This shows that b_k is a (right) eigenvector of PG associated with a_k. Additionally, since η_1 is the first unit-norm right singular vector of PG, it yields that $\eta_1 = b_1/\|b_1\|$.

Consider the fourth claim. Since $\lambda_1 = (K^{-1}\|\theta\|^2)a_1$, it suffices to show that $a_1 > 0$ and that it has a multiplicity 1. This follows immediately from the Perron-Frobenius theorem [31, Theorem 8.4.4] and the assumption that PG is an irreducible matrix.

Consider the last claim. Note that b_1 is the eigenvalue of PG associated with a_1. Since a_1 has a multiplicity 1, $b_1/\|b_1\|$ is unique up to a factor of ± 1 (depending on the choice of ξ_1). By Perron-Frobenius theorem again, $b_1/\|b_1\|$ can be chosen such that all the entries are positive. The associated $\xi_1 = \Theta\Pi b_1$, where $\Theta\Pi$ is a nonnegative matrix with positive row sums. So all the entries of ξ_1 are also positive.

B.1.2. *Proof of Lemma B.1.* By Lemma B.4, all nonzero eigenvalues of Ω are $(K^{-1}\|\theta\|^2)a_1, \ldots, (K^{-1}\|\theta\|^2)a_K$, where a_k is the k-th largest eigenvalue (in magnitude) of PG. By the condition (2.15),

$$a_1 - |a_2| \geq C^{-1}a_1, \qquad C^{-1}\beta_n \leq |a_K| \leq |a_2| \leq C\beta_n.$$

The second and third claims follow immediately.

We then study a_1. For any two matrices A and B, the nonzero eigenvalues of AB are the same as the nonzero eigenvalues of BA. Hence,

$$a_1 = \lambda_1(PG) = \lambda_1(G^{1/2}PG^{1/2}) = \max_{x \neq 0} \frac{x'G^{1/2}PG^{1/2}x}{\|x\|^2}.$$

By (2.14), $\|G\| \leq C$ and $\|G^{-1}\| \leq C$. It is easy to see that $a_1 \leq C\lambda_1(P)$. Additionally, $\lambda_1(P) = \max_{y \neq 0} \frac{y'Py}{\|y\|^2} = \max_{x \neq 0} \frac{x'G^{1/2}PG^{1/2}x}{\|G^{1/2}x\|^2}$. Since $\|G^{1/2}x\|^2 = x'Gx \geq C^{-1}\|x\|^2$, it follows that $\lambda_1(P) \leq \max_{x \neq 0} \frac{x'G^{1/2}PG^{1/2}x}{C^{-1}\|x\|^2} \leq C\lambda_1(PG)$. Together,

$$C^{-1}\lambda_1(P) \leq \lambda_1(PG) \leq C\lambda_1(P).$$

Note that $\lambda_1(P) \leq K\|P\|_{\max} = O(K)$ and $\lambda_1(P) \geq P(k,k) \geq 1$. We plug them into the above inequality to get

$$C^{-1} \leq a_1 \leq CK \tag{B.46}$$

This inequality holds in all cases. If, additionally, $\beta_n \to 0$ as $n \to \infty$, we can get a stronger result. Note that P and G are nonnegative matrices, and for each $1 \leq k \leq K$, $P(k,k) = 1$ and $G(k,k) \geq \lambda_{\min}(G) \geq C^{-1}$. It follows that $(PG)(k,k) \geq P(k,k)G(k,k) \geq C^{-1}$. We thus have

$$\text{trace}(PG) \geq C^{-1}K.$$

At the same time, $\text{trace}(PG) = a_1 + \sum_{k=2}^K a_2 = a_1 + O(K\beta_n) = a_1 + o(K)$. It follows that

$$C^{-1}K \leq a_1 \leq CK, \qquad \text{if } \beta_n = o(1). \tag{B.47}$$

The first claim follows from (B.46) and (B.47).

B.1.3. *Proof of Lemma B.2.* Consider the first claim. From the last item of Lemma B.4, we can choose the sign of ξ_1 such that both (ξ_1, b_1) have strictly positive entries, where this choice of sign corresponds to $\sum_{i=1}^n \xi_1(i) > 0$. Note that $\Xi = \Theta\Pi B$, which implies $\xi_1(i) = \theta(i)\sum_{k=1}^K \pi_i(k)b_1(k)$. Since each π_i is a PMF,

$$\theta(i) \min_{1 \leq k \leq K} b_1(k) \leq \xi_1(i) \leq \theta(i) \max_{1 \leq k \leq K} b_1(k), \qquad 1 \leq i \leq n.$$

Hence, to show the claim, it suffices to show that

$$C^{-1}\|\theta\|^{-1} \leq b_1(k) \leq C\|\theta\|^{-1}, \qquad \text{for all } 1 \leq k \leq K. \tag{B.48}$$

Write $\tilde{G} = K^{-1}\|\theta\|^2 G = \Pi'\Theta^2\Pi$. Since $\Xi = \Theta\Pi B$, we have $B'\Pi'\Theta^2\Pi B = I_K$, or equivalently, $B'\tilde{G}B = I_K$. Multiplying both sides by B from the left and B' from the right, we obtain $BB'\tilde{G}BB' = BB'$. Since BB' is non-singular, it implies

$$BB' = \tilde{G}^{-1} = K\|\theta\|^{-2}G^{-1}. \tag{B.49}$$

We note that $BB' = \sum_{k=1}^{K} b_k b_k' \succeq b_1 b_1'$. So, $\|b_1\|^2 \leq \|B\|^2 \leq K\|\theta\|^{-2}\|G^{-1}\|$. By our assumption of $\|G^{-1}\| \leq C$. It follows that

$$\|b_1\| \leq C\|\theta\|^{-1}\sqrt{K}.$$

At the same time, $1 = \|\xi_1\|^2 = \|\Theta\Pi b_1\|^2$. By direct calculations, $\|\Theta\Pi b_1\|^2 = \sum_i \theta_i^2(\pi_i' b_1)^2 \leq \sum_i \theta_i^2 \|b_1\|_\infty^2 \leq \|\theta\|^2\|b_1\|_\infty^2$. It follows that

$$\|b_1\|_\infty \geq C^{-1}\|\theta\|^{-1}.$$

In Lemma B.4, we have seen that b_1 is the first right singular vector of PG. Hence, $b_1 \propto \eta_1$, where η_1 is the same as in (2.16). By (2.16), all the entries of η_1 are at the same order. So, all the entries of b_1 are at the same order. It follows that $b_1(k) \asymp \|b_1\|_\infty \asymp (1/\sqrt{K})\|b_1\|$. Hence, (B.48) follows from the above inequalities for $\|b_1\|$ and $\|b_1\|_\infty$.

Consider the second claim. Since $\Xi = \Theta\Pi B$, for $1 \leq i \leq n$,

$$\|\Xi_{0,i}\| \leq \theta(i)\|B\pi_i\| \leq C\theta(i)\sqrt{\lambda_{\max}(B'B)} \leq C\sqrt{K}\|\theta\|^{-1}\theta(i),$$

where the last inequality is due to (B.49) and and the condition $\|G^{-1}\| \leq C$.

B.1.4. *Proof of Lemma B.3.* First, we prove the claim about the connection between $\|w_i - w_j\|_1$ and $\|\pi_i - \pi_j\|_1$. Let $\mathcal{S}_0 = \mathcal{S}_0(e_1, e_2, \ldots, e_K) \subset \mathbb{R}^K$ be the standard $(K-1)$-simplex. Define a mapping

$$T_1 : \mathcal{S}_0 \to \mathcal{S}_0, \qquad \text{where} \quad T_1(x) = \frac{x \circ b_1}{\|x \circ b_1\|_1}.$$

By (1.10), $w_i = T_1(\pi_i)$, for $1 \leq i \leq n$. To show the claim, it suffices to show that T_1 and T_1^{-1} are both Lipschitz with respect to the ℓ^1-norm, i.e., for any $x, y \in \mathcal{S}_0$,

$$C^{-1}\|x - y\|_1 \leq \|T_1(x) - T_1(y)\|_1 \leq C\|x - y\|_1. \tag{B.50}$$

We now show (B.50). Fixing any $x, y \in \mathcal{S}_0$, write $x^* = T_1(x)$ and $y^* = T_1(y)$. By definition, $x^*(k) = x(k)b_1(k)/\|x \circ b_1\|_1$ and $y^*(k) = y(k)b_1(k)/\|y \circ b_1\|_1$. We write

$$x^*(k) - y^*(k) = \frac{[x(k) - y(k)]b_1(k)}{\|x \circ b_1\|_1} + y(k)b_1(k)\left[\frac{1}{\|x \circ b_1\|_1} - \frac{1}{\|y \circ b_1\|_1}\right]$$

$$= \frac{b_1(k)}{\|x \circ b_1\|_1}[x(k) - y(k)] + \frac{y^*(k)}{\|x \circ b_1\|_1}\left(\|y \circ b_1\|_1 - \|x \circ b_1\|_1\right).$$

First, by (B.48), $b_1(k) \asymp \|\theta\|^{-1}$ for all $1 \leq k \leq K$. It follows that $|b_1(k)| \leq C\|\theta\|^{-1}$ and $\|x \circ b_1\|_1 \geq \|x\|_1 \cdot C^{-1}\|\theta\|^{-1} \geq C^{-1}\|\theta\|^{-1}$. Hence,

$$\frac{b_1(k)}{\|x \circ b_1\|_1}|x(k) - y(k)| \leq C|x(k) - y(k)|.$$

Second, by the triangle inequality, $|\|y \circ b_1\|_1 - \|x \circ b_1\|_1| \leq \|(y - x) \circ b_1\|_1$. Moreover, since $b_1(k) \asymp \|\theta\|^{-1}$ for all k, we have $\|(y-x)\circ b_1\|_1 \leq C\|\theta\|^{-1}\|x-y\|_1$ and $\|x \circ b_1\|_1 \geq C^{-1}\|\theta\|^{-1}$. It follows that

$$\frac{y^*(k)}{\|x \circ b_1\|_1}\left|\|y \circ b_1\|_1 - \|x \circ b_1\|_1\right| \leq Cy^*(k) \cdot \|x - y\|_1.$$

Combining the above gives

$$|x^*(k) - y^*(k)| \leq C|x(k) - y(k)| + Cy^*(k) \cdot \|x - y\|_1.$$

We sum over k on both sides and note that $\sum_k y^*(k) = 1$. It gives

$$\|x^* - y^*\|_1 \leq C\|x - y\|_1.$$

This shows that T_1 is Lipschitz with respect to the ℓ^1-norm. We then consider T_1^{-1}. Define $\tilde{b}_1 \in \mathbb{R}^K$ by $\tilde{b}_1(k) = 1/b_1(k)$, $1 \leq k \leq K$. We can rewrite

$$T_1^{-1}(x) = \frac{x \circ \tilde{b}_1}{\|x \circ \tilde{b}_1\|_1}.$$

T_1^{-1} has a similar form as T_1, where the vector $\tilde{b}_1$ satisfies that $\tilde{b}_1(k) \asymp \|\theta\|$ for all k. Hence, we can similarly prove that T_1^{-1} is Lipschitz with respect to the ℓ^1-norm. This proves (B.50).

Next, we prove the claim about the connection between $\|r_i - r_j\|$ and $\|w_i - w_j\|$. Let $\mathcal{S}_0$ be the same as before, and let $\mathcal{S}^{ideal} = \mathcal{S}^{ideal}(v_1, v_2, \ldots, v_K) \subset \mathbb{R}^{K-1}$ denote the Ideal Simplex. Let $B = [b_1, b_2, \ldots, b_K]$ be as in (1.7). Define a mapping:

$$T_2 : \mathcal{S}_0 \to \mathcal{S}^{ideal}, \qquad \text{where} \quad \begin{pmatrix} 1 \\ T_2(x) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_K \end{pmatrix}}_{\equiv Q} x.$$

By Theorem 1.1, $r_i = T_2(w_i)$, for all $1 \leq i \leq n$. To show the claim, it suffices to show that T_2 and T_2^{-1} are both Lipschitz with respect to the ℓ^2-norm, whose Lipschitz constants are $\sqrt{K}$ and $1/\sqrt{K}$, respectively. In other words, we want to prove, for any $x, y \in \mathcal{S}_0$,

$$C^{-1}\sqrt{K}\|x-y\| \le \|T_2(x) - T_2(y)\| \le C\sqrt{K}\|x-y\|. \tag{B.51}$$

We now show (B.51). Since $1_K' x = 1_K' y = 1$, we have

$$\|T_2(x) - T_2(y)\|^2 = \|Qx - Qy\|^2 = (x-y)'Q'Q(x-y).$$

It suffices to show that

$$\|Q\| \le C\sqrt{K}, \qquad \text{and} \qquad \|Q^{-1}\| \le C/\sqrt{K}. \tag{B.52}$$

By (1.9), we can re-write

$$Q' = [\mathrm{diag}(b_1)]^{-1}B.$$

By (B.48), $b_1(k) \asymp \|\theta\|^{-1}$ for all k. By (B.49), $BB' = K\|\theta\|^{-2}G^{-1}$; we note that by condition (2.14), $\|G\| \le C$ and $\|G^{-1}\| \le C$; it follows that $\|B\| \le C\sqrt{K}\|\theta\|^{-1}$ and $\|B^{-1}\| \le C\|\theta\|/\sqrt{K}$. Combining them gives (B.52). Then, (B.51) follows.

Last, we prove the claims about the Ideal Simplex (IS). Let $e_1, e_2, \ldots, e_K$ be the standard basis vectors of $\mathbb{R}^K$. It is seen that $v_k = T_2(e_k)$, $1 \le k \le K$. By (B.51), for $k \neq \ell$,

$$\|v_k - v_\ell\| \asymp \sqrt{K}\|e_k - e_\ell\| \asymp \sqrt{K}.$$

By definition of Q and (B.52), for all $1 \le k \le K$,

$$\|v_k\| \le \|Q\| = O(\sqrt{K}).$$

The above give the desired claims.

B.2. Proof of Lemma 1.1. We have rigorously justified in Lemma B.4 that

$$\Xi = \Theta\Pi B, \qquad \text{for a non-singular matrix } B = [b_1, \ldots, b_K] \in \mathbb{R}^{K,K}.$$

Write $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_K)$. Then, $\Omega = \Xi\Lambda\Xi'$. First, plugging in $\Xi = \Theta\Pi B$, we find that $\Omega = \Theta\Pi(B\Lambda B')\Pi'\Theta$. Multiplying both sides by $\Pi'\Theta$ from the left and $\Theta\Pi$ from the right, we have $\Pi'\Theta\Omega\Theta\Pi = \tilde{G}(B\Lambda B')\tilde{G}$, where $\tilde{G} = \Pi'\Theta^2\Pi$ is a non-singular matrix. Second, since $\Omega = \Theta\Pi P\Pi'\Theta'$, we have $\Pi'\Theta\Omega\Theta\Pi = \tilde{G}P\tilde{G}$. Combining the above gives

$$\tilde{G}P\tilde{G} = \tilde{G}(B\Lambda B')\tilde{G} \quad \Longrightarrow \quad P = B\Lambda B'.$$

As a result,

$$1 = P(k,k) = \sum_{\ell=1}^{K} \lambda_\ell b_\ell^2(k) = b_1^2(k)[\lambda_1 + \sum_{\ell=2}^{K} \lambda_\ell^2 v_k(\ell-1)].$$

From the last item of Lemma B.4, the sign of ξ_1 can be chosen such that all entries of b_1 are positive. Hence, the claim follows.

B.3. Proof of Theorem 1.1. By the last two items of Lemma B.4, if we choose the sign of ξ_1 such that $\sum_{i=1}^n \xi_1(i) > 0$, then ξ_1 and b_1 are uniquely determined and have positive entries. This guarantees that $\{r_i\}_{i=1}^n$ and $\{v_k\}_{k=1}^K$ are well-defined.

To show that the Ideal Mixed-SCORE exactly recovers all the π_i, we first show that the simplex structure exists, i.e., (1.10) holds. Since $\Xi = \Theta\Pi B$, $\xi_\ell(i) = \theta(i)\sum_{k=1}^K \pi_i(k)b_\ell(k) = \theta(i)\|b_\ell \circ \pi_i\|_1$ for $1 \leq \ell \leq K$. It follows from $R(i,\ell) = \xi_{\ell+1}(i)/\xi_1(i)$ that

$$R(i,\ell) = \frac{\theta(i)\sum_{k=1}^K \pi_i(k)b_{\ell+1}(k)}{\theta(i)\|b_1 \circ \pi_i\|_1} = \sum_{k=1}^{K} \frac{b_1(k)\pi_i(k)}{\|b_1 \circ \pi_i\|_1} \cdot \frac{b_{\ell+1}(k)}{b_1(k)} = \sum_{k=1}^{K} w_i(k)v_k(\ell).$$

This yields that $r_i = \sum_{k=1}^K w_i(k)v_k = Vw_i$.

Once the simplex structure holds, by applying any convex hull algorithm to rows of R, we can exactly identify $v_1, \cdots, v_K$. Furthermore, for each i, we can recover w_i from r_i and $v_1, \cdots, v_K$ by solving the linear equations ($\mathbf{1}$ is the vector of 1's)

$$Vw_i = r_i, \quad \mathbf{1}'w_i = 1, \qquad \text{where } V \in \mathbb{R}^{K-1,K},\ w_i \in \mathbb{R}^K.$$

Last, Lemma 1.1 guarantees that b_1 can be exactly recovered from $\lambda_1, \ldots, \lambda_K$ and $v_1, \ldots, v_K$. Since $w_i \propto (b_1 \circ \pi_i)$ and π_i is a PMF, π_i is exactly recovered from $\pi_i = \pi_i^*/\|\pi_i^*\|_1$ where $\pi_i^*(k) = w_i(k)/b_1(k)$, $1 \leq k \leq K$.

APPENDIX C: THE REAL CASE AND SPECTRAL ANALYSIS

We consider the real case where A is observed. In Section C.1, we study the eigenvalues of A. In Section C.2, we study the eigenvectors of A and prove Lemma 2.1. In Section C.3, we study the matrix $\hat{R}$ and prove Theorem 2.1. In Section C.4, we state and prove two lemmas that will be used in the proof of the main paper.

C.1. Eigenvalues of A. Let $\hat{\lambda}_1, \ldots, \hat{\lambda}_K$ be the K largest eigenvalues of A (in magnitude), sorted descendingly in magnitude.

LEMMA C.1. *Under conditions of Theorem 2.1, with probability* $1 - o(n^{-3})$, $\max_{1 \le k \le K} |\hat{\lambda}_k - \lambda_k| \le C\sqrt{\theta_{\max}\|\theta\|_1}$.

Proof of Lemma C.1: By Weyl's inequality, $\max_{1\le k\le K} |\hat{\lambda}_k - \lambda_k| \le \|A - \Omega\|$. To show the claim, it suffices to show that with probability $1 - o(n^{-3})$,

$$\|A - \Omega\| \le C\sqrt{\theta_{\max}\|\theta\|_1}. \tag{C.53}$$

The following inequality is useful:

$$(\theta_{\max}\|\theta\|_1)/\log(n) \to \infty. \tag{C.54}$$

To prove (C.54), we rewrite $err_n = (\theta_{\max}/\theta_{\min})\|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1 \log(n)}$. Since $\theta_{\max} \ge \theta_{\min}$ and $\theta_{\max}\|\theta\|_1 \ge \|\theta\|^2$, we immediately have $err_n \ge \|\theta\|^{-1}\sqrt{\log(n)}$. Therefore, the assumption $err_n \to 0$ implies that $\|\theta\|^2/\log(n) \to \infty$. Then (C.54) is also true because $\theta_{\max}\|\theta\|_1 \ge \|\theta\|^2$.

We now prove (C.53). Write

$$A - \Omega = W + \mathrm{diag}(\Omega), \qquad \text{where } \; W \equiv A - E[A].$$

Note that $\pi_i' P \pi_j = \sum_{k,\ell} \pi_i(k)\pi_j(\ell)P_{k\ell} \le \|P\|_{\max}\|\pi_i\|_1\|\pi_j\|_1 \le C$. It follows that

$$\Omega(i,j) \le C\theta_i\theta_j.$$

Note that $\Omega(i,i) = \theta^2(i)(\pi_i' P \pi_i) \le C\theta^2(i)$. As a result,

$$\|\mathrm{diag}(\Omega)\| \le C\theta_{\max}^2 \le C\sqrt{\theta_{\max}\|\theta\|_1}, \tag{C.55}$$

where the last inequality is due to (C.54). We then apply the non-asymptotic bounds for random matrices in [7] to bound $\|W\|$. By Corollary 3.12 and Remark 3.13 of [7], for the $n \times n$ symmetric matrix W whose upper triangle contains independent entries, for any $\epsilon > 0$, there exists a universal constant $\tilde{c}_\epsilon > 0$ such that for every $t \ge 0$,

$$\mathbb{P}\big(\|W\| > (1+\epsilon)2\sqrt{2}\tilde{\sigma} + t\big) \le n e^{-t^2/(\tilde{c}\tilde{\sigma}_*^2)}, \tag{C.56}$$

where

$$\tilde{\sigma} = \max_i \sqrt{\sum_j \mathbb{E}[W(i,j)^2]}, \qquad \tilde{\sigma}_* = \max_{ij} \|W(i,j)\|_\infty.$$

It is seen that $\tilde{\sigma}^2 \leq \max_i\{\sum_j \Omega(i,j)\} \leq C \max_i\{\sum_j \theta(i)\theta(j)\} \leq C\theta_{\max}\|\theta\|_1$ and $\tilde{\sigma}_* \leq 1$. Taking $t = 2\sqrt{\tilde{c}\log(n)}$ for a large enough constant $\tilde{c} > 0$, we have that with probability $1 - o(n^{-3})$,

$$\|W\| \leq C\sqrt{\theta_{\max}\|\theta\|_1} + C\sqrt{\log(n)} \leq C\sqrt{\theta_{\max}\|\theta\|_1}, \tag{C.57}$$

where the last inequality is from (C.54). Combining (C.55) and (C.57) gives (C.53). □

C.2. Proof of Lemma 2.1. We first prove the first two items. The proof is based on the the classical sin-theta theorem [20], where below is a simpler version [14, Theorem 10].

LEMMA C.2. *Let M and $\hat{M}$ be two $n \times n$ symmetric matrices. For $1 \leq k \leq n$, let d_k be the k-th largest eigenvalue of M, η_k and $\hat{\eta}_k$ be the eigenvector associated with the k-th largest eigenvalue of M and $\hat{M}$, respectively. Suppose for some $\delta > 0$ and $1 \leq k_1 \leq k_2 \leq n$, we have $d_{k_1-1} > d_{k_1} + \delta$, $d_{k_2+1} < d_{k_2} - \delta$ and $\|\hat{G} - G\| \leq \delta/2$. Write $U = [\eta_{k_1}, \cdots, \eta_{k_2}]$ and $\hat{U} = [\hat{\eta}_{k_1}, \cdots, \hat{\eta}_{k_2}]$. Then, $\|\hat{U}\hat{U}' - UU'\| \leq 2\delta^{-1}\|\hat{G} - G\|$.*

We divide all eigenvalues of Ω into four groups: (i) λ_1, (ii) positive eigenvalues among $\lambda_2, \ldots, \lambda_K$, (iii) zero eigenvalues, and (iv) negative eigenvalues among $\lambda_2, \ldots, \lambda_K$. Define Ξ_{01} and Ξ_{02} as the submatrices of Ξ_0 by restricting to columns corresponding to eigenvalues in groups (ii) and (iv), respectively. By dividing the empirical eigenvalues and eigenvectors in a similar way, we can define $\hat{\Xi}_{01}$ and $\hat{\Xi}_{02}$. Now, ξ_1, Ξ_{01} and Ξ_{02} contain the eigenvectors associated with eigenvalues in groups (i), (ii) and (iv), respectively. By Lemma B.1, the gap between eigenvalues in group (i) and those in other groups is $\lambda_1 - |\lambda_2| \geq C^{-1}\lambda_1 \geq C^{-1}K^{-1}\|\theta\|^2$, and the eigen-gap between any two remaining groups is $\geq C\beta_n K^{-1}\|\theta\|^2$. It follows from Lemma C.2 that

$$\|\hat{\xi}_1\hat{\xi}_1' - \xi_1\xi_1'\| = O\Big(\frac{K\|A - \Omega\|}{\|\theta\|^2}\Big), \quad \max_{t\in\{1,2\}}\{\|\hat{\Xi}_{0t}\hat{\Xi}_{0t}' - \Xi_{0t}\Xi_{0t}'\|\} = O\Big(\frac{K\|A - \Omega\|}{\beta_n\|\theta\|^2}\Big). \tag{C.58}$$

By elementary linear algebra, $(\hat{\xi}_1\hat{\xi}_1' - \xi_1\xi_1')$ has two nonzero eigenvalues $\pm[1 - (\hat{\xi}_1'\xi_1)^2]^{1/2}$, where $|1 - (\hat{\xi}_1'\xi_1)^2| \geq \min_\pm |1 \pm \hat{\xi}_1'\xi_1| = (\min_\pm \|\hat{\xi}_1 \pm \xi_1\|^2)/2$. It follows that

$$\min_\pm \|\hat{\xi}_1 \pm \xi_1\| \leq \sqrt{2}\|\hat{\xi}_1\hat{\xi}_1' - \xi_1\xi_1'\|. \tag{C.59}$$

Moreover, by [38, Lemma 2.4], there always is an orthogonal matrix X_1 such that $\|\hat{\Xi}_{01} - \Xi_{01}X_1\|_F \leq \|\hat{\Xi}_{01}\hat{\Xi}_{01}' - \Xi_{01}\Xi_{01}'\|_F$. Since the rank of $(\hat{\Xi}_{01}\hat{\Xi}_{01}' -$

$\Xi_{01}\Xi_{01}')$ is at most $2K$, we then have

$$\|\hat{\Xi}_{01} - \Xi_{01}X_1\|_F \leq \sqrt{2K}\|\hat{\Xi}_{01} - \Xi_{01}X_1\|.$$

Similarly, there exists an orthogonal matrix X_2 such that $\|\hat{\Xi}_{02} - \Xi_{02}X_2\|_F \leq \sqrt{2K}\|\hat{\Xi}_{02} - \Xi_{02}X_2\|$. As a result, for the orthogonal matrix $X = \mathrm{diag}(X_1, X_2)$,

$$\|\hat{\Xi}_0 X - \Xi_0\|_F \leq 2\sqrt{K} \max_{t\in\{1,2\}} \{\|\hat{\Xi}_{0t}\hat{\Xi}_{0t}' - \Xi_{0t}\Xi_{0t}'\|\}. \tag{C.60}$$

Plugging (C.59)-(C.60) into (C.58) gives

$$\min_{\pm} \|\hat{\xi}_1 \pm \xi_1\| = O\Big(\frac{K\|A-\Omega\|}{\|\theta\|^2}\Big) = O\Big(\frac{K\sqrt{\theta_{\max}\|\theta\|_1}}{\|\theta\|^2}\Big),$$
$$\|\hat{\Xi}_0 X - \Xi_0\|_F = O\Big(\frac{K\sqrt{K}\|A-\Omega\|}{\beta_n\|\theta\|^2}\Big) = O\Big(\frac{\sqrt{K^3\theta_{\max}\|\theta\|_1}}{\beta_n\|\theta\|^2}\Big),$$

where the last inequality is from (C.53). This gives the first two items.

We then prove the second two items. We apply the row-wise bounds for eigenvectors in [1]. The following lemma is adapted from [1, Theorem 2.1] and is proved below. A direct use of [1, Theorem 2.1] will lead to sub-optimal dependence on β_n in the resulting bound, so we have to modify that theorem accordingly.

Lemma C.3. *Let $M \in \mathbb{R}^{n,n}$ be a symmetric random matrix. Write $M^* = \mathbb{E}M$ and $K_0 = \mathrm{rank}(M^*)$. Let d_k^* and d_k be the k-th largest nonzero eigenvalue of M^* and M, respectively, and let η_k^* and η_k be the corresponding eigenvector, respectively, $1 \leq k \leq K_0$. Let s and r be two integers such that $1 \leq r \leq K_0$ and $0 \leq s \leq K_0 - r$. Write $D = \mathrm{diag}(d_{s+1}, d_{s+2}, \ldots, d_{s+r})$, $D^* = \mathrm{diag}(d_{s+1}^*, d_{s+2}^*, \ldots, d_{s+r}^*)$,*

$$U = [\eta_{s+1}, \eta_{s+2}, \ldots, \eta_{s+r}], \qquad \textit{and} \qquad U^* = [\eta_{s+1}^*, \eta_{s+2}^*, \ldots, \eta_{s+r}^*].$$

Define $\Delta^* = \min\{d_s^* - d_{s+1}^*, d_{s+r}^* - d_{s+r-1}^*, \min_{1\leq j\leq r}|d_{s+j}^*|\}$ and define $\kappa = (\max_{1\leq j\leq r}|d_{s+j}^*|)/\Delta^*$. Below, the notation $\|\cdot\|_{2\to\infty}$ represents the maximum row-wise ℓ^2-norm of a matrix, and $M_{m,\cdot}^*$ is the m-th row of M^*. Suppose for a number $\gamma > 0$, the following assumptions are satisfied:

- *A1 (Incoherence): $\max_{1\leq m\leq n}\|M_{m,\cdot}^*\| \leq \gamma\Delta^*$.*
- *A2 (Independence): For any $1 \leq m \leq n$, the entries of the m-th row and column of M are independent with the other entries.*
- *A3 (Spectral norm concentration): For a number $\delta_0 \in (0,1)$, $\mathbb{P}(\|M - M^*\| \leq \gamma\Delta^*) \geq 1 - \delta_0$.*

- *A4 (Row concentration): There is a number $\delta_1 \in (0,1)$ and a continuous non-decreasing function $\varphi(\cdot)$ with $\varphi(0)=0$ and $\varphi(x)/x$ being non-increasing in $\mathbb{R}^+$ such that, for any $1 \le m \le n$ and non-stochastic matrix $Y \in \mathbb{R}^{n,r}$,*

$$\mathbb{P}\left(\|(M-M^*)_{m,\cdot}Y\|_2 \le \Delta^*\|Y\|_{2\to\infty}\varphi\Big(\frac{\|Y\|_F}{\sqrt{n}\|Y\|_{2\to\infty}}\Big)\right) \ge 1-\delta_1/n.$$

With probability $1-\delta_0-2\delta_1$, for an orthogonal matrix $O \in \mathbb{R}^{r,r}$,

$$\begin{aligned}&\|UO - MU^*(D^*)^{-1}\|_{2\to\infty}\\ &\le C\big[\kappa(\kappa+\varphi(1))(\gamma+\varphi(\gamma)) + \widetilde{\kappa}\gamma\big]\cdot\|\widetilde{U}^*\|_{2\to\infty},\end{aligned} \tag{C.61}$$

where $\widetilde{U}^* = [\eta_1,\ldots,\eta_{K_0}]$ and $\widetilde{\kappa}$ is defined as follows: Let $I_0 = (\{1,\ldots,s-1\}\cup\{s+r+1,\ldots,K_0\})\cap\{j : |d_j^*| > \max_{1\le i\le r}|d_{s+r}^*|\}$ and $\Delta_0^* = \min\{\min_{j\in I_0}|d_j^* - d_s^*|, \min_{j\in I_0}|d_j^*-d_{s+r}^*|\}$. Define $\widetilde{\kappa} = \max_{j\in I_0}|d_j^*|/\Delta_0^*$ if $I_0 \ne \emptyset$, and $\widetilde{\kappa}=0$ otherwise.

Proof of Lemma C.3: Fix $1 \le m \le n$. Let $M^{(m)}$ be the matrix by setting the m-th row and the m-th column of M to be zero. Let $\eta_1^{(m)}, \eta_2^{(m)}, \ldots, \eta_n^{(m)}$ be the eigenvectors of $M^{(m)}$. Write $U^{(m)} = [\eta_{s+1}^{(m)},\ldots,\eta_{s+r}^{(m)}]$. Let $H = U'U^*$, $H^{(m)} = (U^{(m)})'U^*$ and $V^{(m)} = U^{(m)}H^{(m)} - U^*$. We aim to prove

$$\begin{aligned}\|M_{m\cdot}V^{(m)}\| \le& 6(\kappa+\widetilde{\kappa})\gamma\Delta^*\|\widetilde{U}^*\|_{2\to\infty}\\ &+\Delta^*\varphi(\gamma)\big(4\kappa\|UH\|_{2\to\infty} + 6\|U^*\|_{2\to\infty}\big).\end{aligned} \tag{C.62}$$

Once (C.62) is obtained, the proof is almost identical to the proof of (B.26) in [1], except that we plug in (C.62) instead of (B.32) in [1]. This is straightforward, so we omit it.

What remains is to prove (C.62). Without loss of generality, we only consider the case where $I_0 \ne \emptyset$. In the proof of [1, Lemma 5], it is shown that

$$\|M_{m\cdot}V^{(m)}\| \le \|M_m^*V^{(m)}\| + \|(M-M^*)_{m\cdot}V^{(m)}\|,$$
$$\|(M-M^*)_{m\cdot}V^{(m)}\| \le \Delta^*\varphi(\gamma)\big(4\kappa\|UH\|_{2\to\infty} + 6\|U^*\|_{2\to\infty}\big).$$

Combining them gives

$$\|M_{m\cdot}V^{(m)}\| \le \|M_{m\cdot}^*V^{(m)}\| + \Delta^*\varphi(\gamma)\big(4\kappa\|UH\|_{2\to\infty} + 6\|U^*\|_{2\to\infty}\big). \tag{C.63}$$

We further bound the first term in (C.63). Recall that I_0 is the index set of eigenvalues that are not contained in D^* and have an absolute value larger than $\|D^*\|$. Let $\widetilde{M}^* = \sum_{j\in I_0} d_j^*\eta_j^*(\eta_j^*)'$.

$$\|M_{m\cdot}^*V^{(m)}\| \le \|\widetilde{M}_{m\cdot}^*V^{(m)}\| + \|(M_{m\cdot}^* - \widetilde{M}_{m\cdot}^*)V^{(m)}\|$$

$$
\begin{aligned}
&\le \|\widetilde{M}^*_{m\cdot} V^{(m)}\| + \|M^* - \widetilde{M}^*\|_{2\to\infty} \|V^{(m)}\| \\
&\le \|\widetilde{M}^*_{m\cdot} V^{(m)}\| + 6\gamma \|M^* - \widetilde{M}^*\|_{2\to\infty},
\end{aligned}
$$

where the last line uses $\|V^{(m)}\| \le 6\gamma$, by (B.12) of [1]. Note that $M^* - \widetilde{M}^* = \sum_{j\notin I_0} d_j^* \eta_j^* (\eta_j^*)'$. By definition of I_0, for any $j \notin I_0$, $|d_j^*| \le \max_{1\le i\le r} |d^*_{s+r}| \le \kappa\Delta^*$. It follows that

$$
\|M^* - \widetilde{M}^*\|_{2\to\infty} \le \big(\max_{j\notin I_0} |d_j^*|\big) \|\widetilde{U}^*\|_{2\to\infty} \le \kappa\Delta^* \|\widetilde{U}^*\|_{2\to\infty}.
$$

Combining the above gives

$$
\|M^*_{m\cdot} V^{(m)}\| \le \|\widetilde{M}^*_{m\cdot} V^{(m)}\| + 6\kappa\gamma\Delta^* \|\widetilde{U}^*\|_{2\to\infty}. \tag{C.64}
$$

Write $D_0^* = \mathrm{diag}(d_j^*)_{j\in I_0}$, $U_0^* = [\eta_j^*]_{j\in I_0}$, $U_0 = [\eta_j]_{j\in I_0}$, $U_0^{(m)} = [\eta_j^{(m)}]_{j\in I_0}$, and $H_0^{(m)} = (U_0^{(m)})' U_0^*$. We similarly have $\|U_0^{(m)} H_0^{(m)} - U_0^*\| \le 6\gamma_0$, where γ_0 is defined in the same way as γ but is with respect to the eigen-gap Δ_0^*. It is not hard to see that $\gamma_0 = \gamma\Delta^*/\Delta_0^*$. Hence,

$$
\|U_0^{(m)} H_0^{(m)} - U_0^*\| \le 6\gamma\Delta^*/\Delta_0^*.
$$

By mutual orthogonality of eigenvectors, $(U_0^{(m)})' U^{(m)} = 0$ and $(U_0^*)' U^* = 0$. It follows that

$$
\begin{aligned}
\|\widetilde{M}^*_{m\cdot} V^{(m)}\| &= \|e_m' [U_0^* \Lambda_0^* (U_0^*)'][U^{(m)} H^{(m)} - U^*]\| \\
&= \|e_m' [U_0^* \Lambda_0^* (U_0^*)'] U^{(m)} H^{(m)}\| \\
&\le \|e_m' [U_0^* \Lambda_0^* (U_0^*)'] U^{(m)}\| \\
&= \|e_m' U_0^* \Lambda_0^* (U_0^* - U_0^{(m)} H_0^{(m)})' U^{(m)}\| \\
&\le \|e_m' U_0^* \Lambda_0^* (U_0^* - U_0^{(m)} H_0^{(m)})'\| \\
&\le \|\widetilde{U}^*\|_{2\to\infty} \cdot \|\Lambda_0^*\| \cdot \|U_0^* - U_0^{(m)} H_0^{(m)}\| \\
&\le 6(\|\Lambda_0\|^*/\Delta_0^*) \cdot \gamma\Delta^* \|\widetilde{U}^*\|_{2\to\infty}.
\end{aligned}
$$

We plug it into (C.64) and note that $\widetilde{\kappa} = \|\Lambda_0\|^*/\Delta_0^*$. It gives

$$
\|M^*_{m\cdot} V^{(m)}\| \le 6(\kappa + \widetilde{\kappa})\gamma\Delta^* \|\widetilde{U}^*\|_{2\to\infty}. \tag{C.65}
$$

Combining (C.63) and (C.65) gives (C.62). $\square$

We now come back to the proof of Lemma 2.1. We have divided nonzero eigenvalues of Ω into four groups: (i) λ_1, (ii) positive eigenvalues in $\lambda_2, \ldots, \lambda_K$,

(iii) zero eigenvalues, and (iv) negative eigenvalues in $\lambda_2, \dots, \lambda_K$. We shall apply Lemma C.3 to each of the four groups. To save space, we only consider applying it to group (ii). The proof for other groups is similar and omitted.

Now, $M = A$ and $M^* = \Omega = \mathrm{diag}(\Omega) + (A - \mathbb{E}A)$. We check conditions A1-A4. By Lemma B.1, $\Delta^* \geq C\beta_n K^{-1}\|\theta\|^2$ and $\kappa \leq C$. For an appropriately large constant $\tilde{C} > 0$, we take

$$\gamma = \tilde{C}\beta_n^{-1}\|\theta\|^{-2}K\sqrt{\theta_{\max}\|\theta\|_1}.$$

Consider A1. Since $\Omega(i,j) \leq C\theta(i)\theta(j)$, we have $\max_{1\leq i\leq n}\|\Omega_{i,\cdot}\| \leq C\theta_{\max}\|\theta\|$. From the universal inequality $\|\theta\| \leq \sqrt{\theta_{\max}\|\theta\|_1}$ and the assumption $\theta_{\max} = O(1)$, this term is $O(\sqrt{\theta_{\max}\|\theta\|_1})$, which is bounded by $\gamma\Delta^*$ when $\tilde{C}$ is appropriately large. Hence, A1 is satisfied. A2 is satisfied because the upper triangle of A contains independent variables. By (C.53), A3 is satisfied for $\delta_0 = o(n^{-3})$. We then verify A4. Since $\|\mathrm{diag}(\Omega)\| \leq C$,

$$\|\mathrm{diag}(\Omega)_{i,\cdot}Y\|_2 \leq C\|Y\|_{2\to\infty}, \qquad 1 \leq i \leq n. \tag{C.66}$$

Fix $1 \leq i \leq n$ and $1 \leq k \leq r$. Let $y_k \in \mathbb{R}^n$ be the k-th column of Y. Using the Bernstein's inequality, for any $t \geq 0$,

$$\mathbb{P}\big(|y_k'(A-\mathbb{E}A)_{i,\cdot}| > t\big) \leq 2\exp\left(-\frac{t^2/2}{\sum_{j=1}^n \Omega(i,j)y_k^2(j) + t\|y_k\|_\infty/3}\right). \tag{C.67}$$

Note that $\sum_j \Omega(i,j)y_k^2(j) \leq C\|y_k\|_\infty^2\theta_{\max}\|\theta\|_1$. Moreover, $\theta_{\max}\|\theta\|_1 \gg \log(n)$ by (C.54). We take $t = C\|y_k\|_\infty\sqrt{\theta_{\max}\|\theta\|_1\log(n)}$ for a large enough constant $C > 0$. It follows that with probability $1 - o(n^{-4})$,

$$|y_k'(A - \mathbb{E}A)_{i,\cdot}| \leq \|y_k\|_\infty \cdot C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}.$$

Combining it with the probability union bound and (C.66), with probability $1 - o(n^{-3})$,

$$\begin{aligned}\|(A-\Omega)_{i,\cdot}Y\|_2 &\leq C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}\cdot\|Y\|_{2\to\infty}\\ &\leq \Delta^*\|Y\|_{2\to\infty}\cdot\frac{C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}}{K^{-1}\beta_n\|\theta\|^2}.\end{aligned} \tag{C.68}$$

Moreover, in (C.67), if we use an alternative bound $\sum_j \Omega(i,j)y_k^2(j) \leq \|y_k\|^2\theta_{\max}^2$, we obtain a different bound as follows: With probability $1 - o(n^{-4})$,

$$|y_k'(A - \mathbb{E}A)_{i,\cdot}| \leq C\max\big\{\|y_k\|\theta_{\max}\sqrt{\log(n)},\ \|y_k\|_\infty\log(n)\big\}.$$

Due to the probability union bound and (C.66), with probability $1-o(n^{-3})$,

$$\|(A-\Omega)_{i,\cdot}Y\|_2 \le C\max\big\{\|Y\|_F\theta_{\max}\sqrt{\log(n)},\ \|Y\|_{2\to\infty}\log(n)\big\}$$

$$\le \Delta^*\|Y\|_{2\to\infty}\max\left\{\frac{\theta_{\max}\sqrt{n\log(n)}}{K^{-1}\beta_n\|\theta\|^2}\frac{\|Y\|_F}{\sqrt{n}\|Y\|_{2\to\infty}},\ \frac{\log(n)}{K^{-1}\beta_n\|\theta\|^2}\right\}. \tag{C.69}$$

We introduce the quantities $t_1 = C(K^{-1}\beta_n\|\theta\|^2)^{-1}\sqrt{\theta_{\max}\|\theta\|_1\log(n)}$, $t_2 = C(K^{-1}\beta_n\|\theta\|^2)^{-1}\theta_{\max}\sqrt{n\log(n)}$, and $t_3 = C(K^{-1}\beta_n\|\theta\|^2)^{-1}\log(n)$. Define the function

$$\widetilde{\varphi}(x) = \min\{t_1,\ \max\{t_2x, t_3\}\}.$$

Then, (C.68)-(C.69) together imply that with probability $1-o(n^{-3})$,

$$\|(A-\mathbb{E}A)_{i,\cdot}Y\|_2 \le \Delta^*\|Y\|_{2\to\infty}\widetilde{\varphi}\Big(\frac{\|Y\|_F}{\sqrt{n}\|Y\|_{2\to\infty}}\Big). \tag{C.70}$$

We look at the function $\widetilde{\varphi}(x)$. Note that $(\sqrt{n}\|Y\|_{2\to\infty})^{-1}\|Y\|_F$ takes values in the interval $[n^{-1/2},1]$. By (C.54), $t_1 \gg t_3$. Moreover, since $\|\theta\|_1 \le n\theta_{\max}$, when $x=1$, $t_2x \ge Ct_1$. Last, when $x = n^{-1/2}$, $t_2x \ll t_3$. Combining the above, we conclude that in $[n^{-1/2},\infty)$, the function $\widetilde{\varphi}(x)$ first stays flat at t_3, then linearly increases to t_1 and then stays flat at t_1. Hence, we construct a function $\varphi(x)$, which linearly increases from 0 to t_3 for $x\in[0,n^{-1/2}]$, then linear increases from t_3 to t_1 for $x\in[n^{-1/2}, t_2/t_1]$, and then stays constant as t_1 for $x\in[t_2/t_1,\infty)$. It is seen that $\varphi(0)=0$, $\varphi(x)/x$ is non-increasing, and $\widetilde{\varphi}(x)\le\varphi(x)\le t_1$ in the interval $[n^{-1/2},1]$. By (C.70) and that $\widetilde{\varphi}(x)\le\varphi(x)$, A4 is satisfied with $\delta_1 = o(n^{-3})$. Furthermore, since $\varphi(x)\le t_1$,

$$\varphi(\gamma) \le \frac{C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}}{K^{-1}\beta_n\|\theta\|^2}.$$

So far, we have shown that A1-A4 hold.

We now apply Lemma C.3. As mentioned, we only study the eigenvectors in group (ii), which correspond to positive eigenvalues among $\lambda_2,\dots,\lambda_K$. Let Λ_1 be the diagonal matrix consisting of these eigenvalues and let Ξ_{01} be the matrix formed by associated eigenvectors. Define their empirical counterparts, $\hat{\Lambda}_1$ and $\hat{\Xi}_{01}$, in the same way. In Lemma C.3, we take $U=\hat{\Xi}_{01}$, $U^*=\Xi_{01}$, and $\widetilde{U}^*=\Xi$. Since $\lambda_2,\dots,\lambda_K$ are at the same order, $\kappa\le C$. Also, $\widetilde{\kappa}\le\lambda_1/(\lambda_1-|\lambda_2|)\le C$ by our assumption. It follows from (C.61) that there exists an orthogonal matrix O such that

$$\|\hat{\Xi}_{01}O - A\Xi_{01}\Lambda_1^{-1}\|_{2\to\infty} \le \frac{C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}}{K^{-1}\beta_n\|\theta\|^2}\|\Xi\|_{2\to\infty}.$$

By Lemma B.2, $\|\Xi\|_{2\to\infty} = O(\sqrt{K}\|\theta\|^{-1}\theta_{\max})$. Plugging it into the above inequality, we find that

$$\|\hat{\Xi}_{01}O - A\Xi_{01}\Lambda_1^{-1}\|_{2\to\infty} \le \frac{C\theta_{\max}^{3/2}K^{3/2}\sqrt{\|\theta\|_1\log(n)}}{\beta_n\|\theta\|^3}. \tag{C.71}$$

By definition of eigen-decomposition, $\Omega\Xi_{01} = \Xi_{01}\Lambda_1$. It follows that

$$A\Xi_{01}\Lambda_1^{-1} = \Omega\Xi_{01}\Lambda_1^{-1} + (A-\Omega)\Xi_{01}\Lambda_1^{-1} = \Xi_{01} + (A-\Omega)\Xi_{01}\Lambda_1^{-1}.$$

Plugging it into (C.71) yields

$$\|\hat{\Xi}_{01}O - \Xi_{01}\|_{2\to\infty} \le \frac{C\theta_{\max}^{3/2}K^{3/2}\sqrt{\|\theta\|_1\log(n)}}{\beta_n\|\theta\|^3} + \|(A-\Omega)\Xi_{01}\Lambda_1^{-1}\|_{2\to\infty}. \tag{C.72}$$

To bound the second term on the right hand side, we apply the first line of (C.68) by letting $Y = \Xi_{01}$. It turns out that with probability $1 - o(n^{-3})$,

$$\begin{aligned}
\|(A-\Omega)\Xi_{01}\Lambda_1^{-1}\|_{2\to\infty} &\le (\max_{1\le i\le n}\|(A-\Omega)_{i,\cdot}\Xi_{01}\|_2)\cdot\|\Lambda_1^{-1}\| \\
&\le C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}\cdot\|\Xi_{01}\|_{2\to\infty}\cdot\|\Lambda_1^{-1}\| \\
&\le C\sqrt{\theta_{\max}\|\theta\|_1\log(n)}\cdot\sqrt{K}\|\theta\|^{-1}\theta_{\max}\cdot K\beta_n^{-1}\|\theta\|^{-2}, \qquad \text{(C.73)}
\end{aligned}$$

where in the last inequality, the bound of $\|\Lambda_1^{-1}\|$ is from Lemma B.1 and the bound of $\|\Xi_{01}\|_{2\to\infty}$ is from Lemma B.2. Combining (C.72)-(C.73) gives

$$\|\hat{\Xi}_{01}O - \Xi_{01}\|_{2\to\infty} \le \frac{C\theta_{\max}^{3/2}K^{3/2}\sqrt{\|\theta\|_1\log(n)}}{\beta_n\|\theta\|^3}.$$

Note that the left hand side only involves eigenvectors in group (ii). We can prove similar results for the other three groups of eigenvectors. For group (i), $\Delta^* \ge CK^{-1}\|\theta\|^{-1}$ and $\|\widetilde{U}^*\|_{2\to\infty} \le C\|\theta\|^{-1}\theta_{\max}$, and the resulting bound is

$$\|\omega\hat{\xi}_1 - \xi_1\|_\infty \le \frac{C\theta_{\max}^{3/2}K\sqrt{\|\theta\|_1\log(n)}}{\|\theta\|^3}.$$

Furthermore, if $\beta_n = o(1)$, by Lemma B.1, $\lambda_1 - |\lambda_2| \ge C^{-1}\lambda_1 \ge C^{-1}K\|\theta\|^2$. Compared with the case of $\beta_n \ge c$, the Δ^* of group (i) is larger by a factor of K, so all the bounds concerning $\hat{\xi}_1$ are reduced by a factor of K.

C.3. Proof of Theorem 2.1. The second claim is straightforward. We only show the first claim. By Lemma B.2, we can choose the sign of ξ_1 such that it is a strictly positive vector. By definition of err_n, we can re-write

$$err_n = \frac{\|\theta\|}{\theta_{\min}} \cdot \frac{\theta_{\max}^{3/2}\sqrt{\|\theta\|_1 \log(n)}}{\|\theta\|^3}.$$

Then, the last two bullet points of Lemma 2.1 can be re-expressed as

(C.74)
$$\|\omega\hat{\xi}-\xi\|_\infty = O\Big(\frac{\theta_{\min}}{\|\theta\|}Kerr_n\Big), \quad \max_{1\leq i\leq n}\|X'\hat{\Xi}_{i,0}-\Xi_{i,0}\| = O\Big(\frac{\theta_{\min}}{\|\theta\|}K^{3/2}\beta_n^{-1}err_n\Big).$$

We now show the claim. Let (ω, X) be the same as in Lemma 2.1, and define $H = \omega X' \in \mathbb{R}^{K-1,K-1}$. Fix i. By definition of $(r_i, \hat{r}_i)$ and H,

$$r_i = \frac{1}{\xi_1(i)}\Xi_{i,0}, \qquad H\hat{r}_i = \omega X'\hat{r}_i = \frac{1}{\omega\hat{\xi}_1(i)}X'\hat{\Xi}_{i,0}.$$

It follows that

$$\begin{aligned} H\hat{r}_i - r_i &= \frac{1}{\omega\hat{\xi}_1(i)}(X'\hat{\Xi}_{i,0} - \Xi_{i,0}) + \Big[\frac{1}{\omega\hat{\xi}_1(i)} - \frac{1}{\xi_1(i)}\Big]\Xi_{i,0} \\ &= \frac{1}{\omega\hat{\xi}_1(i)}(X'\hat{\Xi}_{i,0} - \Xi_{i,0}) - \frac{\omega\hat{\xi}_1(i) - \xi_1(i)}{\omega\hat{\xi}_1(i)}r_i. \end{aligned}$$

First, by Lemma B.2, $\xi_1(i) \geq C\theta_{\min}/\|\theta\|$; also, by (C.74), $|\omega\hat{\xi}_1(i) - \xi_1(i)| \ll \theta_{\min}/\|\theta\|$. We thus have $\omega\hat{\xi}_1(i) \geq \xi_1(i)/2 \geq C\theta_{\min}/\|\theta\|$. Second, using the first bullet point of Lemma B.3, we have $\|r_i\| \leq \max_k \|v_k\| \leq C\sqrt{K}$. Plugging these results into the above equation gives

$$\|H\hat{r}_i - r_i\| \leq \frac{C\|\theta\|}{\theta_{\min}}\big(\|X'\hat{\Xi}_{i,0} - \Xi_{i,0}\| + \sqrt{K}|\omega\hat{\xi}_1(i) - \xi_1(i)|\big). \tag{C.75}$$

The claim follows by plugging (C.74) into (C.75).

C.4. Two useful lemmas and their proofs.

Lemma C.4. *Let H be the orthogonal matrix in Theorem 2.1. With probability $1-o(n^{-3})$, $\|H\mathrm{diag}(\hat{\lambda}_2,\ldots,\hat{\lambda}_K)-\mathrm{diag}(\hat{\lambda}_2,\ldots,\hat{\lambda}_K)H\| \leq C\sqrt{\theta_{\max}\|\theta\|_1}$.*

Lemma C.5. *Let err_n^* be as defined in Section 3.5. Under conditions of Theorem 2.1, with probability $1-o(n^{-3})$, $n^{-1}\sum_{i=1}^n \|H\hat{r}_i-r_i\|^2 \leq CK^3\beta_n^{-2}(err_n^*)^2$.*

Proof of Lemma C.4: Write for short $\hat{\Lambda}_0 = \mathrm{diag}(\hat{\lambda}_2, \ldots, \hat{\lambda}_K)$. We shall apply [1, Lemma 2]: in our setting, their notations H and $\mathrm{sgn}(H)$ correspond to the matrix $\hat{\Xi}_0'\Xi_0$ and the orthogonal matrix X in Lemma 2.1 (to clarify, our notation H means $\omega X'$). By their Lemma 2,

$$\|\hat{\Xi}_0'\Xi_0 - X\|^{1/2} \leq C\|A - \Omega\|/\Delta^*, \quad \|(\hat{\Xi}_0'\Xi_0)\hat{\Lambda}_0 - \hat{\Lambda}_0(\hat{\Xi}_0'\Xi_0)\| \leq 2\|A - \Omega\|,$$

where Δ^* is the eigen-gap quantity defined in the proof of Lemma 2.1 and it satisfies $\Delta^* \geq C\beta_n K^{-1}\|\theta\|^2$. Additionally, by Lemma B.1 and Lemma C.1, $\|\hat{\Lambda}_0\| \lesssim \|\Lambda_0\| \leq C\beta_n K^{-1}\|\theta\|^2 \leq C\Delta^*$, with probability $1 - o(n^{-3})$. Hence, with probability $1 - o(n^{-3})$,

$$\begin{aligned}
\|H\hat{\Lambda}_0 - \hat{\Lambda}_0 H\| &= \|X\hat{\Lambda}_0 - \hat{\Lambda}_0 X\| \\
&\leq 2\|\hat{\Xi}_0'\Xi_0 - X\| \cdot \|\hat{\Lambda}_0\| + \|(\hat{\Xi}_0'\Xi_0)\hat{\Lambda}_0 - \hat{\Lambda}_0(\hat{\Xi}_0'\Xi_0)\| \\
&\leq C\|A - \Omega\|^2(\Delta^*)^{-2} \cdot C\Delta^* + 2\|A - \Omega\| \\
&\leq 2\|A - \Omega\|\big(1 + C(\Delta^*)^{-1}\|A - \Omega\|\big) \\
&\lesssim 2\|A - \Omega\| \\
&\leq C\sqrt{\theta_{\max}\|\theta\|_1},
\end{aligned}$$

where the last line is from (C.53). □

Proof of Lemma C.5: By definition of err_n^*, we can re-write it as

$$err_n^* = \frac{\|\theta\|}{\theta_{\min}\sqrt{n}} \cdot \frac{\sqrt{\theta_{\max}\|\theta\|_1}}{\|\theta\|^2}.$$

Then, the first two bullet points of Lemma 2.1 can be re-expressed as

$$\|\omega\hat{\xi} - \xi\| = O\bigg(\frac{\theta_{\min}\sqrt{n}}{\|\theta\|} K err_n^*\bigg), \quad \|\hat{\Xi}_0 X - \Xi_0\|_F = O\bigg(\frac{\theta_{\min}\sqrt{n}}{\|\theta\|} K^{3/2}\beta_n^{-1} err_n^*\bigg).$$

Combining it with (C.75) gives

$$n^{-1}\sum_{i=1}^n \|H\hat{r}_i - r_i\|^2 \leq \frac{C\|\theta\|^2}{n\theta_{\min}^2}\big(\|\hat{\Xi}_0 X - \hat{\Xi}_0\|_F^2 + K\|\omega\hat{\xi}_1 - \xi_1\|^2\big) \leq CK^3\beta_n^{-2}(err_n^*)^2.$$

This proves the claim. □

APPENDIX D: VERTEX HUNTING

We prove results in Section 3 of the main paper, including Lemmas 3.1-3.6 (about different vertex hunting algorithms) and Theorem 3.1 (about the faster rate of Mixed-SCORE-SVS).

D.1. Proof of Lemma 3.1. The following lemma is from [25, Theorem 1], which is basically a summary of [24, Theorem 3].

LEMMA D.1. *Fix $m \geq r$ and $n \geq r$. Consider a matrix $Y = SM + Z$, where $S \in \mathbb{R}^{m\times r}$ has a full column rank, $M \in \mathbb{R}^{r\times n}$ is a nonnegative matrix such that the sum of each column is at most 1, and $Z = [Z_1, \dots, Z_n] \in \mathbb{R}^{m\times n}$. Suppose M has a submatrix equal to I_r. Write $\epsilon = \max_{1\leq i\leq n} \|Z_i\|$. Suppose $\epsilon = O(\frac{\sigma_{\min}(S)}{\sqrt{r}\kappa^2(S)})$, where $\sigma_{\min}(S)$ and $\kappa(S)$ are the minimum singular value and condition number of S, respectively. If we apply the SP algorithm to columns of Y, then it outputs an index set $\mathcal{K} \subset \{1, 2, \dots, n\}$ such that $|\mathcal{K}| = r$ and $\max_{1\leq k\leq r} \min_{j\in\mathcal{K}} \|S_k - Y_j\| = O(\epsilon\kappa^2(S))$, where S_k is the k-th column of S.*

We note that the estimated vertices by SP are $\{Y_j\}_{j=1}^r$. Hence, the above lemma says the maximum ℓ^2-error on estimating vertices is

$$O\big(\epsilon\kappa^2(S)\big) = O\Big(\kappa^2(S) \max_{1\leq i\leq n} \|Z_i\|\Big).$$

In our setting, we apply the SP algorithm to $Y_i = (1, \hat{r}_i')'$, $1 \leq i \leq n$. We shall re-write the data in the same form as in Lemma D.1. Recall that H is the orthogonal matrix in Theorem 2.1 and $v_1, \dots, v_K$ are vertices of the Ideal Simplex. By definition,

$$\begin{pmatrix} 1 & \cdots & 1 \\ H^{-1}v_1 & \cdots & H^{-1}v_K \end{pmatrix} w_i = \begin{pmatrix} 1 \\ H^{-1}r_i \end{pmatrix}.$$

Let $\tilde{v}_k = (1, (H^{-1}v_k)')'$, $\tilde{r}_i = (1, (H^{-1}r_i)')'$, $z_i = (0, (\hat{r}_i - H^{-1}r_i)')'$, $1 \leq k \leq K$, $1 \leq i \leq n$. It is seen that

$$(1, \hat{r}_i')' \equiv Y_i = [\tilde{v}_1, \dots, \tilde{v}_K]w_i + z_i.$$

Write $Y = [Y_1, \dots, Y_n] \in \mathbb{R}^{K\times n}$, $\tilde{V} = [\tilde{v}_1, \dots, \tilde{v}_K] \in \mathbb{R}^{K\times K}$, $W = [w_1, \dots, w_n] \in \mathbb{R}^{K\times n}$, and $Z = [z_1, \dots, z_n] \in \mathbb{R}^{K\times n}$. The above can be re-written as

$$Y = \tilde{V}W + Z. \tag{D.76}$$

This reduces to the form in Lemma D.1 with $m = K$. To apply Lemma D.1, we note that $\tilde{V}$ can be re-written as

$$\tilde{V} = \mathrm{diag}(1, H^{-1}) \cdot Q, \qquad \text{where} \quad Q = \begin{pmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_K \end{pmatrix}.$$

Since $\mathrm{diag}(1, H^{-1})$ is an orthogonal matrix, the singular values of $\tilde{V}$ are the same as the singular values of Q. Moreover, by (B.52), all the singular values of Q are at the order of $\sqrt{K}$. It follows that

$$\sigma_{\min}(\tilde{V}) \asymp \sqrt{K}, \qquad \kappa(\tilde{V}) \asymp 1. \tag{D.77}$$

In particular, $\tilde{V}$ has a full rank, and $\frac{\sigma_{\min}(\tilde{V})}{\sqrt{K}\kappa^2(\tilde{V})} \asymp 1$. By by Lemma D.1, when $\max_{1\leq i\leq n}\|Z_i\| = O(1)$, the maximum ℓ^2-error on estimating vertices is

$$O\Big(\max_{1\leq i\leq n}\|Z_i\|\Big).$$

Additionally, $\max_{1\leq i\leq n}\|Z_i\| = \max_{1\leq i\leq n}\|\hat{r}_i - H^{-1}r_i\| = \max_{1\leq i\leq n}\|H\hat{r}_i - r_i\|$. The claim follows immediately.

D.2. Proof of Lemma 3.2. The lemma below is a special case of [33, Theorem 1] with the uniqueness parameter $\alpha = 1$.

Lemma D.2. *Fix $d \geq r-1$ and $n \geq r$. Consider a matrix $Y = SM + Z$, where $S \in \mathbb{R}^{d\times r}$ has affinely-independent columns, $M \in \mathbb{R}^{r\times n}$ is a nonnegative matrix where the sum of each column equals to 1, and $Z = [Z_1, \ldots, Z_n] \in \mathbb{R}^{d\times n}$. Suppose M has a submatrix equal to I_r, and suppose the convex hull of columns of S has internal radius of μ. Write $\epsilon = \max_{1\leq i\leq n}\|Z_i\|$. Suppose $\epsilon = O(\mu/r^{3/2})$. If we apply the AA method to columns of Y with tuning parameter $\delta = \epsilon$, then the solution $\hat{S}$ satisfies $\sqrt{\sum_{k=1}^K \min_{1\leq \ell\leq K}\|\hat{S}_\ell - S_k\|^2} \leq C\mu^{-1}\sigma_{\max}(S)\cdot\max\{1, r^{-1/2}\kappa(S)\}\cdot K^{5/2}\epsilon$, where S_k and $\hat{S}_k$ are the respective k-th column of S and $\hat{S}$.*

Recall that H is the orthogonal matrix in Theorem 2.1 and $v_1, \ldots, v_K$ are vertices of the Ideal Simplex. Write $Y = [\hat{r}_1, \ldots, \hat{r}_n]$, $S = H^{-1}[v_1, \cdots, v_K]$, $M = [w_1, \ldots, w_n]$, and $Z_i = \hat{r}_i - H^{-1}r_i$, $1 \leq i \leq n$. Then, our problem has the same form as that in Lemma D.2, with $r = K$ and $d = K-1$.

It suffices to evaluate $\sigma_{\max}(S)$, $\kappa(S)$, and the internal radius μ. Write

$$V = [v_1, v_2, \ldots, v_K], \qquad Q = \begin{pmatrix} 1 & \cdots & 1 \\ v_1 & \cdots & v_K \end{pmatrix}.$$

Since $S = H^{-1}V$ for an orthogonal matrix H, the singular values of S are the same as the singular values of V. Observing that V is a sub-matrix of Q, we conclude $\sigma_{\max}(V) \leq \|V\| \leq \|Q\|$; additionally, by (B.52), $\|Q\| = O(\sqrt{K})$. It follows that

$$\sigma_{\max}(S) = O(\sqrt{K}). \tag{D.78}$$

Furthermore, the matrix QQ' has a $(K-1)\times(K-1)$ diagonal block equal to VV'. As a result, $\lambda_{\min}(VV') \geq \lambda_{\min}(QQ')$. By (B.52), $\|Q^{-1}\| = O(1/\sqrt{K})$, which implies that $\lambda_{\min}(QQ') \geq C^{-1}K$. It follows that $\sigma_{\min}(S) = \sigma_{\min}(V) \geq \sigma_{\min}(Q) \geq C\sqrt{K}$. Combining it with (D.78) gives

$$\kappa(S) = O(1). \tag{D.79}$$

We then compute the internal radius μ. By the definition in [33], if there is $x_0 \in \mathbb{R}^{K-1}$ and an orthogonal matrix $U \in \mathbb{R}^{(K-1)\times(K-1)}$ such that $\{x_0+Ux : x \in B_{K-1}(\mu)\}$ is contained in the convex hull of S, where $B_{K-1}(\mu)$ is the ball in $\mathbb{R}^{K-1}$ centered at the origin with a radius μ, then the internal radius is at least μ. Here, since H is an orthogonal matrix, it suffices to compute the internal radius of the matrix V. Note that the convex hull of the columns of V are indeed the Ideal Simplex S^{ideal}. Let $\mathcal{S}_0$ be the standard simplex in $\mathbb{R}^K$, whose vertices are the standard basis vectors $e_1, \ldots, e_K$. It is seen that

$$\mathcal{S}^{ideal} = \{Vy : y \in \mathcal{S}_0\}.$$

Suppose $\mathcal{S}_0$ has an internal radius of at least μ_0. Then, for some $y_0 \in \mathcal{S}_0$ and $U_0 \in \mathbb{R}^{K\times(K-1)}$ such that $U_0'U_0 = I_{K-1}$, the set $\{y_0 + U_0 y : y \in B_{K-1}(\mu_0)\}$ is contained in $\mathcal{S}_0$. It follows that $\{Vy_0 + VU_0 y : y \in B_{K-1}(\mu_0)\} \subset S^{ideal}$. Let $V = U\Lambda\Xi$ be the singular value decomposition of $V \in \mathbb{R}^{(K-1)\times K}$, where U is an orthogonal matrix, Λ is a diagonal matrix whose diagonal entries are all at the order of $\sqrt{K}$ (by (D.78)-(D.79)), and $\Xi \in \mathbb{R}^{(K-1)\times K}$ is such that $\Xi\Xi' = I_{K-1}$. Note that $\{\Xi U_0 y : y \in B_{K-1}(\mu_0)\} = B_{K-1}(\mu_0)$. It follows that

$$\{Vy_0 + U\Lambda x : x \in B_{K-1}(\mu_0)\} \subset \mathcal{S}^{ideal}.$$

Let $c_{\min}$ be the minimum diagonal of Λ. Then, $\|\Lambda x\| \geq c_{\min}\|x\|$. It implies $B_{K-1}(c_{\min}\mu_0) \subset \{\Lambda x : x \in B_{K-1}(\mu_0)\}$. Therefore,

$$\big\{Vy_0 + Ux : x \in B_{K-1}(c_{\min}\mu_0)\big\} \subset \mathcal{S}^{ideal}.$$

This means the internal radius of S^{ideal} is at least $c_{\min}\mu_0$. We have seen that $c_{\min} \asymp \sqrt{K}$. Furthermore, the internal radius of a standard simplex is known to be $\mathcal{S}_0$ is $1/\sqrt{K(K-1)}$ (which is basically the distance from $(\frac{1}{K}, \ldots, \frac{1}{K})'$ to $(0, \frac{1}{K-1}, \ldots, \frac{1}{K-1})'$). We thus conclude that

$$\mu \asymp \sqrt{K} \cdot \frac{1}{\sqrt{K(K-1)}} \asymp \frac{1}{\sqrt{K}}. \tag{D.80}$$

Plugging (D.78), (D.79) and (D.80) into Lemma D.2 gives

$$\sqrt{\sum_{k=1}^{K} \|H\hat{v}_k - v_k\|^2} \leq CK^3\sqrt{K}\Big(\max_{1\leq i\leq n} \|H\hat{r}_i - r_i\|\Big).$$

The claim follows by noting that the left hand side above is lower bounded by $\max_{1\leq k\leq K} \|H\hat{v}_k - v_k\|$.

D.3. Proof of Lemma 3.3. Without loss of generality, we only consider the case that H equals to the identity matrix. When H is not the identity matrix, noticing that $\max_{1\leq k\leq K}\|H\hat{v}_k - v_k\| = \max_{1\leq k\leq K}\|\hat{v}_k - H'v_k\|$, we only need to plug $H'v_1, \ldots, H'v_K$ into the proof below.

Write $\hat{h} = \max_{1\leq i\leq n}\|\hat{r}_i - r_i\|$. We aim to show

$$\min_{1\leq \ell\leq K}\|v_k - \hat{v}_\ell\| \leq C_0\hat{h}, \qquad \text{for all } 1\leq k\leq K. \tag{D.81}$$

It means for each true vertex v_k, there is at least one of $\{\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_K\}$ that is within a distance of $C_0\hat{h}$ to v_k. At the same time, since $\hat{h} = o(\sqrt{K})$ and the distance between any two vertices is at least a constant times $\sqrt{K}$ (see Lemma B.3), each $\hat{v}_\ell$ cannot be simultaneously within a distance $C_0\hat{h}$ to two vertices. The above imply that there is a one-to-one correspondence between true and estimated vertices such that for each true vertex the corresponding estimated vertex is within a distance $C_0\hat{h}$ to it. The claim then follows.

We now show (D.81). Fix $1\leq k\leq K$. Recall that w_i is the unique weight vector such that $r_i = \sum_{s=1}^K w_i(s)v_s$, $1\leq i\leq n$. For a constant $C_1 > 0$ to be decided, let

$$\mathcal{V}_{0k} = \big\{1\leq i\leq n : w_i(k) \geq 1 - C_1K^{-1/2}\hat{h}\big\}.$$

Let $\hat{i}_s$ be such that $\hat{v}_s = \hat{r}_{\hat{i}_s}$, $1\leq s\leq K$. We shall first prove that

$$\{\hat{i}_1, \hat{i}_2, \ldots, \hat{i}_K\} \cap \mathcal{V}_{0k} \neq \emptyset. \tag{D.82}$$

This means at least one of the estimated vertices has to come from the point set $\{\hat{r}_i : i\in\mathcal{V}_{0k}\}$. We shall next prove that

$$\max_{i\in\mathcal{V}_{0k}}\|\hat{r}_i - v_k\| \leq C_0\hat{h}. \tag{D.83}$$

Then, the estimated vertex which comes from $\{\hat{r}_i : i\in\mathcal{V}_{0k}\}$ is guaranteed to be within a distance $C_0\hat{h}$ to the true v_k, i.e., (D.81) holds.

It remains to show (D.82)-(D.83). First, consider (D.82). In the proof of Lemma B.3, we introduce a one-to-one linear mapping T_2 from the standard simplex $\mathcal{S}_0$ to the Ideal Simplex $\mathcal{S}^{ideal}$ such that $T_2(w_i) = r_i$ for all $1\leq i\leq n$. We have shown that both T_2 and T_2^{-1} are Lipschitz with the Lipschitz constants at the order of $\sqrt{K}$ and $1/\sqrt{K}$, respectively. As a result, there is a constant $C_2 > 1$ such that, for any $w, \tilde{w}\in\mathcal{S}_0$,

$$C_2^{-1}\sqrt{K}\|w - \tilde{w}\| \leq \|T_2(w) - T_2(\tilde{w})\| \leq C_2\sqrt{K}\|w - \tilde{w}\|. \tag{D.84}$$

Below, we first use (D.84) to show the distance from v_k to the convex hull of $\{r_i : i\notin\mathcal{V}_{0k}\}$ is sufficiently large, and then prove (D.82) by contradiction.

We take $C_1 = 5C_2$. Take an arbitrary point x^* from the convex hull $\mathcal{H}\{r_i : i \notin \mathcal{V}_{0k}\}$. Since T_2 is a linear mapping, $y^* = T_2^{-1}(x^*)$ is a convex combination of $\{w_i : i \notin \mathcal{V}_{0k}\}$. By definition, for each $i \notin \mathcal{V}_{0k}$, $0 \leq w_i(k) \leq 1 - C_1 K^{-1/2}\hat{h}$. As a result, $y^*(k)$, as a convex combination of $\{w_i(k) : i \notin \mathcal{V}_{0k}\}$, also satisfies that $0 \leq y^*(k) \leq 1 - C_1 K^{-1/2}\hat{h}$. This implies

$$\|T_2^{-1}(x^*) - e_k\| = \|y^* - e_k\| \geq C_1 K^{-1/2}\hat{h}, \qquad \text{for any } x^* \in \mathcal{H}\{r_i : i \notin \mathcal{V}_{0k}\}.$$

Combining it with (D.84), we have

$$\|x^* - v_k\| = \|T_2(y^*) - T_2(e_k)\| \geq C_2^{-1}\sqrt{K} \cdot C_1 K^{-1/2}\hat{h} \geq 5\hat{h}.$$

Since x^* is taken arbitrarily from the convex hull $\mathcal{H}\{r_i : i \notin \mathcal{V}_{0k}\}$, we have

$$d\big(v_k, \mathcal{H}\{r_i : i \notin \mathcal{V}_{0k}\}\big) \geq 5\hat{h}. \tag{D.85}$$

Come back to the proof of (D.82). When this claim is not true, the estimated simplex $\widehat{\mathcal{S}}$ is contained in the convex hull of $\{\hat{r}_i : i \notin \mathcal{V}_{0k}\}$. It follows that

$$\begin{aligned} d(v_k, \widehat{\mathcal{S}}) &\geq d\big(v_k, \mathcal{H}\{\hat{r}_i : i \notin \mathcal{V}_{0k}\}\big) \\ &\geq d\big(v_k, \mathcal{H}\{r_i : i \notin \mathcal{V}_{0k}\}\big) - \hat{h} \\ &\geq 4\hat{h}. \end{aligned}$$

Let j_k be a pure node of community k. Then, $\|\hat{r}_{j_k} - v_k\| = \|\hat{r}_{j_k} - r_{j_k}\| \leq \hat{h}$. It follows that

$$\max_{1 \leq i \leq n} d(\hat{r}_i, \widehat{\mathcal{S}}) \geq d(\hat{r}_{j_k}, \widehat{\mathcal{S}}) \geq d(v_k, \widehat{\mathcal{S}}) - \hat{h} \geq 3\hat{h}. \tag{D.86}$$

At the same time, consider the simplex $\widehat{\mathcal{S}}^*$ formed by $\hat{r}_{j_1}, \hat{r}_{j_2}, \ldots, \hat{r}_{j_K}$, where j_s is a pure node of community s, for $1 \leq s \leq K$. Note that $r_{i_1}, r_{i_2}, \ldots, r_{i_K}$ form the Ideal Simplex $\mathcal{S}^*$ and $\max_{1 \leq i \leq n} d(r_i, \mathcal{S}^*) = 0$. It follows that

$$\max_{1 \leq i \leq n} d(\hat{r}_i, \widehat{\mathcal{S}}^*) \leq \max_{1 \leq i \leq n} d(r_i, \mathcal{S}^*) + 2\hat{h} \leq 2\hat{h}. \tag{D.87}$$

Note that $\widehat{\mathcal{S}}$ is the solution of the combinatory search step. It has to satisfy

$$\max_{1 \leq i \leq n} d(\hat{r}_i, \widehat{\mathcal{S}}) \leq \max_{1 \leq i \leq n} d(\hat{r}_i, \widehat{\mathcal{S}}^*).$$

This yields a contradiction to (D.86)-(D.87). Hence, (D.82) must be true.

Next, consider (D.83). It is easy to see that

$$\max_{i \in \mathcal{V}_{0k}} \|\hat{r}_i - v_k\| \leq \max_{i \in \mathcal{V}_{0k}} \|r_i - v_k\| + \hat{h}$$

$$
\begin{aligned}
&= \max_{i \in \mathcal{V}_{0k}} \|T_2(w_i) - T_2(e_k)\| + \hat{h} \\
&\leq C_2\sqrt{K} \max_{i \in \mathcal{V}_{0k}} \|w_i - e_k\| + \hat{h},
\end{aligned}
$$

where we have used (D.84) in the last line. For any $i \in \mathcal{V}_{0k}$, $\|w_i - e_k\|^2 = [1 - w_i(k)]^2 + \sum_{\ell \neq k} w_i^2(\ell) \leq [1 - w_i(k)]^2 + [\sum_{\ell \neq k} w_i(\ell)]^2 \leq 2(C_1 K^{-1/2}\hat{h})^2 = 50C_2^2 K^{-1}\hat{h}^2$. It follows that

$$
\max_{i \in \mathcal{V}_{0k}} \|\hat{r}_i - v_k\| \leq (5\sqrt{2}C_2^2 + 1)\hat{h}.
$$

Hence, (D.83) is true by choosing $C_0 = 5\sqrt{2}C_2^2 + 1$.

D.4. Proof of Lemma 3.4. We aim to prove the following lemma:

LEMMA D.3. *Suppose the conditions of Lemma 3.4 hold. We apply the SVS algorithm to* $\{\hat{r}_i\}_{i=1}^n$ *with* L *being a properly large constant. Write* $\hat{h} = \max_{1 \leq i \leq n} \|H\hat{r}_i - r_i\|$. *The following statements are true.*

- *In the local clustering sub-step, all the local centers output by k-means are within a distance of* $C\hat{h}$ *to the Ideal Simplex. Moreover, for each true vertex* v_k*, there is at least one local center that is within a distance of* $C\hat{h}$ *to it,* $1 \leq k \leq K$.
- *The combinatorial search sub-step selects exactly one local center among those within a distance of* $C\hat{h}$ *to a true* v_k, $1 \leq k \leq K$. *As a result, up to a permutation of estimated vertices,* $\max_{1 \leq k \leq K} \|H\hat{v}_k - v_k\| \leq C\hat{h}$.

The claim of Lemma 3.4 is simply the second bullet point of Lemma D.3.

Proof of Lemma D.3: As explained in the proof of Lemma 3.3, without loss of generality, we consider the case that H equals to the identity matrix. We first argue that, once the first bullet point is proved, the second bullet point follows directly from Lemma 3.3. Let $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L$ be the local centers by k-means. When the first bullet point is true, we have

- $d(\hat{m}_j, \mathcal{S}^{ideal}) \leq C\hat{h}$, $1 \leq j \leq L$.
- For each $1 \leq k \leq K$, there exists j_k such that $\|\hat{m}_{j_k} - v_k\| \leq C\hat{h}$.

By Lemma B.3, the distance between two different v_k and v_ℓ is lower bounded by a constant times $\sqrt{K}$, while $\hat{h} = o(\sqrt{K})$. As a result, any $\hat{m}_j$ cannot be simultaneously within a distance of $C\hat{h}$ to two vertices, which implies that $j_1, j_2, \ldots, j_K$ are distinct. Define

$$
m_j = \begin{cases} \text{argmin}_{x \in \mathcal{S}^{ideal}} \|x - \hat{m}_j\|, & j \notin \{j_1, j_2, \ldots, j_K\}, \\ v_k, & j = j_k, 1 \leq k \leq K. \end{cases}
$$

We then have

- The points $m_1, m_2, \dots, m_L$ are in the Ideal Simplex $\mathcal{S}^{ideal}$.
- $\|\hat{m}_j - m_j\| \le C\hat{h}$, $1 \le j \le L$.
- For each $1 \le k \le K$, there is at least one m_j located at the vertex v_k.

If we view $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_L$ as the data points $\hat{r}_1, \hat{r}_2, \dots, \hat{r}_n$ in Lemma 3.3 and view $m_{j_1}, \dots, m_{j_K}$ as the "pure nodes" , then the assumptions of Lemma 3.3 are satisfied. According to this lemma, by running the combinatorial search on $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_L$, we have

$$\max_{1\le k\le K} \|\hat{v}_k - v_k\| \le C \max_{1\le j\le L} \|\hat{m}_j - m_j\| \le C\hat{h}.$$

It remains to prove the first bullet point. For any $L \ge 1$, define $RSS(L)$ to be the objective achieved by applying k-means to mixed r_i's assuming $\le L$ clusters:

$$RSS(L) = \min_{L \text{ cluster centers}} \sum_{\text{mixed nodes } i} \|r_i - (\text{closest-cluster-center})\|^2.$$

In preparation, we study $RSS(L)$ as a function of L.

We provide an upper bound of $RSS(L)$ by constructing a feasible solution to the k-means problem. In the proof of Lemma B.3, we see that there is a one-to-one mapping $T = T_2 \circ T_1$ from the standard simplex $\mathcal{S}_0$ to the Ideal Simplex $\mathcal{S}^{ideal}$ such that $r_i = T(\pi_i)$ and that (note: we have used that K is a constant)

$$C^{-1}\|x - y\| \le \|T(x) - T(y)\| \le C\|x - y\|, \qquad \text{for any } x, y \in \mathcal{S}_0. \tag{D.88}$$

For an integer $s = \lfloor L^{\frac{1}{K-1}} - 1 \rfloor$, we consider the following choice of centers:

$$\Big\{T(x) : x \in \mathcal{S}_0, \text{ entries of } x \text{ take value on } \{0, \frac{1}{s}, ..., \frac{s-1}{s}, 1\}\Big\}.$$

The total number of centers is bounded by $(s+1)^{K-1} \le L$. We then assign each r_i to the nearest center. The ℓ^∞-distance from each π_i to the nearest x above is at most $1/s$, so the Euclidean distance is at most $\sqrt{K}/s$; combining it with (D.88), the Euclidean distance from $r_i = T(\pi_i)$ to the nearest $T(x)$ above is at most $C\sqrt{K}/s$. It follows that

$$RSS(L) \le n(C\sqrt{K}/s)^2.$$

The choice of s guarantees that $s > L^{\frac{1}{K-1}} - 2$. As a result, for a constant $\tilde{c}$ that does not depend on L,

$$RSS(L) \le n \cdot \tilde{c} L^{-\frac{2}{K-1}}. \tag{D.89}$$

We are now ready to prove the first bullet point. Note that each $\hat{r}_i$ is within a distance $C\hat{h}$ to the corresponding r_i and that all the r_i's are in the Ideal Simplex. Hence, all data points $\{\hat{r}_i\}_{i=1}^n$ are within a distance $C\hat{h}$ to the Ideal Simplex. It is easy to see that all local centers output by k-means must also be within a distance $C\hat{h}$ to the Ideal Simplex. What remains is to show that there is at least one local center within a distance of $C\hat{h}$ to each true vertex v_k. Fix v_k. Our strategy is as follows: for a constant ℓ_0 to be decided,

(a) We first show that there exists at least one local center that is with a distance ℓ_0 to v_k.
(b) We then show that, for each local center within a distance ℓ_0 to v_k, the associated data cluster consists of only pure $\hat{r}_i$ from community k.

Then, by the nature of k-means, such a local center equals to the average of all the $\hat{r}_i$ assigned to this cluster. Since each $\hat{r}_i$ corresponds to a pure node of community k, it is within a distance $C\hat{h}$ to v_k. As a result, the local center must also be within a distance $C\hat{h}$ to v_k. This gives the first bullet point.

What remains is to prove (a) and (b). Fix v_k. Consider (a). Suppose there are no local centers within a distance ℓ_0 to v_k. Then, each pure r_i from community k has a distance $> \ell_0$ to the nearest local center; hence, the distance from $\hat{r}_i$ to the nearest local center is at least $\ell_0 - C\hat{h} \geq \ell_0/2$. At the same time, by the generating process of π_i's, with probability $1-o(n^{-3})$, the number of pure nodes of community k is at least $n\epsilon_k/2$. These pure nodes contribute a sum-of-squares of

$$\geq (n\epsilon_k/2) \cdot (\ell_0/2)^2 = n(\ell_0^2\epsilon_k/8).$$

Additionally, the mixed $\hat{r}_i$'s are assigned to at most L clusters. Since $\|\hat{r}_i - x\|^2 \geq \|r_i - x\|^2/2 - O(\hat{h}^2)$ for any point x, we immediately know that the sum-of-squares contributed by mixed $\hat{r}_i$'s is

$$\geq \frac{1}{2}RSS(L) - O(n\hat{h}^2).$$

Combining the above, the objective attained by k-means is

$$\geq \frac{1}{2}RSS(L) + n(\ell_0^2\epsilon_k/9) \tag{D.90}$$

At the same time, we construct an alternative solution by letting $(L-K)$ of the local centers be those associated with $RSS(L-K)$, letting the remaining K centers be $v_1, v_2, \ldots, v_K$, and assigning each $\hat{r}_i$ to the center closest to the corresponding r_i. Since $\|\hat{r}_i - x\|^2 \leq 2\|r_i - x\|^2 + O(\hat{h}^2)$, the sum of squares attained by this solution is

$$\leq 2RSS(L-K) + O(n\hat{h}^2). \tag{D.91}$$

A contradiction is obtained as long as

$$\begin{aligned} 2RSS(L-K) - \frac{1}{2}RSS(L+K) &< n(\ell_0^2\epsilon_k/9) - O(n\hat{h}^2) \\ &< n(\ell_0^2/10). \end{aligned}$$

According to (D.89), the above is true if we choose $L > (20\tilde{c}/\ell_0^2)^{\frac{K-1}{2}}$. This proves (a).

Consider (b). Fix k. Let $\hat{m}^*$ be a local center such that $\|\hat{m}^* - v_k\| \leq \ell_0$. By the assumption (3.20), for any $\pi_i \neq e_k$, its distance to e_k (e_k is the k-th standard basis of $\mathbb{R}^K$) is at least c_0. Combining it with (D.88), for any node i that is not a pure node of community k, the distance from r_i to v_k is at least $C^{-1}c_0$. As a result, for any such node,

$$\|\hat{r}_i - \hat{m}^*\| \geq C^{-1}c_0 - \ell_0 - C\hat{h}.$$

By taking $\ell_0 = C^{-1}c_0/4.1$, for any node i not pure of community k,

$$\text{the distance from } \hat{r}_i \text{ to the center } \hat{m}^* \text{ is at least } 3\ell_0. \tag{D.92}$$

We shall also show that, for any node i not pure of community k,

$$\text{the distance from } \hat{r}_i \text{ to the nearest center is at most } 2.5\ell_0. \tag{D.93}$$

By (D.92)-(D.93), these nodes cannot be assigned to $\hat{m}^*$. Therefore, the cluster associated with $\hat{m}^*$ consists of only those $\hat{r}_i$ such that i is a pure node of community k. This proves (b).

What remains is to prove (D.93). If i is a pure node of a different community ℓ, then by (a) above, the distance from $r_i = v_\ell$ to the nearest center is $\ell_0 + C\hat{h} < 2.5\ell_0$. Hence, we only need to consider i that is a mixed node. Since $\max_i \|\hat{r}_i - r_i\| \leq C\hat{h} \ll 0.5\ell_0$, it suffices to show that

$$\text{the distance from a mixed } r_i \text{ to the nearest center is at most } 2\ell_0. \tag{D.94}$$

Let $\mathcal{S}_0 = \mathcal{S}_0(e_1, \ldots, e_K) \in \mathbb{R}^K$ be the standard $(K-1)$-simplex, and denote by $\mathcal{B}(x; c)$ an open ball in $\mathcal{S}_0$ centered at x with a radius c; we notice that here an "open ball" means the intersection of $\mathcal{S}_0$ and an open ball in $\mathbb{R}^K$.

Let $\bar{\mathcal{R}}$ be the closure of $\mathcal{R}$, where $\mathcal{R}$ is the support of $f(\cdot)$. We consider the open cover of $\bar{\mathcal{R}}$:

$$\big\{\mathcal{B}(x, C^{-1}\ell_0) : x \in \mathcal{R}\big\}.$$

Since $\bar{\mathcal{R}}$ is closed and bounded, it is a compact set. According to the Borel-Lebesgue covering theorem, the above open cover has a finite sub-cover:

$$\big\{\mathcal{B}(x_1, C^{-1}\ell_0), \mathcal{B}(x_2, C^{-1}\ell_0), \ldots, \mathcal{B}(x_p, C^{-1}\ell_0)\big\}, \qquad \text{where } x_1, \ldots, x_p \in \mathcal{R}.$$

This means each $\pi_i \neq e_k$ is contained in one $\mathcal{B}(x_j, C^{-1}\ell_0)$. Recalling that T is the mapping in (D.88), define

$$\mathcal{B}_j^* = T\big(\mathcal{B}(x_j, C^{-1}\ell_0)\big), \qquad 1 \le j \le p.$$

Then, $r_i = T(\pi_i)$ is contained in $\mathcal{B}_j^*$. Moreover, for any $y, \tilde{y} \in \mathcal{B}_j^*$, $\|y - \tilde{y}\| \le C \max_{x, \tilde{x} \in \mathcal{B}(x_j, C^{-1}\ell_0)} \le 2\ell_0$. Therefore, if we can show that

$$\text{each } \mathcal{B}_j^* \text{ contains at least one local center}, 1 \le j \le p, \tag{D.95}$$

then the distance from r_i to this local center is bounded by $2\ell_0$. This gives (D.94), and in turn gives (D.93).

What remains is to prove (D.95). Note that $\mathcal{R}$ is an open set. By definition of open sets, for each of $x_1, x_2, \ldots, x_p$, there is a $\tau_j > 0$ such that the closed ball $\bar{\mathcal{B}}(x_j, \tau_j)$ is contained in $\mathcal{R}$. We define the closed balls

$$\mathcal{BB}_j \equiv \bar{\mathcal{B}}\big(x_j, \min\{\tau_j, C^{-1}\ell_0/2\}\big), \qquad 1 \le j \le p.$$

Let $\omega_j = \int f(\pi) 1\{\pi \in \mathcal{BB}_j\} d\pi = (1 - \sum_{k=1}^K \epsilon_k) \int g(\pi) 1\{\pi \in \mathcal{BB}_j\} d\pi$, $1 \le j \le p$. Note that each of these closed balls is contained in the support of g with a nonzero radius and that g as a probability density is measurable. We immediately know that $\omega_j > 0$. From the assumption (3.21) and elementary large-deviation inequalities (e.g., the Hoeffding's inequality), we know that with probability $1 - o(n^{-3})$, for $1 \le j \le p$,

$$\text{the number of } \pi_i\text{'s contained in } \mathcal{BB}_j \text{ is at least } n\omega_j/2. \tag{D.96}$$

With (D.96), we now prove (D.95) by contradiction. Suppose (D.95) does not hold, i.e., there exists $\mathcal{B}_j^*$ such that

$$\mathcal{B}_j^* \cap \{\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L\} = \emptyset,$$

where $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L$ are the local centers output by k-means. By definition of $\mathcal{B}_j^*$ and the fact that T is a one-to-one mapping, we have

$$\mathcal{B}(x_j, C^{-1}\ell_0) \cap \big\{T^{-1}(\hat{m}_1), T^{-1}(\hat{m}_2), \ldots, T^{-1}(\hat{m}_L)\big\} = \emptyset.$$

Note that $\mathcal{BB}_j$ is a ball also centered at x_j but with a radius no larger than half of the radius of $\mathcal{B}(x_j, C^{-1}\ell_0)$. As a result, for any $x \in \mathcal{BB}_j$, its distance to the nearest one of $T^{-1}(\hat{m}_1), \cdots, T^{-1}(\hat{m}_L)$ is at least $C^{-1}\ell_0/2$; combining it with (D.88), the distance from $T(x)$ to the nearest one of $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L$ is at least $C^{-2}\ell_0/2$. It follows that

$$\text{for any } \pi_i \in \mathcal{BB}_j, \min_{1 \leq s \leq L} \|r_i - \hat{m}_s\| \geq C^{-2}\ell_0/2.$$

Note that $\max_i \|\hat{r}_i - r_i\| \leq C\hat{h} = o(1)$. We further conclude that

$$\text{for any } \pi_i \in \mathcal{BB}_j, \text{ the distance from } \hat{r}_i \atop \text{to the nearest local center is } \geq C^{-2}\ell_0/3. \tag{D.97}$$

Combining (D.96)-(D.97), the sum-of-squares attained by k-means is

$$\geq (C^{-2}\ell_0/3)^2 \cdot (n\omega_j/2) \geq n(\omega_{\min} C^{-4}\ell_0^2/18),$$

where $\omega_{\min} = \min\{\omega_1, \ldots, \omega_p\}$. At the same time, the objective attained by k-means should be

$$\leq RSS(L) + n(C\hat{h}^2).$$

A contradiction is obtained as long as

$$RSS(L) < n(\omega_{\min} C^{-4}\ell_0^2/18) - n(C\hat{h}^2). \tag{D.98}$$

Comparing it with (D.89), as long as $L > (\frac{19C^4\tilde{c}}{\ell_0^2\omega_{\min}})^{\frac{K-1}{2}}$, the inequality (D.98) will be true. We then have a contradiction, which implies that (D.95) must hold. The proof is now complete.

D.5. Proof of Lemma 3.5. We aim to prove the following lemma:

Lemma D.4. *Suppose the conditions of Lemma 3.5 hold. We apply the SVS algorithm to $\{\hat{r}_i\}_{i=1}^n$ with $L = \hat{L}_n(A)$, where $\hat{L}_n(A)$ is defined in (3.25). Let $\hat{h}^* = \sqrt{n^{-1}\sum_{i=1}^n \|H\hat{r}_i - r_i\|^2}$ and $\hat{h} = \max_{1\leq i\leq n} \|H\hat{r}_i - r_i\|$. With probability $1 - o(n^{-3})$, the following statements are true.*

- *$\hat{L}_n(A) = L_0 + K$.*
- *The local clustering sub-step identifies $(L_0 + K)$ local centers, where there is a unique $(K-1)$-simplex such that K of these centers (denoted by $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_K$) are its vertices, and all other centers are within a distance of $C\hat{h}$ to this simplex. These K local centers will be identified by the combinatorial search sub-step.*

- *The above K local centers satisfy $\hat{v}_k = |\mathcal{N}_k|^{-1}\sum_{i\in\mathcal{N}_k}\hat{r}_i$, $1 \le k \le K$. As a result, up to a permutation of estimated vertices,* $\max_{1\le k\le K}\|H\hat{v}_k - v_k\| \le C\hat{h}^*$.

The claim of Lemma 3.5 is included in the third bullet point of Lemma D.4.

Proof of Lemma D.4: As explained in the proof of Lemma 3.3, we only consider the case that H is the identity matrix. By Theorem 2.1 and Lemma C.5, with probability $1 - o(n^{-3})$,

(D.99)

$$\hat{h} \equiv \max_{1\le i\le n}\|\hat{r}_i - r_i\| \le \frac{Cerr_n}{\beta_n}, \quad n(h^*)^2 \equiv \sum_{i=1}^n \|\hat{r}_i - r_i\|^2 \le \frac{Cn(err_n^*)^2}{\beta_n^2},$$

where we have absorbed the factors of K into the constants. We also note that $err_n^* \le err_n/\sqrt{\log(n)}$. Below, we restrict to the event of (D.99).

First, we study $\hat{L}_n(A)$. Recall that $\gamma_1, \gamma_2, \ldots, \gamma_{L_0}$ are as in (3.23). Let T be the mapping as in (D.88); note that $T(\pi_i) = r_i$ for $1 \le i \le n$. Introduce

$$m_j = T(\gamma_j), \qquad 1 \le j \le L_0.$$

By (D.88), the assumptions (3.23)-(3.24) imply that the distance between any two of $\{v_1, v_2, \ldots, v_K, m_1, m_2, \ldots, m_{L_0}\}$ is at least c, and $\max_{i\in\mathcal{M}_j}\|r_i - m_j\| \le C_1/\log(n)$, where $c > 0$ and $C_1 > 0$ are constants. In particular,

$$\alpha_n^2 \le \frac{C|\mathcal{M}|}{n\log(n)}, \qquad \text{where} \quad \alpha_n^2 \equiv n^{-1}\sum_{j=1}^{L_0}\sum_{i\in\mathcal{M}_j}\|r_i - m_j\|^2.$$

We now study $\epsilon_L(\hat{R})$. When $L = L_0 + K$, by choosing this choice of centers $\{v_1, \ldots, v_K, m_1, \ldots, m_{L_0}\}$, it is easy to see that

$$\epsilon_{L_0+K}(\hat{R}) \le n\alpha_n^2 + C\sum_{i=1}^n \|\hat{r}_i - r_i\|^2 \le \frac{C|\mathcal{M}|}{\log(n)}, \tag{D.100}$$

where the last inequality is due to (D.99) and the assumption that $|\mathcal{M}| \ge n\beta_n^{-2}err_n^2 \ge n\beta_n^{-2}(err_n^*)^2\log(n)$. When $K \le L < L_0 + K$, suppose there are L_1 of $\{v_1, v_2, \ldots, v_K\}$ and L_2 of $\{m_1, m_2, \ldots, m_{L_0}\}$ such that no local centers are within a distance of $c/3$ of them. Since the distance between any two of $\{v_1, v_2, \ldots, v_K, m_1, m_2, \ldots, m_{L_0}\}$ is at least c, we have that $(L_1 + L_2)$ is at least $(L_0 + K) - L$. For any such v_k and $i \in \mathcal{N}_k$ or such m_j and $i \in \mathcal{M}_j$, the distance from $\hat{r}_i$ to the nearest local center is at least $c/3 - \hat{h} \geq c/4$. It follows that

$$\epsilon_L(\hat{R}) \geq (c/4)^2 \cdot (L_1 \min_k |\mathcal{N}_k| + L_2 \min_j |\mathcal{M}_j|) \geq C|\mathcal{M}|, \tag{D.101}$$

where the last inequality is due to $\min_k |\mathcal{N}_k| \geq c_1 n$ and $\min_j |\mathcal{M}_j| \geq c_4|\mathcal{M}|$. At the same time, by choosing the centers to be $\{v_1, v_2, \ldots, v_K\}$ and $(L-K)$ of $\{m_1, m_2, \ldots, m_{L_0}\}$,

$$\epsilon_L(\hat{R}) \leq C(L_0 + K - L)|\mathcal{M}| + C\sum_{i=1}^{n} \|\hat{r}_i - r_i\|^2 \leq C|\mathcal{M}|. \tag{D.102}$$

By (D.100)-(D.102),

$$\epsilon_L(\hat{R})/\epsilon_{L-1}(\hat{R}) \begin{cases} \leq C/\log(n), & L = L_0 + K, \\ \geq C, & K + 1 \leq L \leq L_0 + K. \end{cases}$$

Hence, the definition of $\hat{L}_n(A)$ in (3.25) yields $\hat{L}_n(A) = L_0 + K$. This proves the first bullet point.

Next, we consider the second bullet point. Suppose for L_1 of $\{v_1, v_2, \ldots, v_K\}$ and L_2 of $\{m_1, m_2, \ldots, m_{L_0}\}$, there are no local centers are within a distance of $c/4$ of them. When $L_1 + L_2 \geq 1$, using similar arguments as those for proving (D.101), we can see that the associated sum-of-squares is lower bounded by $C|\mathcal{M}|$. However, in (D.100), we have seen that the sum-of-squares attained by k-means is at most $C|\mathcal{M}|/\log(n)$. Hence, the above situation is impossible, i.e., for each of $\{v_1, v_2, \ldots, v_K, m_1, \ldots, m_{L_0}\}$, there is at least one local center within a distance $c/4$ to it. Since that the distance between any two of $\{v_1, v_2, \ldots, v_K, m_1, \ldots, m_{L_0}\}$ is at least c, these $(L_0 + K)$ local centers must be distinct. Noting that there are at most $\hat{L}_n(A) = L_0 + K$ cluster centers in total, we find that

$$\begin{gathered} \text{there is exactly one local center within a distance } c/4 \\ \text{to each of } \{v_1, v_2, \ldots, v_K, m_1, m_2 \ldots, m_{L_0}\}. \end{gathered} \tag{D.103}$$

Denote by $\hat{m}^*_{(k)}$ the local center nearest to v_k and by $\hat{m}_{(j)}$ the local center nearest to m_j, $1 \leq k \leq K$, $1 \leq j \leq L_0$. For any $i \in \mathcal{N}_k$, the distance from $\hat{r}_i$ to $\hat{m}^*_{(k)}$ is at most $c/4 + O(\hat{h}) \leq c/3$, but its distance to any other local center is at least $c - c/4 - O(\hat{h}) \geq 2c/3$; hence, $\hat{r}_i$ can only be assigned to the cluster associated with $\hat{m}^*_{(k)}$. Similarly, for any $i \in \mathcal{M}_j$, the distance from $\hat{r}_i$ to $\hat{m}_{(j)}$ is at most $c/4 + O(\frac{1}{\log(n)}) + O(\hat{h}) \leq c/3$, but the distance to any other local center is at least $c - c/4 - O(\frac{1}{\log(n)}) - O(\hat{h}) \geq 2c/3$; so $\hat{r}_i$ must be assigned to $\hat{m}_{(j)}$. We have proved that

$$\text{(D.104)} \quad \begin{cases} \text{the cluster associated with } \hat{m}^*_{(k)} \text{ is } \{\hat{r}_i : i \in \mathcal{N}_k\}, 1 \leq k \leq K, \\ \text{the cluster associated with } \hat{m}_{(j)} \text{ is } \{\hat{r}_i : i \in \mathcal{M}_j\}, 1 \leq j \leq L_0. \end{cases}$$

Then, it is easy to see that

- All the local centers are within a distance $\hat{h}$ to the Ideal Simplex.
- Each $\hat{m}^*_{(k)}$ is within a distance $C\hat{h}$ to v_k, $1 \leq k \leq K$.
- Each $\hat{m}_{(j)}$ is within a distance $C/\log(n)$ to m_j, $1 \leq j \leq L_0$.

We now show that $\hat{m}^*_{(1)}, \hat{m}^*_{(2)}, \ldots, \hat{m}^*_{(K)}$ will be selected by the combinatorial search. The proof is similar to that of Lemma 3.3 but is simpler. Suppose one $\hat{m}^*_{(k)}$ is not selected by the combinatorial search. By (D.104), the other local centers are contained in the convex hull $\mathcal{H}\{\hat{r}_i : i \notin \mathcal{N}_k\}$. Hence, the estimated simplex $\hat{\mathcal{S}} \subset \mathcal{H}\{\hat{r}_i : i \notin \mathcal{N}_k\}$. We notice that the distance from e_k to the convex hull of all $\pi_i \neq e_k$ is lower bounded by a constant, as a result of the assumptions (3.23)-(3.24). Using (D.88), we know that the distance from v_k to the convex hull $\mathcal{H}\{r_i : i \notin \mathcal{N}_k\}$ is also lower bounded by a constant. Then,

$$\begin{aligned} d(\hat{m}^*_{(k)}, \hat{\mathcal{S}}) &\geq d\big(\hat{m}^*_{(k)},\ \mathcal{H}\{\hat{r}_i : i \notin \mathcal{N}_k\}\big) \\ &\geq d(v_k, \mathcal{H}\{r_i : i \notin \mathcal{N}_k\}) - O(\hat{h}) \\ &\geq C. \end{aligned}$$

At the same time, if we pick the K local centers $\hat{m}^*_{(1)}, \hat{m}^*_{(2)}, \ldots, \hat{m}^*_{(K)}$,

$$\max_{1 \leq j \leq L_0} d\big(\hat{m}_j, \mathcal{S}(\hat{m}^*_{(1)}, \hat{m}^*_{(2)}, \ldots, \hat{m}^*_{(K)})\big) \leq C\hat{h}.$$

This yields a contradiction since $\hat{h} = o(1)$. As a result, all of $\hat{m}^*_{(1)}, \hat{m}^*_{(2)}, \ldots, \hat{m}^*_{(K)}$ will be selected by the combinatorial search.

Last, we prove the third bullet point. So far, we have seen that $\hat{v}_k = \hat{m}^*_{(k)}$ (up to a label permutation). By (D.104) and the nature of k-means solutions,

$$\hat{v}_k = |\mathcal{N}_k|^{-1} \sum_{i \in \mathcal{N}_k} \hat{r}_i, \qquad 1 \leq k \leq K.$$

We note that $0 \leq \sum_{i \in \mathcal{N}_k} \|\hat{r}_i - \hat{v}_k\|^2 = \sum_{i \in \mathcal{N}_k} \{\|\hat{r}_i - v_k\|^2 - 2(\hat{v}_k - v_k)'(\hat{r}_i - v_k) + \|\hat{v}_k - v_k\|^2)\} = \sum_{i \in \mathcal{N}_k} \|\hat{r}_i - v_k\|^2 - |\mathcal{N}_k|\|\hat{v}_k - v_k\|^2$. As a result,

$$\|\hat{v}_k - v_k\|^2 \leq \frac{1}{|\mathcal{N}_k|} \sum_{i \in \mathcal{N}_k} \|\hat{r}_i - v_k\|^2 \leq \frac{1}{|\mathcal{N}_k|} \sum_{i=1}^{n} \|\hat{r}_i - r_i\|^2, \quad 1 \leq k \leq K.$$

Since $|\mathcal{N}_k| \geq c_1 n$, it follows that

$$\max_{1\leq k\leq K} \|\hat{v}_k - v_k\| \leq C\sqrt{n^{-1}\sum_{i=1}^{n} \|\hat{r}_i - r_i\|^2} \leq C\hat{h}^*. \tag{D.105}$$

This proves the third bullet point. □

Remark. In the above proof, we have used no property of the joint distribution of $\{\hat{r}_i - r_i\}_{i=1}^n$. In other applications, we may encounter VH settings where the noise $(\hat{r}_i - r_i)$ are *iid* sub-Gaussian random vectors. In this case, we can obtain a faster rate. Suppose $\sigma^2 = E(\|\hat{r}_i - r_i\|^2)$. Note that we have proved $\hat{v}_k = |\mathcal{N}_k|^{-1}\sum_{i\in\mathcal{N}_k} \hat{r}_i$. It follows that

$$\hat{v}_k - v_k = \frac{1}{|\mathcal{N}_k|}\sum_{i\in\mathcal{N}_k} (\hat{r}_i - r_i).$$

By large deviation inequalities of *iid* sub-Gaussian random vectors, the right hand side is bounded by $C|\mathcal{N}_k|^{-1/2}\sigma$ with high probability. In comparison, the right hand side of (D.105) is bounded by $C\sigma$ with high probability.

D.6. Proof of Lemma 3.6. As explained in the proof of Lemma 3.3, without loss of generality, we consider the case that H equals to the identity matrix. Recall that K is a constant here. Write $\hat{h}^* = \sqrt{n^{-1}\sum_{i=1}^n \|\hat{r}_i - r_i\|^2}$ and $\hat{h} = \max_{1\leq i\leq n} \|\hat{r}_i - r_i\|$.

By Lemma B.3, there is a constant $c > 0$ such that the distance between any two distinct vertices is at least c. If for some v_k there is no local center within a distance $c/4$ to it, then the sum-of-squares attained by k-means is at least

$$(c/4 - \hat{h})^2 \cdot |\mathcal{N}_k| \geq Cn.$$

However, by considering the choice of centers $\{v_1, v_2, \ldots, v_K\}$ (when $L > K$, we select some of v_k's more than one time), it is seen that the sum-of-squares attained by k-means is

$$\begin{aligned} &\leq \sum_{i\in\mathcal{M}} \min_{1\leq k\leq K} \|r_i - v_k\|^2 + C\sum_{i=1}^n \|\hat{r}_i - r_i\|^2 \\ &\leq C|\mathcal{M}| + Cn(\hat{h}^*)^2 = o(n). \end{aligned}$$

The above yield a contradiction. Hence, we conclude that for each v_k, there is at least one local center within a distance $c/4$ to it.

Since $L = K$, up to a permutation of estimated vertices, $\hat{v}_k$ is the unique local center within a distance $c/4$ to v_k, $1 \leq k \leq K$. For each k and $i \in \mathcal{N}_k$, the distance from $\hat{r}_i$ to any other local center is at least $3c/4 - \hat{h}$, while its distance to $\hat{v}_k$ is at most $c/4 + \hat{h}$. Therefore, the $\hat{r}_i$'s of all pure nodes are associated with their own corresponding $\hat{v}_k$. It means for $1 \leq k \leq K$,

$$\text{the cluster associated with } \hat{v}_k \text{ consists of all } \hat{r}_i \text{ with } i \in \mathcal{N}_k \text{ and some of } \hat{r}_i \text{ with } i \in \mathcal{M}. \tag{D.106}$$

Then, by the nature of k-means solutions, $\hat{v}_k$ equals to the average of all the $\hat{r}_i$'s in this cluster. First, let $\widehat{\mathcal{N}}_k$ be the set of nodes such that $\hat{r}_i$ is assigned to the local center $\hat{v}_k$. By (D.106), $|\mathcal{N}_k| \leq |\widehat{\mathcal{N}}_k| \leq |\mathcal{N}_k| + |\mathcal{M}|$. Second, introduce $\hat{v}_k^* = |\mathcal{N}_k|^{-1} \sum_{i \in \mathcal{N}_k} \hat{r}_i$. In (D.105), we have seen that $\|\hat{v}_k^* - v_k\| \leq C\hat{h}^*$. Last, since $\max_k \|v_k\| \leq C$ and $\max_i \|\hat{r}_i - r_i\| = \hat{h} = o(1)$, we have $\|\hat{r}_i - v_k\| \leq C$ for all i. Combining the above gives

$$\begin{aligned}
\|\hat{v}_k - v_k\| &= \left\| \frac{1}{|\widehat{\mathcal{N}}_k|} \sum_{i \in \widehat{\mathcal{N}}_k} (\hat{r}_i - v_k) \right\| \\
&\leq \left\| \frac{1}{|\widehat{\mathcal{N}}_k|} \sum_{i \in \mathcal{N}_k} (\hat{r}_i - v_k) \right\| + \left\| \frac{1}{|\widehat{\mathcal{N}}_k|} \sum_{i \in \widehat{\mathcal{N}}_k \setminus \mathcal{N}_k} (\hat{r}_i - v_k) \right\| \\
&\leq C \left\| \frac{1}{|\mathcal{N}_k|} \sum_{i \in \mathcal{N}_k} (\hat{r}_i - v_k) \right\| + C|\widehat{\mathcal{N}}_k|^{-1}|\mathcal{M}| \\
&\leq C\|\hat{v}_k^* - v_k\| + C|\mathcal{N}_k|^{-1}|\mathcal{M}| \\
&\leq C\hat{h}^* + C\zeta_n, \qquad \text{(D.107)}
\end{aligned}$$

where the last line is from $|\mathcal{M}| \leq n\zeta_n$ and $|\mathcal{N}_k| \geq c_1 n$. Since we have assumed $\zeta_n \leq C\beta_n^{-1} err_n$, it follows that

$$\max_{1 \leq k \leq K} \|\hat{v}_k - v_k\| \leq C\hat{h}^* + C\beta_n^{-1} err_n.$$

This gives the main claim. Furthermore, on the event F_n, $\hat{h} = \max_i \|\hat{r}_i - r_i\| > \beta_n^{-1} err_n$; it implies that the above right hand side is $\leq C\hat{h}^* + C\hat{h} \leq C\hat{h}$. Hence, the algorithm is efficient.

Remark: The event F_n^c does not affect the error bound of Mixed-SCORE. To see this, note that on the event F_n^c, we have $\max_{1 \leq i \leq n} \|\hat{r}_i - r_i\| \leq \beta_n^{-1} err_n$. Additionally, by (D.107), $\max_{1 \leq k \leq K} \|\hat{v}_k - v_k\| \leq C\beta_n^{-1} err_n$. We can plug them into (6.30) in the proof of Theorem 2.2 to obtain the same conclusion.

D.7. Proof of Theorem 3.1. Following the proof of Theorem 2.2, we can show that

$$\|\hat{\pi}_i - \pi_i\|_1 \leq C\|H\hat{r}_i - r_i\| + C \max_{1 \leq k \leq K} \|H\hat{v}_k - v_k\| + CKerr_n^*. \tag{D.108}$$

The proof of (D.108) is almost identical to the proof of (6.30), where the only difference is that we replace err_n in (6.36) by err_n^*. Note that the last line of (6.36) is due to $err_n \geq \|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1}$. Similarly, we have

$$err_n^* = [\|\theta\|/(\theta_{\min}\sqrt{n})] \cdot \|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1} \geq \|\theta\|^{-2}\sqrt{\theta_{\max}\|\theta\|_1}.$$

Therefore, we can prove (D.108) in the same way.

It follows from (D.108) that

$$n^{-1}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|_1^2 \leq Cn^{-1}\sum_{i=1}^n \|H\hat{r}_i - r_i\|^2 + C\Big(\max_{1\leq k\leq K}\|H\hat{v}_k - v_k\|\Big)^2 + CK^2(err_n^*)^2.$$

In the setting of Lemma 3.5, the second term is bounded by the first term. In the setting of Lemma 3.6, by (D.107) and the condition $\zeta_n \leq C\beta_n^{-1}err_n^*$, the second term is bounded by the first term plus $C\beta_n^{-2}(err_n^*)^2$. Additionally, by Lemma C.5, the first term is upper bounded by $CK^3\beta_n^{-1}(err_n^*)^2$. Combining the above gives

$$n^{-1}\sum_{i=1}^n \|\hat{\pi}_i - \pi_i\|_1^2 \leq CK^3\beta_n^{-1}(err_n^*)^2 \leq CK^3\beta_n^{-1}err_n^2/\log(n),$$

where the last inequality is due to $err_n^* = [\|\theta\|/(\sqrt{n}\theta_{\max})] \cdot err_n/\sqrt{\log(n)} \leq err_n/\sqrt{\log(n)}$. This proves the claim.

J. Jin and S. Luo
Department of Statistics
Carnegie Mellon University
Pittsburgh, Pennsylvania, 15213
USA
E-mail: jiashun@stat.cmu.edu
shengmil@andrew.cmu.edu

Z. Ke
Department of Statistics
Harvard University
Cambridge, Massachusetts, 02138
USA
E-mail: zke@fas.harvard.edu