Optimal Permutation Recovery in Permuted Monotone Matrix Model

Rong Ma[1], T. Tony Cai[2] and Hongzhe Li[1]

Department of Biostatistics, Epidemiology and Informatics[1]

Department of Statistics[2]

University of Pennsylvania

Philadelphia, PA 19104

Abstract

Motivated by recent research on quantifying bacterial growth dynamics based on genome assemblies, we consider a permuted monotone matrix model $Y = \Theta\Pi + Z$, where the rows represent different samples, the columns represent contigs in genome assemblies and the elements represent log-read counts after preprocessing steps and Guanine-Cytosine (GC) adjustment. In this model, Θ is an unknown mean matrix with monotone entries for each row, Π is a permutation matrix that permutes the columns of Θ, and Z is a noise matrix. This paper studies the problem of estimation/recovery of Π given the observed noisy matrix Y. We propose an estimator based on the best linear projection, which is shown to be minimax rate-optimal for both exact recovery, as measured by the 0-1 loss, and partial recovery, as quantified by the normalized Kendall's tau distance. Simulation studies demonstrate the superior empirical performance of the proposed estimator over alternative methods. We demonstrate the methods using a synthetic metagenomics dataset of 45 closely related bacterial species and a real metagenomic dataset to compare the bacterial growth dynamics between the responders and the non-responders of the IBD patients after 8 weeks of treatment.

KEY WORDS: Kendall's tau; Microbiome growth dynamics; Minimax lower bound; Sorting.

1 INTRODUCTION

1.1 A Motivation Example from Microbiome Studies

The statistical problem considered in this paper is motivated by the problem of estimating the bacterial growth dynamics based on shotgun metagenomics data (Myhrvold et al., 2015; Abel et al., 2015; Korem et al., 2015; Brown et al., 2016). The growth dynamics of microbial populations

reflects their physiological states and drives variation of microbial compositions, which provide important feature summary of the microbes in a given community. One way of studying such communities is through shotgun metagenomic sequencing, which involve direct DNA sequencing of all the microbiome genomes in a given microbial community. Korem et al. (2015) presented the first paper on quantifying the bacterial growth dynamics based on shotgun metagenomics data, where the uneven sequencing read coverage resulting from the bidirectional DNA replications provides information on the rates of microbial DNA replications. For bacterial species with known complete genome sequences, Korem et al. (2015) proposed to use the peak-to-trough ratio (PTR) of read coverages to quantify the bacterial growth dynamics after aligning the sequencing reads to the complete genome sequences.

However, in many applications, it is of importance to quantify the bacterial growth dynamics based on genome assemblies for the bacterial species with unknown genomes. These genome assemblies may represent new bacterial species that we have seen or sequenced before. The genome assembly of a bacterium species consists of a collection of contigs (called bin) constructed based on the overlapping of the sequencing reads (Li et al., 2015; Wu et al., 2014). Compared to the complete genome, the genome assembled bins are more fragmented and often contained errors or contaminations. The noisy read coverage data due to intraspecific variations, interspecific/intraspecific repeated sequences, limited sequencing depths and the inability of binning algorithms to correctly cluster all the contigs further complicate the estimation of growth dynamics based on read coverages of the contigs. Besides these noisy count data, one key difficulty in estimating the growth dynamic based on contig counts is that the accurate locations of the contigs on the original genome are unknown. It is therefore not feasible to measure the microbial growth rate directly using peak-to-trough coverage ratio for the assembled genomes (Brown et al., 2016; Gao and Li, 2018).

Brown et al. (2016) presented the first method (called iRep) of estimating the bacterial growth dynamics based on genome assemblies, where the contigs are ordered based on the GC-adjusted counts for each sample separately. However, due to noise in the count data, such an ordering method often leads to wrong ordering of the contigs and therefore inaccurate estimates of the growth dynamics. Gao and Li (2018) developed a computational algorithm, DEMIC, to accurately compare growth dynamics of a given assembled species existing in multiple samples by taking advantage of highly fragmented contigs assembled in typical metagenomics studies. One key step of DEMIC is to apply a principal components analysis (PCA)-based method to recover the true ordering of the contigs along the underlying unknown bacterial complete genomes. Gao and Li (2018) reported excellent empirical performance of DEMIC over existing methods. The goal of this paper is to provide a rigorous statistical framework to study the problem of optimal permutation recovery in a permuted monotone matrix model.

1.2 A Permuted Monotone Matrix Model

For a given genome assembly with p contigs, DEMIC first obtains the read coverage for each of the sliding window of size 5000 bps, denoted by X_{ijl} for the ith sample, jth contig and kth window. In order to account for the GC-content of the kth window, Gao and Li (2018) considered the following mixed-effects model,

$$\log_2 X_{ijk} = \alpha + GC_{jk}\beta + W_{ij} + e_{ijk},$$

where GC_{jk} is the centred GC count of the kth window of the jth contig, W_{ij} is the sample- and contig- specific random intercept, α is the intercept, β is the regression coefficient, and e_{ijk} is the random error. This model is fitted for each contig to obtain the best linear unbiased predictor of W_{ij}, which is used as the GC-adjusted log-read count Y_{ij} for the ith sample and jth contig. Here Y_{ij} can be regarded as average read coverage over non-overlapping windows of a contig and is approximately normally distributed.

Let Y be the GC-adjusted log-contig count matrix of n samples and p contigs of a genome assembly with Y_{ij} as its entries. Given this, we consider the following permuted monotone matrix model:

$$Y = \Theta\Pi + Z, \tag{1.1}$$

where $\Theta \in \mathbb{R}^{n \times p}$ is an unknown nonnegative signal matrix with nondecreasing rows, $Z \in \mathbb{R}^{n \times p}$ is a zero-mean noise matrix, and $\Pi \in \mathbb{R}^{p \times p}$ is a permutation matrix corresponding to some permutation π from the symmetric group $\mathcal{S}_p$. That is, after a suitable permutation of the columns of Y, all the rows of the mean matrix are nondecreasing sequences. In microbiome applications, Θ is the matrix of true log-coverage of n samples over p contigs along the circular genome of the bacterium, which is generally hypothesized to have non-decreasing rows. Π represents a permutation due to unknown locations of the contigs relative to the replication origin. Throughout this paper, we denote the parameter space

$$(\Theta, \pi) \in \mathcal{D} = \left\{ \Theta = (\theta_{ij}) \in \mathbb{R}^{n \times p}, \pi \in \mathcal{S}_p : 0 \leq \theta_{i,j-1} \leq \theta_{i,j} < \infty \text{ for all } 1 \leq i \leq n, 2 \leq j \leq p \right\}.$$

The focus of this paper is to optimally estimate the permutation π from the noisy observation Y.

1.3 Related Problems and Other Applications

The permutation recovery problem under permuted monotone matrix model bears some similarity to other problems studied in machine learning literature, including the feature matching between two sets of observations (Collier and Dalalyan, 2016) and linear regression model with permuted data, where the correspondences between the response and the predictors are unknown (Pananjady et al., 2016; Slawski and Ben-David, 2017; Pananjady et al., 2017). More recently, Flammarion et al. (2019) considered the problem of statistical seriation, which has a close affinity to our model

(1.1). However, the focus of Flammarion et al. (2019) is to optimally estimate the signal matrix Θ rather than the underlying permutation.

Model (1.1) can be thought as a natural extension of the shape constrained matrix denoising model studied in the isotonic regression literature. Specifically, under Model (1.1) with known $\Pi = \mathbf{I}_p$, risk bounds and the minimax rate-optimal estimator for Θ under the Frobenius norm was obtained in Chatterjee et al. (2015) for $n = 1$ and later in Chatterjee et al. (2018) for general $n > 1$. Using the idea of optimal transport, a minimax optimal estimator of the underlying signals was obtained by Rigollet and Weed (2018). However, their goal is not to recover the underlying permutation.

Besides the microbiome applications, the permuted monotone matrix model is generic and has other applications. For instance, the problem of permutation recovery is usually equivalent to statistical ranking/sorting from noisy observations, which arises commonly in finance (Currie and Pandher, 2011), sport analytics (Deshpande and Jensen, 2016), and recommendation systems (Rendle et al., 2009). Specifically, in the latter case, the task of tag recommendation is to provide a user with a personalized ranked list of tags for a specific item. Under the permuted monotone matrix model, we can treat the entries of Y, say Y_{ij}, as an indicator of the jth tag being related to the ith item by a given customer, and Θ as a probability matrix characterizing the customer's tagging preferences across multiple items. As a result, recovering the underlying permutation provides a solution of a tag recommender.

1.4 Main Contributions and Organization

In this paper, we investigate the problem of permutation recovery in the permuted monotone matrix model (1.1), which relies on certain invariance property of the singular subspace of the monotone matrices. The properties of the proposed method in terms of both the exact and partial recovery are studied in detail. In particular, we obtained regions of the signal-to-noise ratio (defined later as Γ/σ) that are subject to exact/partial recovery (Figure 1). For both exact and partial permutation recovery, we obtained the matching minimax lower bounds and established the minimax rate-optimality of the proposed method over a wide range of parameter space (Figure 1). For partial recovery, the proof of the lower bound relies on a version of Fano's lemma and the sphere packing of the symmetric group equipped with the Kendall's tau metric.

The rest of this paper is organized as follows. After a brief introduction of notation and definitions, we present in Section 2 the proposed permutation estimator. The theoretical properties of the proposed method are studied, first under a more illustrative linear growth model in Section 3 and then under a general growth model in Section 4. Section 5 provides results on minimax lower bounds and the optimality of the proposed estimator. We evaluate the methods using both simulated data, synthetic and real microbiome datasets and compare with other methods in Section 6. In Section 7, we discuss some implications and extensions of the methods. Finally, the proofs
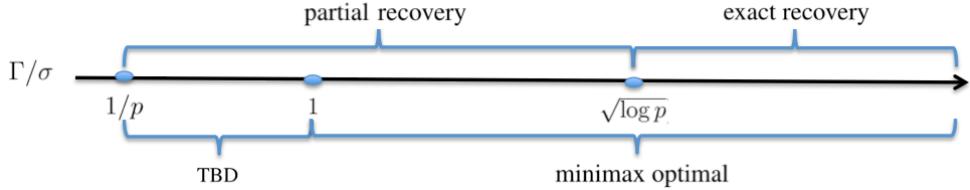
Figure 1: A graphical illustration of the main result obtained in this paper about the regions of the signal-to-noise ratio Γ/σ that correspond to exact/partial recovery, and the region with proved minimax optimality.

of our main results are given in Section 8.

1.5 Notation and Definitions

Throughout, we define the permutation π as a bijection from the set $\{1, 2, ..., p\}$ onto itself. For simplicity, we denote $\pi = (\pi(1), \pi(2), ..., \pi(p))$. All permutations of the set $\{1, 2, ..., p\}$ form a symmetric group, equipped with the function composition operation $\circ$, denoted as $\mathcal{S}_p$. For any $\pi \in \mathcal{S}_p$, we denote $\pi^{-1} \in \mathcal{S}_p$ as its group inverse, so that $\pi \circ \pi^{-1} = \pi^{-1} \circ \pi = id$, and denote $\mathrm{rev}(\pi) = (\pi(p), \pi(p-1), ...\pi(1))$. In particular, we may use π and its corresponding permutation matrix $\Pi \in \mathbb{R}^{p \times p}$ interchangeably, depending on the context. For a vector $\mathbf{a} = (a_1, ..., a_n)^\top \in \mathbb{R}^n$, we define the ℓ_p norm $\|\mathbf{a}\|_p = \left(\sum_{i=1}^n a_i^p\right)^{1/p}$, and the ℓ_∞ norm $\|\mathbf{a}\|_\infty = \max_{1 \leq j \leq n} |a_i|$. For a matrix $\Theta \in \mathbb{R}^{p_1 \times p_2}$, we denote $\Theta_{.i} \in \mathbb{R}^{p_1}$ as its i-th column, $\Theta_{i.} \in \mathbb{R}^{p_2}$ as its i-th row, and denote its (ordered) singular values as $\lambda_1(\Theta) \geq \lambda_2(\Theta) \geq ... \geq \lambda_{\min\{p_1, p_2\}}(\Theta)$. Furthermore, for sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n = o(b_n)$ if $\lim_n a_n/b_n = 0$, and write $a_n = O(b_n)$, $a_n \lesssim b_n$ or $b_n \gtrsim a_n$ if there exists a constant C such that $a_n \leq Cb_n$ for all n. We write $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $a_n \gtrsim b_n$. For a finite set A, we denote $|A|$ as its cardinality. We use the logical symbols $\wedge$ and $\vee$ to represent "and" and "or," respectively. Lastly, $C, C_0, C_1, ...$ are constants that may vary from place to place.

2 PERMUTATION RECOVERY VIA BEST LINEAR PROJECTION

In the following, we first make some key observations about the connection between the underlying permutation π and the column linear projections of the observed matrix Y, which motivate our construction of the proposed estimator.

2.1 Linear Projection

Given the observed noisy matrix Y, we consider the class of the linear projection statistics of the form $w^\top Y \in \mathbb{R}^p$ where $w \in \mathbb{R}^n$ and $\|w\|_2 = 1$. Intuitively, by projecting each column of Y onto the subspace generated by w, the components of $w^\top Y$ (hereafter referred as "projection scores") would quantify the relative position of the columns of Y, so that their order statistics can be used

to recover the original orders of the columns of Θ. To fix ideas, we define the following ranking operator.

Definition 1 (Ranking Operator). The ranking operator $\mathfrak{r} : \mathbb{R}^p \to \mathcal{S}_p$ is defined such that for any vector $\mathbf{x} \in \mathbb{R}^p$, $\mathfrak{r}(\mathbf{x})$ is the vector of ranks for components of $\mathbf{x}$ in increasing order. Whenever there are ties, increasing orders are assigned from left to right.

For example, given a vector $\mathbf{x} = (2, 5, 1, 6, 2)^\top$, we have $\mathfrak{r}(\mathbf{x}) = (2, 4, 1, 5, 3)$. The following proposition concerning the invariance property of the column spacing of Θ is the key to our construction of the minimax optimal estimator.

Proposition 1. Suppose $(\Theta, \pi) \in \mathcal{D}$. For any nonnegative unit vector $w \in \mathbb{R}^n$, we have

$$\mathfrak{r}(w^\top \Theta \Pi) = \pi^{-1}. \tag{2.1}$$

Apparently, under the noiseless setting, any nonnegative unit vector would lead to the exact recovery of the underlying permutation as in this case the relative orders of the columns are exactly coded by the relative magnitudes of the projection scores $w^\top Y = w^\top \Theta \Pi$. However, with the noisy observations, $w^\top Y = w^\top \Theta \Pi + w^\top Z$ so that the relative orders of the columns are only partially preserved by the noisy projection scores $w^\top Y$, up to some random perturbations.

Consequently, the best linear projection vector w_0 would correspond to the case where $w_0^\top \Theta \Pi$ has the most separated components such that their relative orders are most immune to the random noises. Specifically, since for any given $w \in \mathbb{R}^n$, the i-th component of $w^\top \Theta \Pi$ has the expression $w^\top \Theta \Pi e_i$ where $\{e_i\}_{i=1}^p$ is the canonical basis of the Euclidean space $\mathbb{R}^p$, we define

$$w_0 = \underset{\substack{w \in \mathbb{R}^n \\ \|w\|_2 = 1}}{\arg\max} \sum_{\substack{1 \le i,j \le p \\ i \ne j}} (w^\top \Theta \Pi e_i - w^\top \Theta \Pi e_j)^2 = \underset{\substack{w \in \mathbb{R}^n \\ \|w\|_2 = 1}}{\arg\max} \sum_{i=1}^p \left(w^\top \Theta \Pi e_i - \frac{1}{p} \sum_{j=1}^p w^\top \Theta \Pi e_j \right)^2,$$

which maximizes the pairwise distances of the components under the squared distance. Now since w_0 relies on the unknown $\Theta \Pi$ and is not computable from the data, we substitute $\Theta \Pi$ by its sample/noisy counterpart Y and define our data-driven best linear projection vector as

$$\hat{w} = \underset{\substack{w \in \mathbb{R}^n \\ \|w\|_2 = 1}}{\arg\max} \sum_{i=1}^p \left(w^\top Y e_i - \frac{1}{p} \sum_{i=1}^p w^\top Y e_i \right)^2, \tag{2.2}$$

which is actually the first eigenvector of the symmetric matrix

$$A = Y \sum_{i=1}^p \left(e_i - \frac{1}{p} \sum_{i=1}^p e_i \right) \left(e_i - \frac{1}{p} \sum_{i=1}^p e_i \right)^\top Y^\top, \tag{2.3}$$

and can be immediately solved by performing an eigen-decomposition on A. Once $\hat{w}$ is obtained,

we define our proposed permutation estimator as

$$\hat{\pi} = (\mathfrak{r}(\hat{w}^\top Y))^{-1}. \qquad (2.4)$$

Intuitively, the projection vector $\hat{w}$ assigns different weights to the rows of Y so that more weight is given to the rows whose elements are better separated and therefore more informative in distinguishing the columns of Y or Θ.

2.2 Evaluation Criteria

The main focus of this paper is to investigate the theoretical properties of our proposed estimator (2.4) under various loss measures and parameter spaces. For any given estimator $\check{\pi}$, we first consider the 0-1 loss

$$\ell(\check{\pi}, \pi) = 1\{\check{\pi} \neq \pi\},$$

with the corresponding risk $\mathbb{E}\ell(\check{\pi}, \pi) = P(\check{\pi} \neq \pi)$. The 0-1 loss is used to evaluate the exact recovery, which can be a strong requirement for practical applications. As an alternative, we also consider the more flexible partial recovery, where the loss function is given by the normalized Kendall's tau distance (Kendall, 1938) defined as

$$\tau_K(\pi_1, \pi_2) = \frac{\{\# \text{ of discordant pairs between } \pi_1 \text{ and } \pi_2\}}{\binom{n}{2}}. \qquad (2.5)$$

Technically, for two permutations π_1 and π_2, the set of discordant pairs is defined as

$$\mathcal{G}(\pi_1, \pi_2) = \{(i,j) : i < j, [\pi_1(i) < \pi_1(j) \wedge \pi_2(i) > \pi_2(j)] \vee [\pi_1(i) > \pi_1(j) \wedge \pi_2(i) < \pi_2(j)]\}$$

so that the numerator in (2.5) is equal to the cardinality $|\mathcal{G}(\pi_1, \pi_2)|$, which, in fact, is also the minimum number of pairwise adjacent transpositions converting π_1^{-1} into π_2^{-1} (Diaconis, 1988). The denominator $\binom{n}{2}$ ensures that $\tau_K(\pi_1, \pi_2) \in [0, 1]$ where $\tau_K(\pi_1, \pi_2) = 0$ corresponds to $\pi_1 = \pi_2$.

3 A LINEAR GROWTH MODEL

We start with a simpler case where the pair (Θ, π) is from the subspace

$$\mathcal{D}_L = \left\{ (\Theta, \pi) \in \mathcal{D} : \begin{array}{l} \theta_{ij} = a_i \eta_j + b_i, \text{ where } a_i, b_i \geq 0 \text{ for } 1 \leq i \leq n, \\ 0 \leq \eta_j \leq \eta_{j+1} \text{ for } 1 \leq j \leq p - 1 \end{array} \right\}. \qquad (3.1)$$

In other words, each row of Θ has a linear growth pattern with possibly different intercepts and different slopes. In the context of bacterial growth dynamics, this model is sometimes referred as the Cooper-Helmstetter model (Cooper and Helmstetter, 1968; Bremer and Churchward, 1977)

that associates the copy number of genes with their relative distances to the replication origin. Specifically, a_i is the ratio of genome replication time and doubling time, which can be used to quantify the bacterial growth dynamics for the ith sample, η_j is related to distance from the replication origin for the jth contig, and b_i is related to the read counts at the replication origin and the sequencing depth. If the bacterium is non-dividing in sample i, a_i is zero.

For the linear growth model (3.1), there are two key quantities that are relevant to permutation recovery.

Definition 2. For any $\Theta \in \mathcal{D}_L$, we define

$$\Gamma = \left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} \cdot \min_{1 \leq i < j \leq p} |\eta_i - \eta_j| \tag{3.2}$$

as the *local minimal signal gap* of Θ, and define

$$\Lambda = \left(\sum_{i=1}^{n} a_i^2 \right) \cdot \frac{1}{p} \sum_{1 \leq i < j \leq p} (\eta_i - \eta_j)^2 = \left(\sum_{i=1}^{n} a_i^2 \right) \cdot \sum_{j=1}^{p} (\eta_j - \bar{\eta})^2 \tag{3.3}$$

as the *global signal strength* of Θ, where $\bar{\eta} = \sum_{j=1}^{p} \eta_j / p$.

Intuitively, both quantities involve the set $\{|\eta_j - \eta_i|\}_{1 \leq i < j \leq p}$ and the ℓ_2 norm of the vector $a = (a_1, ..., a_n)^\top$, which characterize the column spacings and the growth rates (slopes) of Θ, respectively. Throughout this paper, we assume

(**A1**) the additive noise matrix $Z \in \mathbb{R}^{n \times p}$ has i.i.d. entries $z_{ij} \sim N(0, \sigma^2)$.

The Gaussian assumption simplifies our theoretical analysis. But this is not essential because all the theoretical results remain true if Z has independent sub-Gaussian entries with parameters bounded by σ^2. The following theorem provides conditions on Γ and Λ such that exact recovery of π can be obtained by $\hat{\pi}$ in (2.4).

Theorem 1 (Exact Recovery, Linear). *Suppose (A1) hold, $(\Theta, \pi) \in \mathcal{D}_L$ and Θ satisfies*

$$\Gamma > C_0 \sigma \sqrt{\log p}, \quad \Lambda > C_1 \sigma^2 (n \max\{\sigma^2 n / \Gamma^2, 1\} + \sqrt{np \max\{\sigma^2 n / \Gamma^2, 1\}}) \tag{3.4}$$

for some $C_0, C_1 > 0$. Then with probability at least $1 - O(p^{-c})$ for some constant $c > 0$, up to a permutation reversion, we have $\hat{\pi} = \pi$.

Remark 1. Due to non-identifiability between $\hat{w}$ and $-\hat{w}$ defined in (2.2), in Theorem 1, as well as all the other theoretical results concerning $\hat{\pi}$, the statement is up to a possible reversion of $\hat{\pi}$. For example, for permutation $\pi = (2, 4, 1, 5, 3)$, its reversion would be $\text{rev}(\pi) = (4, 2, 5, 1, 3)$. In fact, such indeterminacy can be avoided by noting that $a_i \geq 0$ for all i's, but we will not pursue such a direction in this study as the practical interest only concerns relative orders of the permuted elements.

Since Γ depends on a only through its ℓ_2 norm $\|a\|_2$, the local minimal signal gap (MSG) condition $\Gamma \geq C\sigma\sqrt{\log p}$ allows for the presence of non-informative signals in the sense that some components of a can be 0. In contrast, the condition on Λ (GSS) depends on a trade-off between Γ and $\sigma\sqrt{n}$. One the one hand, when $\Gamma > \sigma\sqrt{n}$, the condition on Λ becomes $\Lambda \geq \sigma^2(C_0 n + C_1\sqrt{np})$, which is independent of Γ, and is minimax optimal for left singular subspace estimation (Cai and Zhang, 2018). On the other hand, when $\Gamma < \sigma\sqrt{n}$, a stronger condition on Λ is posed, as a compensation for small Γ.

In some cases, the GSS condition in (3.4) can be implied by the MSG condition. We summarize our results in the following proposition.

Proposition 2. *Suppose $\Gamma/\sigma > 1/p$ and the MSG condition hold. Then the GSS condition can be implied by either one of the following conditions*

(i) $\Gamma \gtrsim \sigma\sqrt{n}$;

(ii) $\Gamma \lesssim \sigma\sqrt{n}$, and either $(\sigma^4 n^2/\Gamma^4)^{1/3} \lesssim p \lesssim \sigma^2 n^2/\Gamma^2$ or $p \gtrsim \sigma^2 n^2/\Gamma^2 + (\sigma^3 n/\Gamma^3)^{2/5}$.

We next turn to the partial recovery and study the rate of convergence of $\hat{\pi}$ measured by the normalized Kendall's tau distance under the linear growth model. In particular, we will assume an approximate uniform assignment of $\{\eta_j\}_{j=1}^p$ over some subinterval of $[0, \infty)$. In other words, the minimal element and maximal element of the set $\{|\eta_j - \eta_{j+1}|\}_{j=1}^{p-1}$ should have roughly the same magnitude, so that $\Gamma = \|a\|_2 \cdot \min_{1 \leq j \leq p-1} |\eta_j - \eta_{j+1}| \asymp \|a\|_2 \cdot \max_{1 \leq j \leq p-1} |\eta_j - \eta_{j+1}|$. This is equivalent to assuming that the contigs in genome assemblies are approximately uniformly spaced along the circular genome.

Theorem 2 (Partial Recovery, Linear). *Suppose (A1) hold, $(\Theta, \pi) \in \mathcal{D}_L$, and Θ satisfies*

(i) there exist some $C_0 > 0$ such that $\max_{1 \leq j \leq p-1} |\eta_j - \eta_{j+1}| < C_0 \min_{1 \leq j \leq p-1} |\eta_j - \eta_{j+1}|$ for all $p > 0$, and

(ii) $\Lambda > C_1\sigma^2\big(\max\big\{\frac{\sigma^2(n+\log p)^2}{\Gamma^2}, n\big\} + \sqrt{p}\max\big\{\frac{\sigma(n+\log p)}{\Gamma}, \sqrt{n}\big\}\big)$ for some $C_1 > 0$.

Then, up to a permutation reversion,

$$\mathbb{E}[\tau_K(\hat{\pi}, \pi)] \leq 1 \wedge \left(\frac{c_0\sigma}{p\Gamma}\min\left\{1, e^{-\Gamma^2/2\sigma^2}\log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{c_1 e^{-\Gamma^2/2\sigma^2}}{p(\Gamma/\sigma + \sqrt{8/\pi})} + \frac{c_2}{p^{c+2}}\right)$$

for some $c, c_0, c_1, c_2 > 0$.

Remark 2. The risk upper bound derived in the above theorem can be simplified as

$$\mathbb{E}[\tau_K(\hat{\pi}, \pi)] \lesssim \begin{cases} \frac{\sigma}{p\Gamma} \wedge 1 & \text{if } \Gamma/\sigma \to 0 \\ \frac{\sigma}{p\Gamma}e^{-\Gamma^2/2\sigma^2} + 1/p^{c+2} & \text{otherwise} \end{cases}$$

for some $c > 0$. In the case of $\Gamma/\sigma \to \infty$, simple calculation yields $e^{-\Gamma^2/2\sigma^2}\sigma/(p\Gamma) + 1/p^{c+2} \asymp e^{-\Gamma^2/2\sigma^2}\sigma/\Gamma$ when $\Gamma < \sigma\sqrt{2(c+1)\log p}$, whereas $e^{-\Gamma^2/2\sigma^2}\sigma/(p\Gamma) + 1/p^{c+2} \asymp 1/p^{c+2}$ when $\Gamma \geq \sigma\sqrt{2(c+1)\log p}$. As a result, we also have

$$\mathbb{E}[\tau_K(\hat{\pi}, \pi)] \lesssim \begin{cases} 1/p^{c+2} & \text{if } \Gamma/\sigma \geq \sqrt{2(c+1)\log p} \\ \frac{\sigma}{p\Gamma}e^{-\Gamma^2/2\sigma^2} & \text{if } 1 \lesssim \Gamma/\sigma < \sqrt{2(c+1)\log p} \\ \frac{\sigma}{p\Gamma} \wedge 1 & \text{if } \Gamma/\sigma \lesssim 1 \end{cases} . \tag{3.5}$$

See Figure 2 for an illustration.

In general, Theorem 2 shows that, even with a weaker condition on Γ that is below the requirement for the exact recovery, our proposed estimator $\hat{\pi}$ is still able to obtain a partial recovery of π with an exponential rate of convergence if $\Gamma/\sigma \gtrsim 1$ and a polynomial rate of convergence if $1/p < \Gamma/\sigma \lesssim 1$. As for Λ, the requirement is essentially the same as the exact recovery, except for an additional $\log p$ term, which is negligible in the exact recovery scenario.
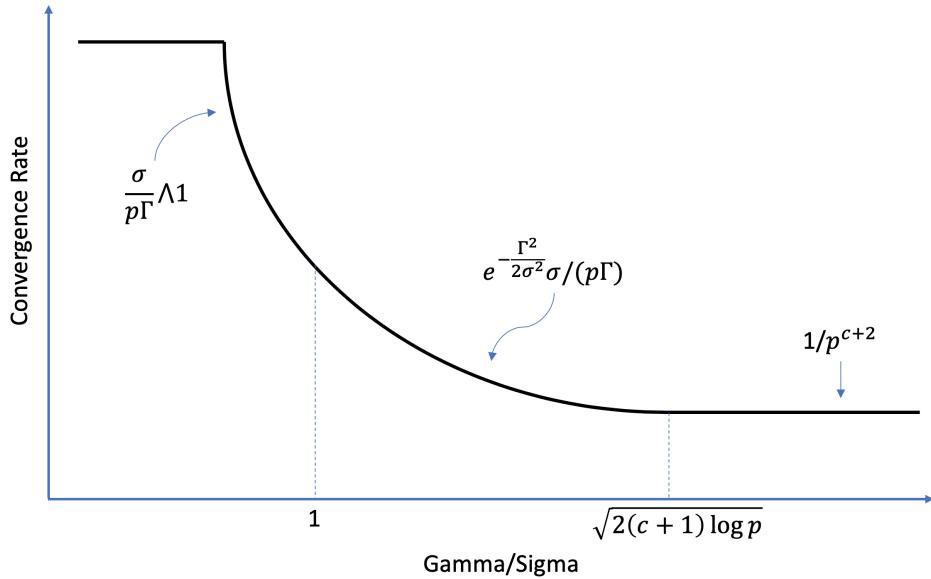

Figure 2: A graphical illustration of the risk upper bound for $\mathbb{E}[\tau_K(\hat{\pi}, \pi)]$, as a function of signal-to-noise ratio Γ/σ.

Some implications about the practically preferable settings of n and p should be clarified. Firstly, although Theorem 1 implies that the difficulty for exact recovery increases as p grows (see also Table 1 from our simulations), our theory suggests a wide range of feasible choices for p. For example, if the underlying signals θ_{ij} and the noise level σ^2 are of constant order, then we have $\Gamma \asymp \sqrt{n}$ and $\Lambda \asymp np^3$, so the conditions of Theorem 1 imply that the exact recovery can be guaranteed as long as $\log p \lesssim n$. In other words, p is allowed to grow exponentially with n, which is in line with the modern high-dimensional setting. Secondly, our Theorem 2 implies that,

even if some conditions (such as MSG) for the exact recovery are not satisfied, one can still hope to partially recover the underlying permutation. In accordance to our theoretical result (3.5), our numerical results (Figure 4) show that, for the partial recovery, increasing p indeed reduces the overall risk of the proposed estimator. Finally, as to the sample size n, we argue that, without assuming additional structural assumptions such as row-sparsity, it is very unlikely that including more samples will result in a worse estimate (see Table 1 and Figure 4 for numerical evidences).

4   A GENERAL GROWTH MODEL

In this section we study the permutation recovery over the general parameter space $\mathcal{D}$ where the growth pattern is not necessarily linear and therefore is more realistic inasmuch as the noisy nature of the shotgun metagenomic datasets (Boulund et al., 2018; Gao and Li, 2018). The analysis relies on a deeper understanding of the relationship between the row-monotonic matrices and its leading singular vectors.

Specifically, for any $\Theta \in \mathcal{D}$, we define the row-centered matrix

$$\Theta' = \Theta(I - p^{-1}ee^\top) \in \mathbb{R}^{n \times p} \tag{4.1}$$

whose singular value decomposition (SVD) is given by $\Theta' = \sum_{i=1}^r \lambda_i(\Theta')u_i'v_i'^\top$, with $r \leq \min\{n, p\}$. The following proposition is essential to our analysis of the general growth model.

Proposition 3. *Let Θ' be defined as above, then its first right singular vector v_1' is a monotone vector, i.e., either $v_{11}' \leq v_{12}' \leq ... \leq v_{1p}'$ or $v_{11}' \geq v_{12}' \geq ... \geq v_{1p}'$.*

Together with Proposition 1, the above proposition justifies our construction of the permutation estimator $\hat{\pi}$ using a PCA based approach. To overcome the identifiability issue, we further assume $\lambda_1(\Theta')$ has multiplicity one. We first introduce the several quantities that play the key roles in permutation recovery over $\mathcal{D}$.

Definition 3. For any $\Theta \in \mathcal{D}$ and the corresponding Θ' defined as above, we define

$$\Gamma = \min_{1 \leq i < j \leq p} |u_1'^\top(\Theta'_{\cdot i} - \Theta'_{\cdot j})| = \lambda_1(\Theta') \min_{1 \leq i < j \leq p} |v_{1i}' - v_{1j}'|,$$

as the *local minimal signal gap*, define

$$\Xi = \max_{1 \leq i \leq p-1} \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2 = \max_{1 \leq i \leq p-1} \left(\sum_{j=1}^r \lambda_j^2(\Theta')|v_{ji}' - v_{j,i+1}'|^2 \right)^{1/2},$$

as the *local maximal signal gap*, and define

$$\Lambda = \lambda_1^2(\Theta') - \lambda_2^2(\Theta')$$

as the *global signal strength* of Θ.

In particular, the above definitions of Γ and Λ generalize the ones given earlier in the linear growth model as these quantities coincide for $\Theta \in \mathcal{D}_L$. The following theorem concerns the exact permutation recovery with $\hat{\pi}$ over $\mathcal{D}$.

Theorem 3 (Exact Recovery, General). *Suppose (A1) hold, $n \lesssim p$, $(\Theta, \pi) \in \mathcal{D}$, and Θ satisfies $\Gamma > C_0 \sigma \sqrt{\log p}$ and*

$$\Lambda > C_1 \sigma^2 \left[\left(n + \frac{\Xi^2}{\sigma^2} \right) \max\left\{ \frac{(n + \log p)\sigma^2}{\Gamma^2}, 1 \right\} + \sqrt{p} \left(\sqrt{n} + \frac{\Xi}{\sigma} \right) \max\left\{ \frac{\sigma\sqrt{n + \log p}}{\Gamma}, 1 \right\} \right]$$

for some $C_0, C_1 > 0$. Then with probability at least $1 - O(p^{-c})$ for some constant $c > 0$, up to a permutation reversion, we have $\hat{\pi} = \pi$.

As in the case of linear growth model (Theorem 1), in Theorem 3, to guarantee exact recovery, we need the MSG condition $\Gamma > C_0 \sigma \sqrt{\log p}$. Unlike the linear growth model, here Γ only implicitly depends on the elements of Θ through its spectral quantities, which makes its interpretation less clear. To address this issue, we make the following observation that links the minimal singular vector gap $\min_{1 \le i < j \le p} |v'_{1i} - v'_{1j}|$ in the definition of Γ to the elements of Θ.

Proposition 4. *Let Θ' in (4.1) be such that there exists a $\delta > 0$ being the lower bound of the normalized minimum gap between any two entries in the same row, i.e.*

$$\min_{1 \le k \le n} \frac{|\theta'_{k,i} - \theta'_{k,j}|}{\|\Theta'_{k.}\|_2} \ge \delta \quad \text{for some } i \ne j.$$

Then the first singular vector $v'_1 \in \mathbb{R}^p$ of Θ' satisfies $|v'_{1,i} - v'_{1,j}| \ge \delta$.

Consequently, the implicit requirement that $\min_{1 \le i < j \le p} |v'_{1i} - v'_{1j}|$ is large can be guaranteed when the normalized minimum distance $\min_{1 \le i < j \le p} \min_{1 \le k \le n} |\theta'_{k,i} - \theta'_{k,j}|/\|\Theta'_{k.}\|_2$ is large. Our next theorem concerns the partial recovery over the general parameter space $\mathcal{D}$.

Theorem 4 (Partial Recovery, General). *Suppose (A1) hold, $n \lesssim p$, $(\Theta, \pi) \in \mathcal{D}$, and Θ satisfies*

(i) *there exits some $C_0 > 0$ such that $\max_{1 \le j \le p-1} |v'_{1j} - v'_{1,j+1}| < C_0 \min_{1 \le j \le p-1} |v'_{1j} - v'_{1,j+1}|$ for all $p > 0$, and*

(ii) $\Lambda > C_1 \sigma^2 \left[\max\left\{ \frac{(n + \log p)^2 \sigma^2}{\Gamma^2}, n + \frac{\Xi^2}{\sigma^2} \right\} + \sqrt{p} \max\left\{ \frac{\sigma(n + \log p)}{\Gamma}, \sqrt{n} + \frac{\Xi}{\sigma} \right\} \right]$ *for some $C_1 > 0$.*

Then, up to a permutation reversion,

$$\mathbb{E}[\tau_K(\hat{\pi}, \pi)] \le 1 \wedge \left(\frac{c_0 \sigma}{p\Gamma} \min\left\{ 1, e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2} \right) \right\} + \frac{c_1 e^{-\Gamma^2/2\sigma^2}}{p(\Gamma/\sigma + \sqrt{8/\pi})} + \frac{c_2}{p^{c+2}} \right)$$

for some $c, c_0, c_1, c_2 > 0$.

Condition (i) of Theorem 4 parallels the one given in Theorem 2. It essentially requires an even distancing of the elements (the projected columns of Θ) whose ordering is to be tracked by $\hat{\pi}$. In contrast, in both Theorem 3 and 4, the conditions on Λ are slightly more complicated than those in Theorem 1 and 2, as it further depends on the relative magnitude between Ξ/σ and $\sqrt{n}$. In particular, if $\Xi/\sigma \lesssim \sqrt{n}$, the conditions reduce to the ones required in the linear growth models. Interestingly, the risk upper bound obtained in Theorem 4 remains the same as in the linear growth model, which only depends on p and the signal-to-noise ratio Γ/σ.

5 MINIMAX LOWER BOUNDS AND OPTIMALITY

In this section, we establish the minimax lower bounds for both exact and partial recovery considered in previous sections, in relation to different levels of the signal-to-noise ratio Γ/σ. In the following theorem, we show the MSG condition for exact recovery is asymptotically sharp.

Theorem 5. *Suppose (A1) hold. Let $\mathcal{D}_1 = \mathcal{D}_L \cap \{(\Theta, \pi) : \Gamma \leq \frac{\sigma}{4}\sqrt{\log p}\}$ and $\mathcal{D}_1' = \mathcal{D} \cap \{(\Theta, \pi) : \Gamma \leq \frac{\sigma}{4}\sqrt{\log p}\}$. Then for any $p \geq 10$, we have*

$$\inf_{\hat{\pi}} \sup_{(\Theta, \pi) \in \mathcal{D}_1'} P(\hat{\pi} \neq \pi) \geq \inf_{\hat{\pi}} \sup_{(\Theta, \pi) \in \mathcal{D}_1} P(\hat{\pi} \neq \pi) \geq 0.3,$$

where the infimum is over all the permutation estimators $\hat{\pi}$.

This theorem along with Theorem 1 and Theorem 3 indicates that our proposed estimator is minimax rate-optimal over $\mathcal{D}_L$ and $\mathcal{D}$ in terms of the MSG condition on Γ. In light of Proposition 2, in some situations the MSG condition can be both necessary and sufficient for the exact recovery, which includes practically important cases such as $n \asymp p$, $n < \log p$, etc. Using the information-theoretic language, we have therefore obtained both the achievability result, i.e., the existence of an algorithm or estimator that exactly recovers signal with high probability, and the converse result, namely, an upper bound on the probability of exact recovery that applies to any estimators (Cullina and Kiyavash, 2016). See Figure 3 for an illustration.
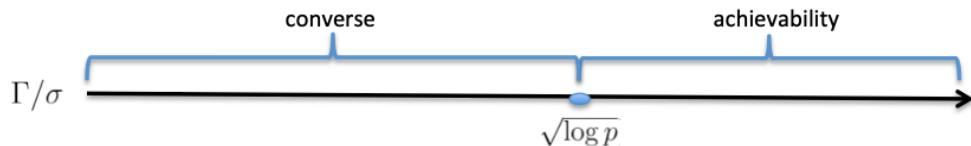

Figure 3: A graphical illustration of the achievability/converse result for exact recovery.

Our next theorem establishes a minimax lower bound for the expected rate of convergence for the partial recovery.

Theorem 6. *Suppose (A1) hold, $\mathcal{D}_2(t) = \mathcal{D}_L \cap \{(\Theta, \pi) : \Gamma \asymp t\}$ and $\mathcal{D}'_2(t) = \mathcal{D} \cap \{(\Theta, \pi) : \Gamma \asymp t\}$ and $t/\sigma \geq 2$. Then there exist constants $C_1, C_2 > 0$ such that*

$$\inf_{\hat{\pi}} \sup_{(\Theta, \pi) \in \mathcal{D}'_2(t)} \mathbb{E}[\tau_K(\hat{\pi}, \pi)] \geq \inf_{\hat{\pi}} \sup_{(\Theta, \pi) \in \mathcal{D}_2(t)} \mathbb{E}[\tau_K(\hat{\pi}, \pi)] \geq \frac{C_1 \sigma}{pt} e^{-t^2/2\sigma^2} + \frac{C_2}{p^2}.$$

Comparing the above minimax lower bound to the risk upper bounds obtained in Theorem 2 and 4, we conclude that our proposed estimator $\hat{\pi}$ is minimax rate-optimal in terms of the partial recovery for both the linear growth model and the general growth model over the range whenever Γ/σ does not diminish (Figure 1). In particular, in Theorem 5 and 6, since the minimax lower bounds only concern the worst-case scenarios, the same lower bounds should hold for any parameter spaces whenever the same worst cases are included. Similarly, the assumption (A1) does not pose a restriction to the general applicability of such results.

6 NUMERICAL STUDIES

6.1 Simulation with Model-Generated Data

To demonstrate our theoretical results and compare with alternative methods, we generate data from model (1.1) with various configurations of the signal matrix Θ. We compare the empirical performance of our proposed estimator $\hat{\pi}$ with the following alternatives:

- π_{mean} : Order the columns of Y by the magnitude of its column means;

- π_{max}: Order the columns of Y by the magnitude of its column maximums.

We use both the 0-1 loss and the normalized Kendall's tau distance in comparing these methods. Due to the identifiability issue, the performance of each estimator is evaluated up to a complete reversion of the permutation. For example, we use $\min\{\tau_K(\hat{\pi}, \pi), \tau_K(\hat{\pi}, \text{rev}(\pi))\}$ as the empirical Kendall's tau distance. By symmetry, we set the underlying permutation $\pi = id$. The signal matrix $\Theta = (\theta_{ij}) \in \mathbb{R}^{n \times p}$ is generated under the following four regimes:

(i) $S_1(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = \log(1 + j\alpha_i + \beta_i)$ where $\alpha_i \sim \text{Unif}(\alpha/2, \alpha)$ for $1 \leq i \leq n/2$, $\alpha_i \sim \text{Unif}(0, 0.01)$ for $n/2 < i \leq n$, and $\beta_i \sim \text{Unif}(1, 3)$ for all $1 \leq i \leq n$;

(ii) $S_2(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = j\alpha_i + \beta_i$ where $\alpha_i \sim \text{Unif}(\alpha/2, \alpha)$ for $1 \leq i \leq n/2$, $\alpha_i \sim \text{Unif}(0, \alpha/10)$ for $n/2 < i \leq n$, and $\beta_i \sim \text{Unif}(1, 3)$ for all $1 \leq i \leq n$;

(iii) $S_3(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = \log(1 + j\alpha_i + \beta_i)$ where $\alpha_i \sim \text{Unif}(\alpha/2, \alpha)$ for $1 \leq i \leq 3$, $\alpha_i \sim \text{Unif}(0, 0.01)$ for $4 < i \leq n$, and $\beta_i \sim \text{Unif}(1, 3)$ for all $1 \leq i \leq n$;

(iv) $S_4(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = j\alpha_i + \beta_i$ where $\alpha_i \sim \text{Unif}(\alpha/2, \alpha)$ for $1 \leq i \leq 3$, $\alpha_i \sim \text{Unif}(0, \alpha/10)$ for $4 < i \leq n$, and $\beta_i \sim \text{Unif}(1, 3)$ for all $1 \leq i \leq n$.

14

Specifically, under each regime, the sample-specific "growth rate" parameter α_i is randomly and uniformly generated either from the interval $[\alpha/2, \alpha]$ or an interval with much smaller values, namely, $[0, \alpha/10]$ in S_2 and S_4 and $[0, 0.01]$ in S_1 and S_3. By construction, the four regimes consist of the nonlinear growth model where the signals spread out over many samples (S_1) or concentrate at a few rows (S_3) and the linear growth model where the signals spread out over many samples (S_2) or concentrate at a few rows (S_4). In particular, in accordance to our theory, for the supposedly "non-informative" samples, we allow the corresponding growth rates to be small but non-zero, which shows the flexibility of our proposed method. The entries of Z are drawn from i.i.d. centred normal distributions whose variance σ^2 will be given explicitly. In each setting, we evaluate the empirical performance of each method over a range of n, p or α. Each setting is repeated for 200 times.

For the exact recovery, in Table 1, we reported the empirical risks of the estimators under the 0-1 loss for various regimes and parameter combinations. The noise level σ^2 is chosen for each regime to better illustrate the differences in the empirical risks among the estimators. From our simulation results, in consistent to our theory, our proposed estimator has the smallest empirical risk over all the settings, and the estimation risk decreases as we increase α, n or decrease p.

Table 1: The empirical risks of the estimators under the 0-1 loss based on 200 simulations for various combinations of the parameters (p, n, α). $\hat{\pi}$: proposed method; π_{mean}: mean-based method; π_{max}: max-based method.

$p = 75$	$S_1(\sigma^2 = 0.025)$		$S_2(\sigma^2 = 0.1)$		$S_3(\sigma^2 = 0.0075)$		$S_4(\sigma^2 = 0.025)$	
$n = 40$	$\alpha = 0.1$	0.2	0.1	0.2	0.1	0.2	0.1	0.2
$\hat{\pi}$	0.775	0.575	0.415	0.000	0.025	0.020	0.025	0.000
π_{mean}	0.925	0.815	0.955	0.015	0.155	0.135	0.880	0.005
π_{max}	1.000	1.000	1.000	0.995	0.995	0.970	0.840	0.430
$n = 40$	$S_1(\sigma^2 = 0.025)$		$S_2(\sigma^2 = 0.1)$		$S_3(\sigma^2 = 0.0075)$		$S_4(\sigma^2 = 0.025)$	
$\alpha = 0.1$	$p = 60$	90	60	90	60	90	60	90
$\hat{\pi}$	0.410	0.930	0.340	0.470	0.010	0.115	0.000	0.010
π_{mean}	0.720	0.985	0.910	0.980	0.070	0.245	0.775	0.900
π_{max}	1.000	1.000	1.000	1.000	0.975	1.000	0.815	0.875
$p = 75$	$S_1(\sigma^2 = 0.025)$		$S_2(\sigma^2 = 0.1)$		$S_3(\sigma^2 = 0.0075)$		$S_4(\sigma^2 = 0.025)$	
$\alpha = 0.1$	$n = 40$	60	40	60	40	60	40	60
$\hat{\pi}$	0.765	0.440	0.475	0.095	0.050	0.020	0.010	0.005
π_{mean}	0.920	0.645	0.940	0.700	0.175	0.045	0.900	0.905
π_{max}	1.000	1.000	1.000	1.000	0.995	0.995	0.855	0.820

For partial recovery, in Figure 4, we show boxplots of the empirical normalized Kendall's tau between each estimator and the true permutation π. Again, our proposed method outperforms the alternatives in all the cases. As expected from our theory, under all the four regimes, increasing p while keeping other parameters fixed results to smaller estimation risk. As for the dependence on

n, under S_1 and S_2, increasing n leads to smaller risk as it is equivalent to increasing Γ, whereas under S_3 and S_4, the risk roughly remains the same across different n's as in these case Γ doesn't change much.

To offer more intuitive interpretation of why $\hat{\pi}$ performs better than the alternative methods, we assessed the weight vectors $\hat{w}$ of our proposed estimator $\hat{\pi}$ under each regime after 200 rounds of simulations (Figure 3 in Supplemented Material). In comparison, the weight vector for π_{mean} is simply $(1/\sqrt{n}, ..., 1/\sqrt{n})$, which assigns equal weight to all the samples. On the other hand, since $\pi_{\max}$ cannot be written in the form of $(\mathfrak{r}(w^\top Y))^{-1}$ for some weight vector w and therefore does not belong to the class of linear projection estimators, we reported instead the pseudo-weight vector $\tilde{w} \in \mathbb{R}^n$ where the i-th component is the proportion that the i-th sample is used among the p coordinates. In general, we found that $\tilde{w} \in \mathbb{R}^n$ assigns larger weights to only a few samples among those with higher signal strength, and the weight vector for π_{mean} fails to distinguish the informative samples from the non-informative ones. In contrast, the weight vectors $\hat{w}$ for our proposed estimator $\hat{\pi}$ would automatically adapt to the varying signal strengths across the samples and assign larger weights to the samples with more significant signal changes. This also explains the interesting phenomenon in Figure 4 that, under the regime S_1 and S_2, $\hat{\pi}$ and $\pi_{\mean}$ perform better than $\pi_{\max}$, whereas under S_3 and S_4, $\hat{\pi}$ and $\pi_{\max}$ perform better. In summary, methods that are able to detect and assign larger weight to the more informative samples would perform better than methods that are not. Observably, $\hat{\pi}$ combines the advantages of $\pi_{\mean}$ and $\pi_{\max}$ in that it finds the best weights (projection scores) in a data-driven manner.

6.2 Evaluation Using Synthetic Metagenomic Data

We evaluate the empirical performance of our proposed method using a synthetic metagenomic sequencing dataset used in Gao and Li (2018) by generating sequencing reads based on 45 bacterial genomes. Instead of estimating the PTRs, which was the focus of Gao and Li (2018), our goal is to recover the unknown relative orders of the contigs assembled in typical metagenomics studies. In addition to assisting the estimation of PTRs, such ordering of the contigs could be of independent interest for other applications, including genome assemblies based on shotgun metagenomics data.

Gao and Li (2018) presented a synthetic shotgun metagenomic sequencing dataset of a community of 45 phylogenetically related species from 15 genera of five different phyla with known RefSeq ID, taxonomy and replication origin (Gao et al., 2013) (see Figure 2 in our Supplementary Material). To generate metagenomics reads, reference genome sequences of randomly selected three species in each genus were downloaded from NCBI. Read coverages were generated along the genome based on an exponential distribution with a specified peak-to-trough ratio and a function of accumulative distribution of read coverages along the genome was calculated. Sequencing reads were next generated using the above accumulative distribution function and a random location of each read on the genome, until the total read number achieved a randomly assigned average cover-
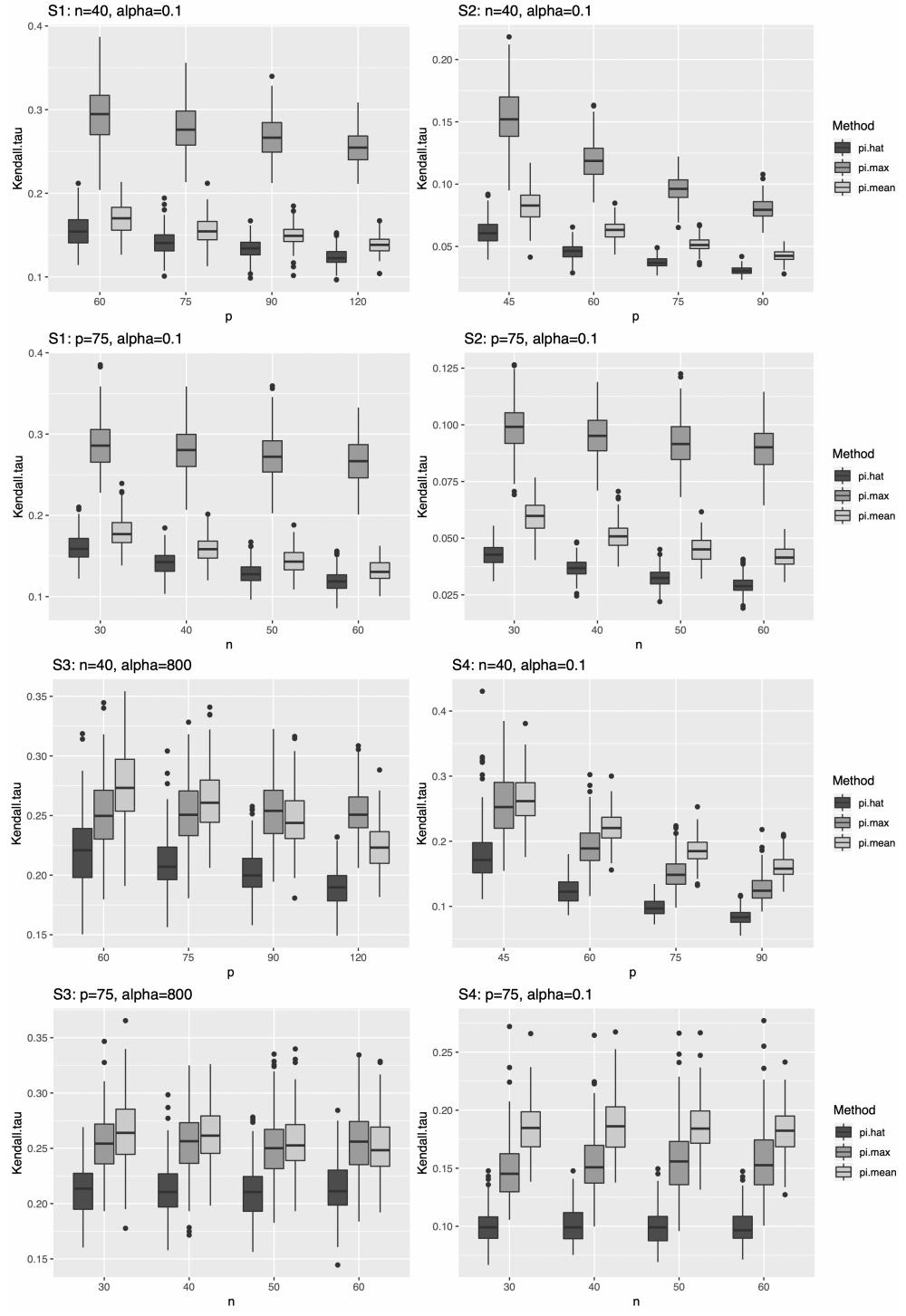
Figure 4: Boxplots of the empirical normalized Kendall's distance between the estimated and true permutations under different models. $\hat{\pi}$: proposed estimator; π_{mean}: mean-based estimator; π_{max}: max-based estimatior.

age between 0.5 and 10 folds for the species in a sample. Sequencing errors including substitution, insertion and deletion were simulated in a position- and nucleotide-specific pattern according to a recent study on metagenomic sequencing error profiles of Illumina.

For the final dataset, the average nucleotide identities (ANI) between species within each genus ranged from 66.6% to 91.2% The probability of one species existing in each of the 50 simulated samples was set as 0.6, and a total of 1,336 average coverages and the corresponding PTRs were randomly and independently assigned. After the same processing and filtering steps and CG-adjustment step as in Gao and Li (2018), the final dataset included genome assemblies of 41 species. For each species, we obtained the permuted matrix of log-contig counts, with the number of samples ranging from 29 to 46, and the number of contigs ranging from 47 to 482.

Our proposed method ($\hat{\pi}$) was used to estimate the unknown orders of the contigs for each species and each sample. As a comparison, we also considered the iRep estimator proposed in Brown et al. (2016), where the contigs of a given species were ordered for each sample separately based on the read counts observed. We evaluate these methods by comparing the estimated contig orders to their true orders as measured by the normalized Kendall's tau distance. To generalize our evaluation to diverse metagenomic datasets, we also evaluate the effect of sample size as well as contig numbers by randomly selecting subsets of samples or contigs from each dataset. The selection was made with replacement.

The results are summarized in Figure 5 by comparing the normalized Kendall's tau distances. As n or p varies, our proposed estimator performs consistently better than iRep in recovering the true contig orders, which explains partially why the DEMIC algorithm worked better in estimating the bacterial growth dynamics. The results of our methods are not sensitive to the sample size and the number of contigs from the genome assemblies. Our estimator also shows smaller variability.
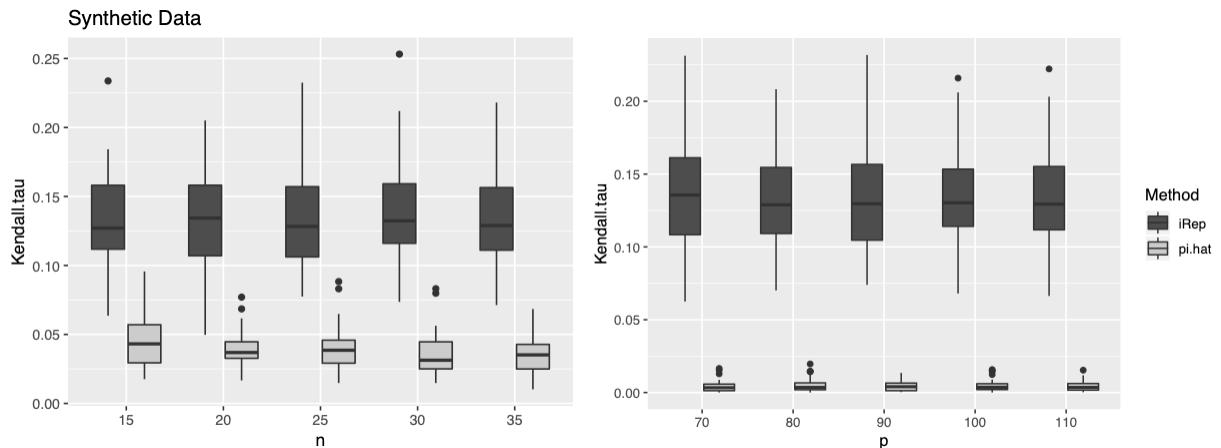

Figure 5: Boxplots of the normalized Kendall's distance between the estimated contig orders and the true orders for different sample sizes n and different numbers of contigs p. The lighter ones correspond to our proposed method and the darker ones correspond to the iRep estimation method.

6.3 Analysis of a Real Microbiome Metagenomic Data Set

Finally, we complete our numerical studies by analyzing a real metagenomic dataset from the Pediatric Longitudinal Study of Elemental Diet and Stool Microbiome Composition (PLEASE) study, a prospective cohort study to investigate the treatment effects on the gut microbiome and reduction of inflammation in pediatric Crohn's disease patients (Lewis et al., 2015). In particular, sequencing data from the fecal samples of 86 Crohn's disease children were obtained at baseline, 1 week and 8 weeks after antiTNF or enteral diet treatment. In our analysis, the sequencing data at the 8th week after treatment was used to compare the bacterial growth dynamics for non-responders ($n = 34$) and responders ($n = 47$). The reads were downloaded from NCBI short read archive (SRP057027) with the corresponding metadata. After the same coassembly, alignment and binning steps as in Gao and Li (2018), the DEMIC algorithm was applied to estimate the bacterial growth rate of a given species represented by a contig cluster (bin) for each sample. In particular, DEMIC applied our proposed method to the GC-adjusted contig coverage data to recover the original order of the contigs. After obtaining the ordered contigs, a simple linear regression was fitted to obtain estimates of the PTRs (ePTRs).

In order to compare the baterial growth rates between responders and non-responders, our analysis focused on ePTRs of 8 contig clusters over subsets of the non-responders (n_1) and the responders (n_2) after 8 weeks of treatment with $\min\{n_1, n_2\} > 5$. Other contig clusters were rare and only appeared in a few samples. For each contig cluster, we compare the ePTRs of the responders and non-responders Wilcoxon rank sum test (Table S.1 in Supplementary Material). The taxonomic annotations of these eight contig clusters were obtain by applying the BAT algorithm (von Meijenfeldt et al., 2019) that compares the metagenomic assembled bins to a taxonomy database. In Table S.1, we show the final taxonomic annotations for each bin to the finest possible resolution, with the lineage scores indicating the quality of each taxonomic classification.

Among the 8 contig clusters, bin.026 showed a significant difference in ePTRs between responders and non-responders after either antiTNF or enteral diet treatment for 8 weeks (p=0.0418), where the growth rate was higher in Crohn's disease patients who did not respond to the treatment. The taxonomic classification (Table S.1) shows that this contig cluster belongs to the phylum *Firmicutes* and the order *Clostridiales*. Since BAT algorithm was not able to classify the order *Clostridiales* to finer taxonomic level of known species, this contig cluster may represent a new species that is important to the treatment outcome of Crohn's disease patients.

7 DISCUSSION

In this paper, partial recovery was studied under the normalized Kendall's tau distance. Another commonly used metric is the normalized Spearman's footrule distance defined by

$$\rho(\pi_1, \pi_2) = \frac{2}{p(p-1)} \sum_{i=1}^{p} |\pi_1(i) - \pi_2(i)|, \qquad \pi_1, \pi_2 \in \mathcal{S}_p.$$

A celebrated result by Diaconis and Graham (1977) shows that $\tau_K(\pi_1, \pi_2) \leq \rho(\pi_1, \pi_2) \leq 2\tau_K(\pi_1, \pi_2)$, which means the two distances are equivalent. As a consequence, all the theoretical results presented in this paper concerning the Kendall's tau distance also hold for the Spearman's footrule distance without any change.

The minimax optimality of the proposed estimator $\hat{\pi}$ was investigated in Section 5 by examining the asymptotic sharpness of the MSG condition for exact recovery, and by obtaining the matching minimax risk lower bound for partial recovery. There are a few issues that deserve further investigation. For both exact and partial recovery, it is unclear to what extent the GSS condition is necessary. In our risk analysis, the perturbation bound for the left singular subspace (Cai and Zhang, 2018) was used. In fact, similar results can be obtained using the concentration bound for the linear functionals of singular vectors (Koltchinskii and Xia, 2016). Nevertheless, it remains to show whether the GSS condition is also asymptotically sharp. In addition, in Theorem 6, the matching minimax lower bound was obtained only for nonvanishing Γ/σ. It remains to show whether the rate $\sigma/(p\Gamma)$ is minimax optimal when $\Gamma/\sigma \to 0$. The difficulty lies in finding a $p^{1+\delta}$-sphere packing of the group $\mathcal{S}_p$ equipped with the Kendall's tau distance for any $0 < \delta < 1$, while the pairwise ℓ_2 distances of the packing elements are also well controlled. Some initial steps have been made in the so-called rank modulation theory (Barg and Mazumdar, 2010; Mazumdar et al., 2013).

There are several related problems that are also of significant theoretical and practical interest. Firstly, although we used the Kendall's tau distance or the equivalent Spearman's footrule distance as the metric for partial recovery, other distances such the Hamming distance, Spearman's rank correlation distance, and Ulam's distance have also been used as the performance metrics for partial recovery in other permutation estimation problems (Göloğlu et al., 2015; Mukherjee, 2016). It is therefore of interest to see how $\hat{\pi}$ performs under these losses. Secondly, our proposed estimator $\hat{\pi}$ implicitly performs a (linear) dimension reduction technique and only uses the information contained in the first eigenvector of A in (2.3). A natural extension is to consider the eigen-subspace spanned by the first k eigenvectors and to estimate the permutation in a sequential manner.

The present paper focuses on the estimation of the permutation matrix Π. It is also of interest to estimate the underlying signal matrix Θ or some functionals of it. For example, in microbiome growth dynamics studies, it is of significant interest to estimate the peak-to-trough ratio $\exp(\theta_{kp} -$

θ_{k1}) for $k = 1, ..., n$, which measures the microbial growth rate for the kth sample, and to identify the samples with peak-to-trough ratio of 1. It is also interesting to identify the bacteria that show differential growth dynamics between disease and normal individuals. Finally, robust permutation recovery methods that can relax the Gaussian or sub-Gaussian assumption of the noise in the permuted monotone matrix model are needed. For example, in some applications, the columns of the noise matrix are not independent, or the variance levels across the noise matrix are not identical. In these cases, we argue that, as long as the marginal distributions of the noise matrix entries remain sub-Gaussian, the analytical framework of the current paper can still be applied, but with more efforts to control the underlying heteroskedasticity. Toward this end, results from the recent work of Zhang et al. (2018) can be very useful, in terms of the new technical tools that parallel the ones used in the current paper to analyse the homoskedastic PCA (cf. Lemma 2 and 3). Finally, to account for non-informative samples, sparse PCA (Cai et al., 2013; Yuan and Zhang, 2013) can be considered. These are interesting problems left for future research.

8 PROOFS OF THE MAIN THEOREMS

In this section, we prove Theorems 1 and 2 in detail and briefly sketch the proofs of Theorems 3 and 4. We also prove the minimax lower bounds in Theorems 5 and 6. Proofs of other results including the technical lemmas can be found in the online Supplementary Materials.

Proof of Theorem 1. Let $X = \Theta + Z$. It follows that $Y = X\Pi$. By right invariance of the 0-1 loss with respect to permutation composition, we have

$$\ell((\mathfrak{r}(\hat{w}^\top Y))^{-1}, \pi) = \ell((\mathfrak{r}(\hat{w}^\top X\Pi))^{-1}, \pi) = \ell((\mathfrak{r}(\hat{w}^\top X))^{-1} \circ \pi, \pi) = \ell((\mathfrak{r}(\hat{w}^\top X))^{-1}, id).$$

Thus it suffices to study the risk $\mathbb{E}\ell((\mathfrak{r}(\hat{w}^\top X))^{-1}, id) = P((\mathfrak{r}(\hat{w}^\top X))^{-1} \neq id)$. In fact,

$$P((\mathfrak{r}(\hat{w}^\top X))^{-1} \neq id) \leq P\left(\bigcup_{i=1}^{p-1}\left\{\sum_{k=1}^n \hat{w}_k X_{ki} \geq \sum_{k=1}^n \hat{w}_k X_{k,i+1}\right\}\right)$$

$$\leq \sum_{i=1}^{p-1} P\left(\sum_{k=1}^n \hat{w}_k X_{ki} \geq \sum_{k=1}^n \hat{w}_k X_{k,i+1}\right) = \sum_{i=1}^{p-1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right), \quad (8.1)$$

which further reduces to obtaining an upper bound for $P_i = P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right)$. By definition, $\hat{w}$ is the first eigenvector of $A = Y\left(I - \frac{1}{p}ee^\top\right)\left(I - \frac{1}{p}ee^\top\right)Y^\top$. Simple calculation yields $\Pi\left(I - \frac{1}{p}ee^\top\right)\left(I - \frac{1}{p}ee^\top\right)\Pi^\top = \left(I - \frac{1}{p}ee^\top\right)\left(I - \frac{1}{p}ee^\top\right)$ for any $\Pi \in \mathcal{S}_p$. So $\hat{w}$ is also the first eigenvector of $A = X\left(I - \frac{1}{p}ee^\top\right)\left(I - \frac{1}{p}ee^\top\right)X^\top \equiv TT^\top$, where $T \in \mathbb{R}^{n \times p}$. Note that T admits the decomposition $T = \Theta' + E \in \mathbb{R}^{n \times p}$ where $E_{ij} \sim N(0, (p-1)\sigma^2/p)$ and $\Theta' = a\eta'^\top, \eta_j' = \eta_j - \frac{1}{p}\sum_{i=1}^p \eta_i$. In particular, $T_{\cdot i} = X_{\cdot i} - \bar{X}_{row}$ where $\bar{X}_{row} = p^{-1}\sum_{i=1}^p X_{\cdot i} \in \mathbb{R}^n$ is the vector of row means of X. We denote

$\phi_{ij} = T_{\cdot i} - T_{\cdot j} = X_{\cdot i} - X_{\cdot j}$ and denote $w = a/\|a\|_2 \in \mathbb{R}^n$ as the first eigenvector of the rank-one matrix $\Theta'\Theta'^\top$. Now following (8.1), we have

$$P_i = P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) = P\left(w^\top\phi_{ij} + (\hat{w} - w)^\top\phi_{ij} \geq 0\right)$$

$$= P\left(w^\top\phi_{i,i+1} + (\hat{w} - w)^\top\phi_{ij} \geq 0, |1 - (\hat{w}^\top w)^2| \leq \delta\right)$$

$$+ P\left(w^\top\phi_{i,i+1} + (\hat{w} - w)^\top\phi_{ij} \geq 0, |1 - (\hat{w}^\top w)^2| > \delta\right)$$

for some $\delta > 0$. By definition, up to a change of sign for $\hat{w}$, we have $0 \leq \hat{w}^\top w \leq 1$. Then $|1 - (\hat{w}^\top w)^2| \leq \delta$ implies $|(\hat{w} - w)^\top\phi_{i,i+1}| \leq \|\hat{w} - w\|_2\|\phi_{i,i+1}\|_2 \leq \sqrt{2\delta}\|\phi_{i,i+1}\|_2$, where the first inequality follows from Cauchy-Schwartz and the second inequality used $\|\hat{w} - w\|_2 = \sqrt{2(1 - \hat{w}^\top w)} \leq \sqrt{2(1 - (\hat{w}^\top w)^2)}$. Thus

$$P_i \leq P\left(w^\top\phi_{i,i+1} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|_2\right) + P\left(|1 - (\hat{w}^\top w)^2| > \delta\right). \tag{8.2}$$

The following lemmas provide upper bounds for the two probability events in the last expression.

Lemma 1. *Under the conditions of Theorem 1, denote $\Gamma_i = \|a\|_2(\eta_i - \eta_{i+1})$, then for any $\delta > 0$, we have*

$$P\left(w^\top\phi_{i,i+1} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|\right) \leq \Phi\left(C_1\sqrt{\delta}\Psi_i^{1/2} + \frac{\Gamma_i}{\sigma}\right) + \frac{C_2}{p^c} \tag{8.3}$$

for $\Psi_i = (\sqrt{n} + \sqrt{\log p})^2 + \frac{\Gamma_i^2}{\sigma^2} + \frac{|\Gamma_i|}{\sigma}\sqrt{\log p}$ and some constants $C_1, C_2, c > 0$.

Lemma 2. *Suppose $\lambda_1^2(\Theta') \geq C\sigma^2(n + \sqrt{np})$ for some $C > 0$, it follows that*

$$P\left(|1 - (\hat{w}^\top w)^2| \leq C_1\frac{\sigma^2(\lambda_1^2(\Theta') + \sigma^2 p)(n + \log p)}{\lambda_1^4(\Theta')}\right) \geq 1 - \frac{C_2}{p^c}$$

for some $C_1, C_2, c > 0$.

Now since $\frac{1}{p}\sum_{1 \leq i < j \leq p}(\eta_i - \eta_j)^2 = \sum_{j=1}^{p}\left(\eta_j - \frac{1}{p}\sum_{i=1}^{p}\eta_i\right)^2 = \sum_{j=1}^{p}\eta_j'^2$, we have $\lambda_1^2(\Theta') = \Lambda > C_0\sigma^2(n + \sqrt{np})$ for some $C_0 > 0$. Set $\delta = C_0\sigma^2\frac{(\lambda_1^2(\Theta') + \sigma^2 p)(n + \log p)}{\lambda_1^4(\Theta')}$. It follows that $\delta = o(1)$. Combining Lemma 1 and Lemma 2, we have

$$P_i \leq \Phi\left(C\sqrt{\delta}\left[(\sqrt{n} + \sqrt{\log p})^2 + \frac{\Gamma_i^2}{\sigma^2} + \frac{|\Gamma_i|}{\sigma}\sqrt{\log p}\right]^{1/2} + \frac{\Gamma_i}{\sigma}\right) + \frac{C}{p^c} \tag{8.4}$$

for some $C, c > 0$. The rest of the analysis is divided into several cases.

22

Case 1. $\log p \lesssim n$. In this case, we have $P_i \leq \Phi\big(C\sqrt{\delta}\big[n + \frac{\Gamma_i^2}{\sigma^2} + \frac{|\Gamma_i|}{\sigma}\sqrt{\log p}\big]^{1/2} + \frac{\Gamma_i}{\sigma}\big) + \frac{C}{p^c}$. In addition, if $|\Gamma_i|/\sigma \lesssim \sqrt{n}$, we have $P_i \leq \Phi\big(C\sqrt{\delta n} + \frac{\Gamma_i}{\sigma}\big) + \frac{C}{p^c} \leq \frac{C'}{p^c}$, where the last inequality follows from $\sqrt{\log p} \lesssim \Gamma/\sigma \leq |\Gamma_i|/\sigma \lesssim \sqrt{n}$ and $\Lambda \gtrsim \sigma^2 n\big(\frac{\sigma^2 n}{\Gamma^2} + \frac{\sigma\sqrt{p}}{\Gamma}\big)$. If instead $|\Gamma_i|/\sigma \gtrsim \sqrt{n}$, we have $P_i \leq \Phi\big(C\sqrt{\delta}\frac{|\Gamma_i|}{\sigma} + \frac{\Gamma_i}{\sigma}\big) + \frac{C}{p^c} \leq \frac{C'}{p^c}$, where the last inequality follows from $|\Gamma_i|/\sigma \gtrsim \sqrt{n} \gtrsim \sqrt{\log p}$ and $\delta = o(1)$. Hence, in Case 1, (8.4) can be bounded by $O(p^{-c})$.

Case 2. $\log p \gtrsim n$. In this case, we have $P_i \leq \Phi\big(C\sqrt{\delta}\big[\log p + \frac{\Gamma_i^2}{\sigma^2} + \frac{|\Gamma_i|}{\sigma}\sqrt{\log p}\big]^{1/2} + \frac{\Gamma_i}{\sigma}\big) + \frac{C}{p^c}$. In addition, since $|\Gamma_i| \geq \Gamma \gtrsim \sigma\sqrt{\log p}$ and $\delta = o(1)$, we have $P\big(\sum_{k=1}^n \hat{u}_k(X_{ki} - X_{k,i+1}) \geq 0\big) \leq \Phi\big(\frac{C\sqrt{\delta}}{\sigma}|\Gamma_i| + \frac{\Gamma_i}{\sigma}\big) + \frac{C}{p^c} \leq \frac{C'}{p^c}$. This shows that, in Case 2, (8.4) can also be bounded by $O(p^{-c})$.

As a result, it follows that, up to a change of sign for $\hat{w}$, $P((\mathfrak{r}(\hat{w}^\top X))^{-1} \neq id) = O(p^{-c})$ for some constant $c > 0$. $\qquad\square$

Proof of Theorem 2. Firstly, by invariance property of Kendall's tau distance, $\mathbb{E}[\tau_K(\hat{\pi}, \pi)] = \mathbb{E}[\tau_K((\mathfrak{r}(\hat{w}^\top X))^{-1}, id)] = \mathbb{E}[\tau_K((\mathfrak{r}(\hat{w}^\top X)), id)]$. It then follows

$$\mathbb{E}[\tau_K(\hat{\pi}, \pi)] = \frac{2}{p(p-1)} \sum_{i<j} P([\mathfrak{r}(\hat{w}^\top X)]_i \geq [\mathfrak{r}(\hat{w}^\top X)]_j)$$

$$= \frac{2}{p(p-1)} \sum_{i<j} P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\Big).$$

The summation in the last expression can be divided into two parts, namely, the consecutive differences and non-consecutive differences, i.e.,

$$\sum_{i<j} P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\Big) = \sum_{(i,j):j=i+1} P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\Big)$$

$$+ \sum_{(i,j):j>i+1} P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\Big).$$

In the following, we first show

$$P_i = P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\Big) \leq \frac{ce^{-\Gamma^2/2\sigma^2}}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c} \qquad (8.5)$$

so that

$$\sum_{(i,j):j=i+1} P\Big(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\Big) \leq \frac{cpe^{-\Gamma^2/2\sigma^2}}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c}. \qquad (8.6)$$

23

for some $C, c > 0$. Then we show that

$$\sum_{(i,j):j>i+1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq C\frac{p\sigma}{\Gamma} \min\left\{1, e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{C}{p^c}. \quad (8.7)$$

Combining (8.6) and (8.7), we conclude that

$$\mathbb{E}[\tau_K(\hat{w}^\top Y, \pi)] \leq \frac{C\sigma}{p\Gamma} \min\left\{1, e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{Ce^{-\Gamma^2/2\sigma^2}}{p(\Gamma/\sigma + \sqrt{8/\pi})} + \frac{C}{p^{c+2}},$$

which completes the proof, as the bound $\mathbb{E}[\tau_K(\hat{w}^\top Y, \pi)] \leq 1$ is trivial.

Proof of (8.5). Following the same argument as the proof of Theorem 1, we have for $1 \leq i \leq p - 1$ and $\delta = \frac{\sigma^2(n + \log p)(\lambda_1^2 + \sigma^2 p)}{\lambda_1^4}$, $P_i \leq P\left(w^\top \phi_{ij} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|_2\right) + P\left(|1 - (\hat{w}^\top w)^2| > \delta\right)$, where the second term can be bounded using Lemma 2. For the first term, by Lemma 1, for $\lambda_1^4(\Theta') \geq \sigma^2(p + \sigma^2\lambda_1^2(\Theta'))(\log p + n)$, we have

$$P_i \leq \Phi\left(C\sqrt{\delta}\left[(\sqrt{n} + \sqrt{\log p})^2 + \frac{\Gamma_i^2}{\sigma^2} + \frac{|\Gamma_i|}{\sigma}\sqrt{\log p}\right]^{1/2} + \frac{\Gamma_i}{\sigma}\right) + \frac{C}{p^c}.$$

Using same argument as the proof of Theorem 1, it holds that $P_i \leq \Phi\left(\frac{\Gamma_i}{\sigma}\right) + \frac{C}{p^c}$. Equation (8.5) then follows by using formula 7.1.13 of Abramowitz and Stegun (1965) that $\Phi(-t) < \frac{2}{t + \sqrt{t^2 + 8/\pi}}\phi(t)$ for $t \geq 0$.

Proof of (8.7). For the set of indices $S = \{(i,j) : 1 \leq i < j \leq p, j > i + 1\}$, we further divide it into two subsets $S_1 = \{(i,j) : 1 \leq i < j \leq p, j > i + \lfloor \sigma\sqrt{C \log p}/\Gamma \rfloor\}$ and $S_2 = \{(i,j) : 1 \leq i < j \leq p, i + 1 < j \leq i + \lfloor \sigma\sqrt{C \log p}/\Gamma \rfloor\}$ for some constant $C > 0$. Apparently we have the decomposition

$$\sum_{(i,j):j>i+1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) = \sum_{(i,j)\in S_1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right)$$

$$+ \sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \quad (8.8)$$

For the first term, by construction, it can be shown using the same argument (see supplementary materials) in Theorem 1 that

$$\sum_{(i,j)\in S_1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{C|S_1|}{p^c} \leq \frac{C}{p^{c_0}}. \quad (8.9)$$

Now for the second term in (8.8), similar argument yields, for $(i, j) \in S_2$,

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{c \exp(-|i-j|^2 \Gamma^2/(2\sigma^2))}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c}.$$

Note that, on the one hand,

$$\frac{e^{-|i-j|^2\Gamma^2/(2\sigma^2)}}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} \leq \frac{\sigma e^{-|i-j|^2\Gamma^2/2\sigma^2}}{|i-j|\Gamma}.$$

We have

$$\sum_{(i,j)\in S_2} \frac{\sigma e^{-|i-j|^2\Gamma^2/2\sigma^2}}{|i-j|\Gamma} = \frac{\sigma}{\Gamma} \sum_{k=2}^{p \wedge \lfloor \sqrt{\log p}/\Gamma \rfloor} \left(\frac{e^{-k^2\Gamma^2/2\sigma^2}p}{k} - e^{-k^2\Gamma^2/2\sigma^2}\right) = T_1 - T_2.$$

For the rest of the proof, we assume $\Gamma \leq \sigma\sqrt{\log p}/2$, otherwise the set S_2 will vanish. Then

$$T_1 = \frac{\sigma p}{\Gamma} \sum_{k=2}^{p \wedge \lfloor \sigma\sqrt{\log p}/\Gamma \rfloor} \frac{e^{-k^2\Gamma^2/2\sigma^2}}{k} \leq \frac{\sigma p}{\Gamma} \int_1^{p \wedge \sigma\sqrt{\log p}/\Gamma} \frac{e^{-x^2\Gamma^2/2\sigma^2}}{x} dx$$

where the last inequality used monotonicity of the integrand. The integral in the last inequality, after change of variable, can be bounded by an exponential integral $\mathrm{Ei}(\Gamma^2/2\sigma^2)$, which has an upper bound

$$\int_1^{p \wedge \sigma\sqrt{\log p}/\Gamma} \frac{e^{-x^2\Gamma^2/2\sigma^2}}{x} dx = \frac{1}{2} \int_{\Gamma^2/2\sigma^2}^{(p^2\Gamma^2/2\sigma^2) \wedge (\log p/2)} \frac{e^{-t}}{t} dt \leq \frac{1}{2} \int_{\Gamma^2/2\sigma^2}^{\infty} \frac{e^{-t}}{t} dx \leq e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)$$

so that $T_1 \leq \frac{\sigma p}{\Gamma} e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)$. For T_2, we have $T_2 = \frac{\sigma}{\Gamma} \sum_{k=2}^{p \wedge \lfloor \sigma\sqrt{\log p}/\Gamma \rfloor} e^{-k^2\Gamma^2/2\sigma^2} \geq \frac{\sigma}{\Gamma} e^{-2\Gamma^2/\sigma^2}$. Therefore,

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{C\sigma p}{\Gamma} e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right) + \frac{C}{p^c}. \tag{8.10}$$

On the other hand, note that

$$\frac{e^{-|i-j|^2\Gamma^2/(2\sigma^2)}}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} \leq c e^{-|i-j|^2\Gamma^2/(2\sigma^2)}.$$

We have

$$\sum_{(i,j)\in S_2} e^{-|i-j|^2\Gamma^2/(2\sigma^2)} = \sum_{k=2}^{p\wedge\lfloor\sigma\sqrt{\log p}/\Gamma\rfloor} pe^{-k^2\Gamma^2/(2\sigma^2)} - \sum_{k=2}^{p\wedge\lfloor\sigma\sqrt{\log p}/\Gamma\rfloor} ke^{-k^2\Gamma^2/(2\sigma^2)}$$

$$\leq p\int_1^\infty e^{-k^2\Gamma^2/2\sigma^2}dk - 2e^{-2\Gamma^2/\sigma^2} \leq Cp\sigma/\Gamma.$$

Thus

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki}-X_{kj})\geq 0\right) \leq Cp\sigma/\Gamma + \frac{C'}{p^c}. \tag{8.11}$$

Combining (8.10) and (8.11), we have

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki}-X_{kj})\geq 0\right) \leq C\frac{p\sigma}{\Gamma}\min\left\{1, e^{-\Gamma^2/2\sigma^2}\log\left(1+\frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{C}{p^c} \tag{8.12}$$

Combining (8.9) and (8.12), we have (8.7). □

Proof of Theorem 3 and Theorem 4. Here we only provide a sketch of the proofs. We refer the readers to our Supplementary Material for detailed proofs. The proofs follow essentially from the same argument as the proofs of Theorem 1 and Theorem 2, respectively. However, in place of Lemma 2 used therein, we need the following lemma that provides a perturbation bound for the leading eigenvector of approximate rank-one matrices, which could be of independent interest.

Lemma 3. *Suppose $p \gtrsim n$ and $\lambda_1^2(\Theta') \geq \lambda_2^2(\Theta') + C\sigma^2(n+\sqrt{np})$ for some $C > 0$. Let $w = u'_1$ be the first left singular vector of Θ', it follows that,*

$$P\left(|1-(\hat{w}^\top w)^2| \leq \frac{C\sigma^2(\lambda_1^2(\Theta')+\sigma^2 p)(n+\log p)}{(\lambda_1^2(\Theta')-\lambda_2^2(\Theta'))^2}\right) \geq 1 - \frac{C}{p^c}.$$

The proof of Lemma 8 is nontrivial, which depends on a combination of the generic perturbation bound obtained by Cai and Zhang (2018) and new concentration inequalities of approximate rank-one matrices (see Supplementary Materials).

Proof of Theorem 5. The proof relies on the following lemma adapted from (Tsybakov, 2009).

Lemma 4. *Assume that for some integer $M \geq 2$ there exist distinct parameters $\theta_0,...,\theta_M$ from the parameter space Θ and mutually absolutely continuous probability measures $P_0,...,P_M$ with $P_j = P_{\theta_j}$ for $j = 0,1,...,M$, defined on a common probability space $(\Omega,\mathcal{F})$ such that the averaged*

K-L divergence $\frac{1}{M}\sum_{j=1}^{M} D(P_j, P_0) \le \frac{1}{8}\log M$. Then, for every measurable mapping $\hat\theta : \Omega \to \Theta$,

$$\max_{j=0,...,M} P_j(\hat\theta \ne \theta_j) \ge \frac{\sqrt{M}}{\sqrt{M}+1}\left(\frac{3}{4} - \frac{1}{2\sqrt{\log M}}\right).$$

We construct the $(M+1) = p$ points parameter space as follows. We define p permutations from $\mathcal{S}_p$ as an identity plus $(p-1)$ consecutive swaps, i.e., $\pi_0 = id$, $\pi_k = (k, k+1)$ for $k = 1, ..., p-1$. The signal matrix $\Theta_0 = a\eta^\top$ where $a = (1, ..., 1)^\top \in \mathbb{R}^n$ and $\eta = (0, \delta, ..., (p-1)\delta)^\top \in \mathbb{R}^p$, $\delta = \frac{\sigma}{4}\sqrt{\log p/n}$. In this way, we have $\Gamma = \|a\|_2 \cdot \min_{1 \le i \le p-1}|\eta_j - \eta_{j+1}| = \frac{\sigma}{4}\sqrt{\log p}$. Let P_k corresponds to the joint probability measure of Y under (Θ_0, π_k) for $k = 0, 1, ..., p-1$, and let p_k be the pdf of P_k, we have $p_0(x) = \prod_{i=1}^{n}\prod_{j=1}^{p}\phi_{\eta_j}(x_{ij})$, $p_k(x) = \prod_{i=1}^{n}\prod_{j=1}^{p}\phi_{\eta_{\pi_k(j)}}(x_{ij})$ for $k = 1, ..., p-1$, where ϕ_μ is the pdf of Gaussian distribution $N(\mu, \sigma^2)$. Now we calculate the KL-divergence

$$D(P_k, P_0) = \int \log\left(\frac{p_k(x)}{p_0(x)}\right)p_0(x)dx = \int \frac{n}{2\sigma^2}\sum_{i=1}^{p}[(x_{1j} - \eta_{\pi_k(j)})^2 - (x_{1j} - \eta_j)^2]p_0(x)dx$$

$$= \frac{n\delta^2}{\sigma^2} = \frac{\log p}{16}.$$

Then, we have for $p \ge 10$, $\frac{1}{p-1}\sum_{k=1}^{p-1} D(P_k, P_0) = \frac{\log p}{16} \le \frac{1}{8}\log(p-1)$. It follows from Lemma 4 that, $\inf_{\hat\pi}\sup_{(\pi,\Theta)\in\mathcal{D}_1} P(\hat\pi \ne \pi) \ge \inf_{\hat\pi}\max_{j=0,...,p-1} P_j(\hat\pi \ne \pi_j) \ge 0.3$ as long as $p \ge 10$. In addition, $\inf_{\hat\pi}\sup_{(\pi,\Theta)\in\mathcal{D}_1'} P(\hat\pi \ne \pi) \ge \inf_{\hat\pi}\sup_{(\pi,\Theta)\in\mathcal{D}_1} P(\hat\pi \ne \pi)$ as $\mathcal{D}_1 \subset \mathcal{D}_1'$. $\qquad\square$

Proof of Theorem 6. The proof relies on the following lemma from Tsybakov (2009).

Lemma 5. *Assume that $M \ge 2$ and suppose that Θ contains elements $\theta_0, \theta_1, ..., \theta_M$ such that: (i) $d(\theta_j, \theta_k) \ge 2s > 0$ for any $0 \le j < k \le M$; (ii) for any $j = 1, ..., M$, $\frac{1}{M}\sum_{j=1}^{M} D(P_j, P_0) \le \alpha\log M$ with $0 < \alpha < 1/8$ and $P_j = P_{\theta_j}$ for $j = 0, 1, ..., M$. Then*

$$\inf_{\hat\theta}\sup_{\theta\in\Theta} P_\theta(d(\hat\theta, \theta) \ge s) \ge \frac{\sqrt{M}}{1+\sqrt{M}}\left(1 - 2\alpha - \sqrt{\frac{2\alpha}{\log M}}\right) > 0.$$

We also need the following sphere packing lemma proved by Mao et al. (2017), which is a direct consequence of the well-celebrated Varshamov-Gilbert bound.

Lemma 6. *For any $r < p/2$, there exists a subset $\mathcal{Q}_r$ of $\mathcal{S}_p$ such that (i) $\log|\mathcal{Q}_r| \ge \frac{r}{5}\log(p/r)$, (ii) for any elements $\pi_1, \pi_2 \in \mathcal{Q}_r$, we have $\binom{p}{2} \cdot \tau_K(\pi_1, \pi_2) \ge r$, and (iii) for any $\pi \in \mathcal{Q}_r$, we have $\|\pi - id\|_2^2 \le 2r$.*

For $t/\sigma \ge 2$, we set $r = \frac{p\sigma}{t}e^{-t^2/2\sigma^2} < p/2$. Let $\pi_0 = id$ and $\pi_1, ..., \pi_{|\mathcal{Q}_r|}$ be the elements of $\mathcal{Q}_r$. The signal matrix $\Theta_0 = a\eta^\top$ where $a = (1/\sqrt{320n}, ..., 1/\sqrt{320n})^\top \in \mathbb{R}^n$ and $\eta = (t, ..., pt)^\top \in \mathbb{R}^p$.

Let P_k be the joint probability measure of Y under (Θ_0, π_k) for $k = 0, 1, ..., |\mathcal{Q}_r|$, and let p_k be the pdf of P_k. By Lemma 6, the KL-divergence

$$D(P_k, P_0) = \int \log\left(\frac{p_k(x)}{p_0(x)}\right) p_0(x) dx = \frac{t^2}{320\sigma^2} \|\pi_k - id\|_2^2 \leq \frac{pt}{160\sigma} e^{-t^2/2\sigma^2}.$$

and therefore

$$\frac{1}{p-1} \sum_{k=1}^{p-1} D(P_k, P_0) \leq \frac{pt}{160\sigma} e^{-t^2/2\sigma^2} \leq \frac{p\sigma}{80t} e^{-t^2/2\sigma^2} \log\left(\frac{t}{\sigma} e^{t^2/2\sigma^2}\right) \leq \frac{1}{16} \log|\mathcal{Q}_r|.$$

Without loss of generality, we assume $|\mathcal{Q}_r| \geq 2$. By Lemma 5, it then follows that,

$$\inf_{\hat{\pi}} \sup_{(\Theta,\pi)\in\mathcal{D}_2(t)} P\left(\tau_K(\hat{\pi}, \pi) \geq \frac{\sigma}{2pt} e^{-t^2/2\sigma^2}\right) \geq C_1,$$

for some absolute constant $C_1 > 0$. By Markov's inequality, we have

$$\inf_{\hat{\pi}} \sup_{(\Theta,\pi)\in\mathcal{D}_2(t)} \mathbb{E}[\tau_K(\hat{\pi}, \pi)] \geq \frac{\sigma}{2pt} e^{-t^2/2\sigma^2} \inf_{\hat{\pi}} \sup_{(\Theta,\pi)\in\mathcal{D}_2(t)} P\left(\tau_K(\hat{\pi}, \pi) \geq \frac{\sigma}{2pt} e^{-t^2/2\sigma^2}\right) \geq \frac{C_1\sigma}{pt} e^{-t^2/2\sigma^2}.$$

The relationship $\inf_{\hat{\pi}} \sup_{(\Theta,\pi)\in\mathcal{D}_2'(t)} \mathbb{E}[\tau_K(\hat{\pi}, \pi)] \geq \inf_{\hat{\pi}} \sup_{(\Theta,\pi)\in\mathcal{D}_2(t)} \mathbb{E}[\tau_K(\hat{\pi}, \pi)]$ follows from $\mathcal{D}_L \subset \mathcal{D}$. The rate $1/p^2$ follows by setting $t = C_2\sigma\sqrt{\log p}$ for some $C_2 > 0$. $\qquad\square$

SUPPLEMENTARY MATERIALS

In our online Supplemental Materials, we prove Theorem 3-4, Proposition 1-4, as well as the technical lemmas. Some supplementary simulations, figures and tables are included in the appendix.

References

Abel, S., P. A. Zur Wiesch, H.-H. Chang, B. M. Davis, M. Lipsitch, and M. K. Waldor (2015). Sequence tag–based analysis of microbial population dynamics. *Nature Methods 12*(3), 223.

Abramowitz, M. and I. A. Stegun (1965). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Volume 55. Courier Corporation.

Barg, A. and A. Mazumdar (2010). Codes in permutations and error correction for rank modulation. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pp. 854–858. IEEE.

Boulund, F., M. B. Pereira, V. Jonsson, and E. Kristiansson (2018). Computational and statistical considerations in the analysis of metagenomic data. In *Metagenomics*, pp. 81–102. Elsevier.

Bremer, H. and G. Churchward (1977). An examination of the cooper-helmstetter theory of dna replication in bacteria and its underlying assumptions. *Journal of Theoretical Biology 69*(4), 645–654.

Brown, C. T., M. R. Olm, B. C. Thomas, and J. F. Banfield (2016). Measurement of bacterial replication rates in microbial communities. *Nature Biotechnology 34*(12), 1256.

Cai, T. T., Z. Ma, and Y. Wu (2013). Sparse PCA: Optimal rates and adaptive estimation. *The Annals of Statistics 41*, 3074–3110.

Cai, T. T. and A. Zhang (2018). Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics. *The Annals of Statistics 46*(1), 60–89.

Chatterjee, S., A. Guntuboyina, and B. Sen (2015). On risk bounds in isotonic and other shape restricted regression problems. *The Annals of Statistics 43*(4), 1774–1800.

Chatterjee, S., A. Guntuboyina, and B. Sen (2018). On matrix estimation under monotonicity constraints. *Bernoulli 24*(2), 1072–1100.

Collier, O. and A. S. Dalalyan (2016). Minimax rates in permutation estimation for feature matching. *The Journal of Machine Learning Research 17*(1), 162–192.

Cooper, S. and C. E. Helmstetter (1968). Chromosome replication and the division cycle of escherichia coli b/r. *Journal of Molecular Biology 31*(3), 519–540.

Cullina, D. and N. Kiyavash (2016). Improved achievability and converse bounds for erdös-renyi graph matching. In *ACM SIGMETRICS Performance Evaluation Review*, Volume 44, pp. 63–72. ACM.

Currie, R. R. and G. S. Pandher (2011). Finance journal rankings and tiers: An active scholar assessment methodology. *Journal of Banking & Finance 35*(1), 7–20.

Deshpande, S. K. and S. T. Jensen (2016). Estimating an nba player's impact on his team's chances of winning. *Journal of Quantitative Analysis in Sports 12*(2), 51–72.

Diaconis, P. (1988). *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics Lecture Notes–Monograph Series (11).

Diaconis, P. and R. L. Graham (1977). Spearman's footrule as a measure of disarray. *Journal of the Royal Statistical Society. Series B (Methodological)*, 262–268.

Flammarion, N., C. Mao, and P. Rigollet (2019). Optimal rates of statistical seriation. *Bernoulli 25*(1), 623–653.

Gao, F., H. Luo, and C.-T. Zhang (2013). Doric 5.0: an updated database of oric regions in both bacterial and archaeal genomes. *Nucleic Acids Research 41*, D90.

Gao, Y. and H. Li (2018). Quantifying and comparing bacterial growth dynamics in multiple metagenomic samples. *Nature Methods 15*, 1041–1044.

Göloğlu, F., J. Lember, A.-E. Riet, and V. Skachek (2015). New bounds for permutation codes in ulam metric. In *Information Theory (ISIT), 2015 IEEE International Symposium on*, pp. 1726–1730. IEEE.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika 30*(1/2), 81–93.

Koltchinskii, V. and D. Xia (2016). Perturbation of linear forms of singular vectors under gaussian noise. In *High Dimensional Probability VII*, pp. 397–423. Springer.

Korem, T., D. Zeevi, J. Suez, A. Weinberger, T. Avnit-Sagi, M. Pompan-Lotan, E. Matot, G. Jona, A. Harmelin, and N. Cohen (2015). Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science*, aac4812.

Laurent, B. and P. Massart (2000). Adaptive estimation of a quadratic functional by model selection. *The Annals of Statistics 28*(5), 1302–1338.

Lewis, J. D., E. Z. Chen, R. N. Baldassano, A. R. Otley, A. M. Griffiths, D. Lee, K. Bittinger, A. Bailey, E. S. Friedman, C. Hoffmann, et al. (2015). Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric crohns disease. *Cell Host & Microbe 18*(4), 489–500.

Li, D., C. Liu, R. Luo, K. Sadakane, and T. Lam (2015). Megahit: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics 15*, 1674–1676.

Mao, C., J. Weed, and P. Rigollet (2017). Minimax rates and efficient algorithms for noisy sorting. *arXiv preprint arXiv:1710.10388*.

Mazumdar, A., A. Barg, and G. Zemor (2013). Constructions of rank modulation codes. *IEEE Transactions on Information Theory 59*(2), 1018–1029.

Mukherjee, S. (2016). Estimation in exponential families on permutations. *The Annals of Statistics 44*(2), 853–875.

Myhrvold, C., J. W. Kotula, W. M. Hicks, N. J. Conway, and P. A. Silver (2015). A distributed cell division counter reveals growth dynamics in the gut microbiota. *Nature Communications 6*, 10039.

Pananjady, A., M. J. Wainwright, and T. A. Courtade (2016). Linear regression with an unknown permutation: Statistical and computational limits. *arXiv preprint arXiv:1608.02902*.

Pananjady, A., M. J. Wainwright, and T. A. Courtade (2017). Denoising linear models with permuted data. In *Information Theory (ISIT), 2017 IEEE International Symposium on*, pp. 446–450. IEEE.

Rendle, S., L. Balby Marinho, A. Nanopoulos, and L. Schmidt-Thieme (2009). Learning optimal ranking with tensor factorization for tag recommendation. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 727–736. ACM.

Rigollet, P. and J. Weed (2018). Uncoupled isotonic regression via minimum wasserstein deconvolution. *arXiv preprint arXiv:1806.10648*.

Rudelson, M. and R. Vershynin (2013). Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability 18*.

Slawski, M. and E. Ben-David (2017). Linear regression with sparsely permuted data. *arXiv preprint arXiv:1710.06030*.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York.

von Meijenfeldt, F. B., K. Arkhipova, D. D. Cambuy, F. H. Coutinho, and B. E. Dutilh (2019). Robust taxonomic classification of uncharted microbial sequences and bins with cat and bat. *bioRxiv*, 530188.

Wu, Y., Y. Tang, S. Tringe, B. Simmons, and S. Singer (2014). Maxbin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome 2*, 26.

Yuan, X. and T. Zhang (2013). Truncated power method for sparse eigenvalue problems. *Journal of Machine Learning Research 14*, 899–925.

Zhang, A., T. T. Cai, and Y. Wu (2018). Heteroskedastic pca: Algorithm, optimality, and applications. *arXiv preprint arXiv:1810.08316*.

Supplement to "Optimal Permutation Recovery in Permuted Monotone Matrix Model"

Rong Ma[1], T. Tony Cai[2] and Hongzhe Li[1]
Department of Biostatistics, Epidemiology and Informatics[1]
Department of Statistics[2]
University of Pennsylvania
Philadelphia, PA 19104

Abstract

In this Supplementary Material, we prove Theorem 3 & 4 and Proposition 1 to 4 in the main paper and the technical lemmas. The numerical comparisons of $\hat{\pi}$ and an SVD-based estimator $\tilde{\pi}$ are also included as an Appendix.

1 Proofs of Theorem 3 and 4

Proof of Theorem 3. As in the proof of Theorem 1, it suffices to bound the probability $P\big(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\big)$. Note that $\hat{w}$ is the first eigenvector of

$$A = X\left(I - \frac{1}{p}ee^\top\right)\left(I - \frac{1}{p}ee^\top\right)X^\top = TT^\top$$

where $T \in \mathbb{R}^{n \times p}$ and T admits the decomposition

$$T = \Theta\left(I - \frac{1}{p}ee^\top\right) + Z\left(I - \frac{1}{p}ee^\top\right) = \Theta' + E \in \mathbb{R}^{n \times p}$$

where $E_{ij} \sim N(0, (p-1)\sigma^2/p)$ and Θ' has SVD

$$\Theta' = U'D'V'^\top.$$

Define $w = u'_1 \in \mathbb{R}^n$ as the first eigenvector of $\Theta'\Theta'^\top$. Following the same argument that leads to (13) of the main paper, we have, up to a change of sign for $\hat{w}$,

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq P\left(w^\top \phi_{i,i+1} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|_2\right) + P\left(|1 - (\hat{w}^\top w)^2| > \delta\right).$$

The following lemmas parallel Lemma 1 and Lemma 2 in the proof of Theorem 1. In particular, Lemma 8 provides a perturbation bound for the leading eigenvector of approximate rank-one matrices, which could be of independent interest.

Lemma 7. *Under the conditions of Theorem 3, let $\Xi_i = \|\Theta'_{.i} - \Theta'_{.i+1}\|_2$. Then for any $\delta > 0$, we have*

$$P\left(w^\top \phi_{i,i+1} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\| \right) \tag{1.1}$$

$$\leq \Phi\left(C\sqrt{\delta}\left[(\sqrt{n} + \sqrt{\log p})^2 + \frac{\Xi_i^2}{\sigma^2} + \sqrt{\log p}\frac{\Xi_i}{\sigma}\right]^{1/2} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c} \tag{1.2}$$

for some universal constant $C > 0$.

Lemma 8. *Suppose $n \lesssim p$ and $\lambda_1^2(\Theta') \geq \lambda_2^2(\Theta') + C\sigma^2(n + \sqrt{np})$ for some $C > 0$, it follows that*

$$P\left(|1 - (\hat{w}^\top w)^2| \leq \frac{C\sigma^2(\lambda_1^2(\Theta') + \sigma^2 p)(n + \log p)}{(\lambda_1^2(\Theta') - \lambda_2^2(\Theta'))^2} \right) \geq 1 - \frac{C}{p^c}.$$

Combining Lemma 7 and Lemma 8, since

$$\lambda_1^2(\Theta') - \lambda_2^2(\Theta') > C_0\sigma^2(n + \sqrt{np}) \quad \text{and let} \quad \delta = \frac{C_0\sigma^2(\lambda_1^2(\Theta') + \sigma^2 p)(n + \log p)}{(\lambda_1^2(\Theta') - \lambda_2^2(\Theta'))^2} \tag{1.3}$$

for some $C_0 > 0$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0 \right)$$

$$\leq \Phi\left(C\sqrt{\delta}\left[(\sqrt{n} + \sqrt{\log p})^2 + \frac{\Xi_i^2}{\sigma^2} + \frac{\Xi_i}{\sigma}\sqrt{\log p}\right]^{1/2} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

for some $C, c > 0$. The rest of the analysis is divided into several cases.

Case 1. $\log p \lesssim n$. In this case, under (1.3), we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0 \right)$$

$$\leq \Phi\left(C\sqrt{\delta}\left[n + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma}\sqrt{\log p}\right]^{1/2} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

then

33

1. if $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \lesssim \sigma\sqrt{n}$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta n} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c} \leq \frac{C'}{p^c},$$

where the last inequality follows from

$$\sqrt{\log p} \lesssim \Gamma/\sigma, \qquad \lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2 n(n + \log p)}{\Gamma^2/\sigma^2} + \frac{\sigma^2\sqrt{np(n + \log p)}}{\Gamma/\sigma}.$$

2. $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \gtrsim \sigma\sqrt{n}$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta}\frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c} \leq \frac{C'}{p^c},$$

where the last inequality follows from

$$\Gamma/\sigma \gtrsim \sqrt{n} \gtrsim \sqrt{\log p}, \qquad \lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2\Xi^2(n + \log p)}{\Gamma^2} + \frac{\sigma^2\Xi\sqrt{p(n + \log p)}}{\Gamma}.$$

This completes the proof of case 1.

Case 2. $\log p \gtrsim n$. In this case, under (1.3), we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right)$$

$$\leq \Phi\left(C\sqrt{\delta}\left[\log p + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma}\sqrt{\log p}\right]^{1/2} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

In addition, since $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \gtrsim \sigma\sqrt{\log p}$, we have

$$P\left(\sum_{k=1}^n \hat{u}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(\frac{C\sqrt{\delta}}{\sigma}\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c} \leq \frac{C'}{p^c}$$

where the last inequality follows from

$$\Gamma/\sigma \gtrsim \sqrt{\log p} \gtrsim \sqrt{n}, \qquad \lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2\Xi^2(n + \log p)}{\Gamma^2} + \frac{\sigma^2\Xi\sqrt{p(n + \log p)}}{\Gamma}.$$

This completes the proof of case 2. As a result, it follows that, up to a change of sign of $\hat{w}$,

$$P((\mathfrak{r}(\hat{w}^\top X))^{-1} \neq id) \leq \sum_{i=1}^{p-1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \frac{C}{p^c}$$

34

for some constant $C, c > 0$. $\qquad\square$

Proof of Theorem 4. Similar to the proof of Theorem 2, we write

$$\mathbb{E}[\tau_K(\hat{w}^\top Y, \pi)] = \frac{2}{p(p-1)} \sum_{i<j} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right)$$

where

$$\sum_{i<j} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) = \sum_{(i,j):j=i+1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right)$$
$$+ \sum_{(i,j):j>i+1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right).$$

In the following, we first show

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \frac{ce^{-\Gamma^2/2\sigma^2}}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c} \qquad (1.4)$$

for some constant $c > 0$ and therefore

$$\sum_{(i,j):j=i+1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{cpe^{-\Gamma^2/2\sigma^2}}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c}$$

Then we show that

$$\sum_{(i,j):j>i+1} P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq C\frac{p\sigma}{\Gamma} \min\left\{1, e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{C}{p^c} \qquad (1.5)$$

and conclude that

$$\mathbb{E}[\tau_K(\hat{w}^\top Y, \pi)] \leq \frac{C\sigma}{p\Gamma} \min\left\{1, e^{-\Gamma^2/2\sigma^2} \log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{Ce^{-\Gamma^2/2\sigma^2}}{p(\Gamma/\sigma + \sqrt{8/\pi})} + \frac{C}{p^{c+2}}.$$

The bound $\mathbb{E}[\tau_K(\hat{w}^\top Y, \pi)] \leq 1$ is trivial.

Proof of (1.4). Following the same argument as in the proof of Theorem 3, we have for $1 \leq i \leq p-1$ and $\delta = \frac{\sigma^2(n+\log p)(\lambda_1^2(\Theta')+\sigma^2 p)}{(\lambda_1^2(\Theta')-\lambda_2^2(\Theta'))^2}$,

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq P\left(w^\top \phi_{ij} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|_2\right) + P\left(|1 - (\hat{w}^\top w)^2| > \delta\right)$$

$$\leq P\left(w^\top \phi_{i,i+1} \geq -\sqrt{2\delta}\|\phi_{i,i+1}\|_2\right) + \frac{C}{p^c}$$

where the last inequality follows from Lemma 8. For the first term, by Lemma 7, for $\delta < 1$,

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta}\left((\sqrt{n} + \sqrt{\log p})^2 + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2^2}{\sigma^2}\right.\right.$$

$$\left.\left. + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma}\sqrt{\log p}\right)^{1/2} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}.$$

The rest of the analysis is divided into several cases.

Case 1. $\log p \leq n$. In this case, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta}\left(n + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma}\sqrt{\log p}\right)^{1/2}\right.$$

$$\left. + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}.$$

In addition,

1. if $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \lesssim \sigma\sqrt{n}$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{k,i+1}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta n} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

$$\leq \Phi\left(\frac{C\Gamma}{\sigma}\right) + \frac{C}{p^c}$$

where the last inequality follows from

$$\lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2 n(n + \log p)}{\Gamma^2/\sigma^2} + \frac{\sigma^2\sqrt{np(n + \log p)}}{\Gamma/\sigma}.$$

Thus, using Formula 7.1.13 from Abramowitz and Stegun (1965),

$$\Phi(-x) < \frac{2}{t + \sqrt{t^2 + 8/\pi}}\phi(x), \tag{1.6}$$

for $1 \le i \le p-1$, we have

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \ge 0\right) \le \frac{C}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} \exp(-\Gamma^2/(2\sigma^2)) + \frac{C}{p^c}.$$

2. if $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \gtrsim \sigma\sqrt{n}$, we have

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \ge 0\right) \le \Phi\left(\frac{C\sqrt{\delta}}{\sigma}\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

$$\le \frac{C}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} \exp(-\Gamma^2/(2\sigma^2)) + \frac{C}{p^c},$$

where the last inequality follows from (1.6) and

$$\lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2 \Xi^2 (n + \log p)}{\Gamma^2} + \frac{\sigma^2 \Xi \sqrt{p(n + \log p)}}{\Gamma}.$$

This completes the proof of case 1.

Case 2. $\log p > n$. In this case, we have

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{k,i+1}) \ge 0\right) \le \Phi\left(C\sqrt{\delta}\left(\log p + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.i+1}\|_2}{\sigma}\sqrt{\log p}\right)^{1/2}\right.$$

$$\left. + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}.$$

In addition,

1. if $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \gtrsim \sigma\sqrt{\log p}$, we have

$$P\left(\sum_{k=1}^{n} \hat{u}_k(X_{ki} - X_{k,i+1}) \ge 0\right) \le \Phi\left(\frac{C\sqrt{\delta}}{\sigma}\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

$$\le \frac{C}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} \exp(-\Gamma^2/(2\sigma^2)) + \frac{C}{p^c}.$$

2. if $\|\Theta'_{.i} - \Theta'_{.i+1}\|_2 \lesssim \sigma\sqrt{\log p}$, we have

$$P\left(\sum_{k=1}^{n} \hat{u}_k(X_{ki} - X_{k,i+1}) \ge 0\right) \le \Phi\left(C\sqrt{\delta \log p} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma}\right) + \frac{C}{p^c}$$

$$\le \frac{\sqrt{2/\pi}}{\Gamma/\sigma + \sqrt{\Gamma^2/\sigma^2 + 8/\pi}} \exp(-\Gamma^2/(2\sigma^2)) + \frac{C}{p^c}$$

37

where the last inequality follows from

$$\lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2 \log p(n + \log p)}{\Gamma^2/\sigma^2} + \frac{\sigma^2 \sqrt{p \log p(n + \log p)}}{\Gamma/\sigma}.$$

This completes the proof of Case 2 and in sum we have proven (1.4).

Proof of (1.5). Following the same construction of S_1 and S_2 in the proof of (18) of the main paper, we have

$$\sum_{(i,j):j>i+1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right) = \sum_{(i,j)\in S_1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right)$$

$$+ \sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right) \qquad (1.7)$$

For the first term, for any $(i,j) \in S_1$, we have

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right) \leq P\left(w^\top \phi_{ij} \geq -\sqrt{2\delta}\|\phi_{ij}\|_2 \right) + P\left(|1 - (\hat{w}^\top w)^2| > \delta \right)$$

$$\leq P\left(w^\top \phi_{ij} \geq -\sqrt{2\delta}\|\phi_{ij}\|_2 \right) + \frac{C}{p^c}$$

where the last inequality follows from Lemma 8. To bound $P\left(w^\top \phi_{ij} \geq -\sqrt{2\delta}\|\phi_{ij}\|_2 \right)$, similar argument as in Theorem 3 implies, for $\delta < 1$,

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right) \leq \Phi\left(C\sqrt{\delta}\left((\sqrt{n} + \sqrt{\log p})^2 + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2}{\sigma}\sqrt{\log p} \right)^{1/2} \right.$$

$$\left. + \frac{w^\top(\Theta'_{.i} - \Theta'_{.i+1})}{\sigma} \right) + \frac{C}{p^c}.$$

The rest of the analysis is divided into two cases.

Case 1. $\log p \lesssim n$. In this case, we have

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0 \right) \leq \Phi\left(C\sqrt{\delta}\left(n + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2}{\sigma}\sqrt{\log p} \right)^{1/2} \right.$$

$$\left. + \frac{w^\top(\Theta'_{.i} - \Theta'_{.j})}{\sigma} \right) + \frac{C}{p^c}.$$

In addition, let $\Gamma_{ij} = \lambda_1(\Theta')|v'_{1i} - v'_{1j}| \geq |i - j|\Gamma$,

38

1. if $\|\Theta'_{.i} - \Theta'_{.j}\|_2 \lesssim \sigma\sqrt{n}$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta n} + \frac{w^\top(\Theta'_{.i} - \Theta'_{.j})}{\sigma}\right) + \frac{C}{p^c} \leq \frac{C}{p^c},$$

since

$$\lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2 n(n + \log p)}{\Gamma^2/\sigma^2} + \frac{\sigma^2\sqrt{np(n + \log p)}}{\Gamma/\sigma}$$

$$\geq \lambda_2^2(\Theta') + \frac{\sigma^2 n(n + \log p)}{\Gamma_{ij}^2/\sigma^2} + \frac{\sigma^2\sqrt{np(n + \log p)}}{\Gamma_{ij}/\sigma},$$

and $\Gamma_{ij} \geq \sigma\sqrt{C\log p}$ for $(i,j) \in S_1$.

2. if $\|\Theta'_{.i} - \Theta'_{.j}\|_2 \gtrsim \sigma\sqrt{n}$, then let $\Xi_{ij} = \|\Theta'_{.i} - \Theta'_{.j}\|_2 \leq |i - j|\Xi$, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \Phi\left(\frac{C\sqrt{\delta}}{\sigma}\|\Theta'_{.i} - \Theta'_{.j}\|_2 + \frac{w^\top(\Theta'_{.i} - \Theta'_{.j})}{\sigma}\right) + \frac{C}{p^c} \leq \frac{C}{p^c},$$

where the last inequality follows from

$$\lambda_1^2(\Theta') \gtrsim \lambda_2^2(\Theta') + \frac{\sigma^2\Xi^2(n + \log p)}{\Gamma^2} + \frac{\sigma^2\Xi\sqrt{p(n + \log p)}}{\Gamma}$$

$$\geq \lambda_2^2(\Theta') + \frac{\sigma^2\Xi_{ij}^2(n + \log p)}{\Gamma_{ij}^2} + \frac{\sigma^2\Xi_{ij}\sqrt{p(n + \log p)}}{\Gamma_{ij}},$$

and $\Gamma_{ij} \geq \sigma\sqrt{C\log p}$.

This completes the proof of case 1.

Case 2. $\log p \gtrsim n$. In this case, we have

$$P\left(\sum_{k=1}^n \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \Phi\left(C\sqrt{\delta}\left(\log p + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2^2}{\sigma^2} + \frac{\|\Theta'_{.i} - \Theta'_{.j}\|_2}{\sigma}\sqrt{\log p}\right)^{1/2}\right.$$

$$\left. + \frac{w^\top(\Theta'_{.i} - \Theta'_{.j})}{\sigma}\right) + \frac{C}{p^c}.$$

In addition, since $\Gamma_{ij} \geq \sigma\sqrt{C\log p}$, similar arguments yield

$$P\left(\sum_{k=1}^n \hat{u}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{C}{p^c}.$$

This completes the proof of Case 2. Combining Case 1 and Case 2, we have proven

$$\sum_{(i,j)\in S_1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{C|S_1|}{p^c} \leq \frac{C}{p^{c_0}}. \tag{1.8}$$

Now for the second term in (1.7), since $\Gamma|i - j| \leq \Gamma_{ij}$, we have for $(i,j) \in S_2$,

$$P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{ce^{-|i-j|^2\Gamma^2/(2\sigma^2)}}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} + \frac{C}{p^c}.$$

Note that, on the one hand,

$$\frac{ce^{-|i-j|^2\Gamma^2/(2\sigma^2)}}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} \leq \frac{c\sigma e^{-|i-j|^2\Gamma^2/2\sigma^2}}{|i-j|\Gamma}.$$

Following similar argument that leads to (22) of the main paper, we have

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq \frac{C\sigma p}{\Gamma}e^{-\Gamma^2/2\sigma^2}\log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right) + \frac{C}{p^c}. \tag{1.9}$$

On the other hand, note that

$$\frac{e^{-|i-j|^2\Gamma^2/(2\sigma^2)}}{|i-j|\Gamma/\sigma + \sqrt{|i-j|^2\Gamma^2/\sigma^2 + 8/\pi}} \leq ce^{-|i-j|^2\Gamma^2/(2\sigma^2)},$$

we have

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq Cp\sigma/\Gamma + \frac{C'}{p^c}. \tag{1.10}$$

Combining (1.9) and (1.10), we have

$$\sum_{(i,j)\in S_2} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq C\frac{p\sigma}{\Gamma}\min\left\{1, e^{-\Gamma^2/2\sigma^2}\log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{C}{p^c} \tag{1.11}$$

Combining (1.8) and (1.11), we have

$$\sum_{(i,j):j>i+1} P\left(\sum_{k=1}^{n} \hat{w}_k(X_{ki} - X_{kj}) \geq 0\right) \leq C\frac{p\sigma}{\Gamma}\min\left\{1, e^{-\Gamma^2/2\sigma^2}\log\left(1 + \frac{2\sigma^2}{\Gamma^2}\right)\right\} + \frac{C}{p^c},$$

which leads to (1.5). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2 Proofs of Proposition 1-4

Proof of Proposition 1. Since $\Theta \in \mathcal{D}$, for any nonnegative unit vector $w \in \mathbb{R}^n$, the vector $w^\top \Theta \in \mathbb{R}^p$ is monotonic increasing, so that $\mathfrak{r}(w^\top \Theta) = id$. It then follows that

$$\mathfrak{r}(w^\top \Theta \Pi) = \mathfrak{r}(w^\top \Theta) \circ \pi^{-1} = \pi^{-1}.$$

$\square$

Proof of Proposition 2. The key observation is, for any Γ, we have

$$\Lambda \geq C\Gamma^2 p^3. \tag{2.1}$$

Then if $\Gamma \gtrsim \sigma\sqrt{n}$ holds, apparently GSS can be implied by (2.1) and MSG; if $\Gamma \lesssim \sigma\sqrt{n}$, then the results depend on the relative magnitude of $\sigma^2 n^2/\Gamma^2$ and p. Specifically, if $p \gtrsim \sigma^2 n^2/\Gamma^2$, then $p \gtrsim (\sigma^3 n/\Gamma^3)^{2/5}$ implies GSS; if $p \lesssim \sigma^2 n^2/\Gamma^2$, then $p \gtrsim (\sigma^4 n^2/\Gamma^4)^{1/3}$ implies GSS. $\square$

Proof of Proposition 3. Note that by SVD of Θ', we have

$$v_1' = \arg\max_{\|x\|_2 = 1} x^\top \Theta^\top \Theta x = \arg\max_{\|x\|_2 = 1} \sum_{i=1}^{n} \left(\sum_{j=1}^{p} x_j \theta_{ij}' \right)^2. \tag{2.2}$$

To prove that v_1' is monotone, we need the following rearrangement inequality.

Lemma 9 (Rearrangement Inequality). *If $a_1 \geq a_2 \geq ... \geq a_n$ and $b_1 \geq b_2 \geq ... \geq b_n$, then*

$$a_n b_1 + ... + a_1 b_n \leq a_{\sigma(1)} b_1 + ... + a_{\sigma(n)} b_n \leq a_1 b_1 + ... + a_n b_n,$$

where σ is any permutation in S_n.

Lemma 10. *For any Θ' defined above, its first left singular vector u_1' is either nonpositive or nonnegative.*

Now since u_1' is either nonnegative or nonpositive, then we know that $\sum_{j=1}^{p} v_j \theta_{ij}'$ have the same sign for all $i = 1, ..., n$. By Lemma 9, we have that the components of v_1' are either in increasing order or in decreasing order. $\square$

Proof of Proposition 4. Note that

$$v_1' = \arg\max_{\|x\|_2 = 1} x^\top {\Theta'}^\top \Theta' x = \arg\max_{\|x\|_2 = 1} \sum_{i=1}^{n} (x^\top \Theta_{i.}')^2. \tag{2.3}$$

Let $f(x, \Theta', \lambda) = \sum_{i=1}^{n}(x^\top \Theta'_{i.})^2 + \lambda(\|x\|_2 - 1)$. By Lagrange's multiplier method, we have

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^{n} 2\theta_{ij}(x^\top \Theta'_{i.}) + 2\lambda x_j = 0,$$

for $j = 1, ..., p$. It follows that

$$\frac{\sum_{i=1}^{n} \theta'_{ij}(v'^{\top}_1 \Theta'_{i.})}{v'_{1j}} = -\lambda, \qquad \text{for all } j = 1, ..., p,$$

where

$$\lambda^2 = \sum_{j=1}^{p}\left[\sum_{i=1}^{n} \theta'_{ij}(v'^{\top}_1 \Theta'_{i.})\right]^2 = \|v'^{\top}_1 \Theta'^{\top}\Theta'\|_2^2.$$

Thus, by Lemma 10, we have

$$
\begin{aligned}
|v'_{1,j+1} - v'_{1,j}| &= \frac{\sum_{i=1}^{n} |\theta'_{i,j+1} - \theta'_{i,j}||v'^{\top}_1 \Theta'_{i.}|}{|\lambda|} \\
&\geq \frac{\sum_{i=1}^{n} \delta \|\Theta_{i.}\|_2 |v'^{\top}_1 \Theta'_{i.}|}{|\lambda|} \\
&\geq \delta,
\end{aligned}
$$

where the last inequality follows from

$$|\lambda| = \|v'^{\top}_1 \Theta'^{\top}\Theta'\|_2 = \left\|\sum_{i=1}^{n}(v'^{\top}_1 \Theta'_{i.})\Theta'_{i.}\right\|_2 \leq \sum_{i=1}^{n}\|(v'^{\top}_1 \Theta'_{i.})\Theta'_{i.}\|_2 \leq \sum_{i=1}^{n}(v'^{\top}_1 \Theta'_{i.})\|\Theta'_{i.}\|_2.$$

This completes the proof. $\square$

3   Proofs of Technical Lemmas

Proof of Lemma 1.   Note that

$$\phi^\top_{i,i+1}w = \sum_{k=1}^{n} w_k(T_{ki} - T_{k,i+1}) \sim N\left(\|a\|_2(\eta_i - \eta_{i+1}), \frac{2(p-1)\sigma^2}{p}\right),$$

and

$$\|\phi_{i,i+1}\|_2^2 = \sum_{k=1}^{n}(T_{ki} - T_{k,i+1})^2, \quad \text{where} \quad T_{ki} - T_{k,i+1} \sim N(a_k(\eta_i - \eta_{i+1}), 2(p-1)\sigma^2/p).$$

We construct the standardized chi-square statistic

$$Y_k^2 = \left[\frac{T_{k,i} - T_{k,i+1} - a_k(\eta_i - \eta_{i+1})}{\sigma\sqrt{2(p-1)/p}}\right]^2.$$

Define

$$Q = \frac{p-1}{p}2\sigma^2 \cdot \sum_{k=1}^{n}(Y_k^2 - 1),$$

we have

$$Q = \|\phi_{i,i+1}\|_2^2 + \|a\|_2^2(\eta_i - \eta_{i+1})^2 - 2(\eta_i - \eta_{i+1})\sum_{k=1}^{n}(T_{ki} - T_{k,i+1})a_k - \frac{p-1}{p} \cdot 2n\sigma^2.$$

Our next result follows from an exponential inequality for chi-square random variables proved by Laurent and Massart (2000).

Lemma 11. *Let $(Y_1, ..., Y_n)$ be i.i.d. Gaussian variables with mean 0 and variance 1. Let $a_1, ..., a_n$ be nonnegative. We set $\|a\|_\infty = \sup_{1 \le i \le n}|a_i|, \|a\|_2^2 = \sum_{i=1}^{n}a_i^2$. Let $Z = \sum_{i=1}^{n}a_i(Y_i^2 - 1)$. Then the following inequalities hold for any positive x:*

$$P(Z \ge 2\|a\|_2\sqrt{x} + 2\|a\|_\infty x) \le \exp(-x), \tag{3.1}$$

$$P(Z \le -2\|a\|_2\sqrt{x}) \le \exp(-x). \tag{3.2}$$

By choosing $x = c\log p$, we have, with probability at least $1 - \frac{1}{p^c}$,

$$\|\phi_{i,i+1}\|_2^2 \le c\sigma^2\sqrt{n\log p} + c\sigma^2\log p - \|a\|_2^2(\eta_i - \eta_{i+1})^2 + 2(\eta_i - \eta_{i+1})\sum_{k=1}^{n}(T_{ki} - T_{k,i+1})a_k + \frac{p-1}{p} \cdot 2n\sigma^2$$

$$\le c\sigma^2(\sqrt{n} + \sqrt{\log p})^2 - \|a\|_2^2(\eta_i - \eta_{i+1})^2 + 2(\eta_i - \eta_{i+1})\sum_{k=1}^{n}(T_{ki} - T_{k,i+1})a_k \tag{3.3}$$

Note that the term

$$(\eta_i - \eta_{i+1})\sum_{k=1}^{n}(T_{ki} - T_{k,i+1})a_k \sim N\left(\|a\|_2^2(\eta_i - \eta_{i+1})^2, 2\|a\|_2^2(\eta_i - \eta_{i+1})^2 \cdot \frac{p-1}{p}\sigma^2\right)$$

It follows from the tail bound of standard Gaussian distribution $\Phi(-x) \le \frac{1}{x}\phi(x)$ that

$$P\left((\eta_i - \eta_{i+1})\sum_{k=1}^{n}(T_{ki} - T_{k,i+1})a_k \ge \|a\|_2^2(\eta_i - \eta_{i+1})^2 + 2\sigma\sqrt{c\log p} \cdot \|a\|_2|\eta_i - \eta_{i+1}|\right) \le \frac{1}{p^c}. \tag{3.4}$$

Combining (3.3) and (3.4), we have

$$P\left(\|\phi_{i,i+1}\|_2^2 \le c'\sigma^2(\sqrt{n} + \sqrt{\log p})^2 + \|a\|_2^2(\eta_i - \eta_{i+1})^2 + c'\sigma\sqrt{\log p}\|a\|_2|\eta_i - \eta_{i+1}|\right) \ge 1 - \frac{C}{p^c}$$

for some constant $C, c, c' > 0$. Thus,

$$P\left(w^\top\phi_{i,i+1} \ge -\sqrt{2\delta}\|\phi_{i,i+1}\|\right)$$

$$\le \Phi\left(\frac{C\sqrt{\delta}}{\sigma}\left[\sigma^2(\sqrt{n} + \sqrt{\log p})^2 + \|a\|_2^2(\eta_i - \eta_{i+1})^2 + \|a\|_2|\eta_i - \eta_{i+1}|\sigma\sqrt{\log p}\right]^{1/2} + \frac{\|a\|_2(\eta_i - \eta_{i+1})}{\sigma}\right) + \frac{C}{p^c}$$

$$\square$$

Proof of Lemma 2. The proof follows from the following deterministic perturbation bound and concentration inequalities adapted from Cai and Zhang (2018).

Lemma 12. *Suppose $A \in \mathbb{R}^{n \times p}$, $\tilde{V} = \begin{bmatrix} V & V_\perp \end{bmatrix} \in \mathbb{O}_n$ are left singular vectors of A, $V \in \mathbb{O}_{n,r}, V_\perp \in \mathbb{O}_{n,n-r}$ correspond to the first r and last $(n-r)$ singular vectors respectively. $\tilde{W} = \begin{bmatrix} W & W_\perp \end{bmatrix} \in \mathbb{O}_n$ is any orthogonal matrix with $W \in \mathbb{O}_{n,r}$, $W_\perp \in \mathbb{O}_{n,n-r}$. Given that $\lambda_r(W^\top A) > \lambda_{r+1}(A)$, we have*

$$\|\sin\Theta(V,W)\| \le \frac{\lambda_r(W^\top A)\|\mathbb{P}_{A^\top W}A^\top W_\perp\|^2}{\lambda_r^2(W^\top A) - \lambda_{r+1}^2(A)} \wedge 1.$$

Lemma 13. *Suppose $X \in \mathbb{R}^{n \times p}$ is a rank-r matrix with left singular space as $V \in \mathbb{O}_{n,r}, Z \in \mathbb{R}^{n \times p}$, Z is a i.i.d. sub-Gaussian random matrix with sub-Gaussian parameter τ. $Y = X + Z$. Then there exists constants C, c such that for any $x > 0$,*

$$P\left(\lambda_r^2(V^\top Y) \le (\lambda_r^2(X) + \tau^2 p)(1-x)\right) \le C\exp\left\{Cr - c\tau^{-2}(\lambda_r^2(X) + \tau^2 p)(x^2 \wedge x)\right\}, \qquad (3.5)$$

$$P\left(\lambda_{r+1}^2(Y) \ge \tau^2 p(1+x)\right) \le C\exp\left\{Cn - cp \cdot (x^2 \wedge x)\right\} \qquad (3.6)$$

Moreover, there exists C_0, C, c such that whenever $\lambda_r^2(X) \ge C_0\tau^2 n$, for any $x > 0$ we have

$$P(\|\mathbb{P}_{V^\top Y}V_\perp^\top Y\| \le x) \ge 1 - C\exp\left\{Cn - c\tau^{-2}\min(x^2, x\sqrt{\lambda_r^2(X) + \tau^2 p})\right\} - C\exp\left\{-c\tau^{-2}(\lambda_r^2(X) + \tau^2 p)\right\}.$$
$$(3.7)$$

Note that $T = \Theta' + E$ where $\Theta' = \|a\| \cdot w\eta'^\top$ is rank 1, and $\hat{w}$ is the first left singular vector of T, by Lemma 12, we have

$$|1 - (\hat{w}^\top w)^2| = \|\sin\Theta(w, \hat{w})\|^2 \le \frac{\lambda_1^2(w^\top T)\|\mathbb{P}_{w^\top T}w_\perp^\top T\|^2}{(\lambda_1^2(w^\top T) - \lambda_2^2(T))^2} \wedge 1. \qquad (3.8)$$

To bound the quantity $\lambda_1^2(w^\top T)$, using Lemma 13, by choosing $x = \frac{\lambda^2}{3(\lambda^2+\sigma^2 p)}$ in (3.5), since $\lambda^2 \geq C\sigma^2(n + \sqrt{pn})$ for some sufficiently large C, or $\sigma^2(\lambda^2 + \sigma^2 p)(\sigma^2 n + \log p)/\lambda^4 \leq 1$, we have

$$c\sigma^{-2}\min\left\{\frac{\lambda^4}{\lambda^2+\sigma^2 p}, \lambda^2\right\} - C \geq \frac{c_0\sigma^{-2}\lambda^4}{\lambda^2+\sigma^2 p} - C \geq c_1\log p.$$

Thus

$$P\left(\lambda_1^2(w^\top T) \leq \frac{2\lambda^2}{3} + \sigma^2 p\right) \leq C\exp\left\{C - \frac{c_0\sigma^{-2}\lambda^4}{\lambda^2 + p}\right\} \leq C\exp\{-c_1\log p\}. \tag{3.9}$$

To bound $\lambda_2^2(T)$, if $p \leq n + \log p$, by choosing $x = \frac{n+\log p}{p}$ in (3.6), we have

$$c\min\left\{\frac{(n+\log p)^2}{p}, n+\log p\right\} - Cn \geq c\log p$$

for some $C, c > 0$, so that

$$P\left(\lambda_2^2(T) \geq \sigma^2(p + n + \log p)\right) \leq C\exp\left\{Cn - cp\cdot x^2 \wedge x\right\} \leq C\exp\{-c\log p\} = \frac{C}{p^c}. \tag{3.10}$$

If $p > n + \log p$, by choosing $x = \sqrt{\frac{n+\log p}{p}}$ in (3.6), we have

$$c\min\left\{\sqrt{p(n+\log p)}, n+\log p\right\} - Cn \geq c\log p$$

for some $C, c > 0$, so that

$$P\left(\lambda_2^2(T) \geq \sigma^2(p + \sqrt{p(n+\log p)})\right) \leq C\exp\left\{Cn - cp\cdot x^2 \wedge x\right\} \leq C\exp\{-c\log p\} = \frac{C}{p^c}. \tag{3.11}$$

Lastly, to bound $\|\mathbb{P}_{w^\top T}w_\perp^\top T\|$, choosing $x = \sigma\sqrt{n + \log p}$ in (3.7), since $\lambda^2 \geq \sigma^2 n$, we have

$$\sigma^2(n + \log p) \leq c(\lambda^2 + \sigma^2 p)$$

for sufficiently large $c > 0$ and therefore

$$c\sigma^{-2}\min\left\{x^2, x\sqrt{\lambda^2 + \sigma^2 p}\right\} - Cn \geq c\log p.$$

So that

$$P(\|\mathbb{P}_{w^\top T}w_\perp^\top T\| \leq \sigma\sqrt{n+\log p}) \geq 1 - \frac{C}{p^c} - C\exp\left\{-c\sigma^{-2}(\lambda^2 + \sigma^2 p)\right\}. \tag{3.12}$$

Combining (3.9) (3.10) (3.11) and (3.12), by the fact that $x/(x-y)^2$ is decreasing in x and increasing

in y for $x > y > 0$, we have

$$P\left(\frac{\lambda^2(w^\top T)\|\mathbb{P}_{w^\top T}w_\perp^\top T\|^2}{(\lambda_r^2(w^\top T) - \lambda_2^2(T))^2} \leq C'\frac{\sigma^2(\lambda^2 + \sigma^2 p)(n + \log p)}{\lambda^4}\right) \geq 1 - Cp^{-c}.$$

$\square$

Proof of Lemma 7. Note that

$$\phi_{i,i+1}^\top w = \sum_{k=1}^n w_k(T_{ki} - T_{k,i+1}) \sim N\left(w^\top(\Theta'_{\cdot i} - \Theta'_{\cdot i+1}), \frac{2(p-1)\sigma^2}{p}\right),$$

and

$$\|\phi_{i,i+1}\|_2^2 = \sum_{k=1}^n (T_{ki} - T_{k,i+1})^2, \quad \text{where} \quad T_{ki} - T_{k,i+1} \sim N(\theta'_{ki} - \theta'_{k,i+1}, 2(p-1)\sigma^2/p).$$

We construct the standardized chi-square statistic

$$Y_k^2 = \left[\frac{T_{k,i} - T_{k,i+1} - (\theta'_{ki} - \theta'_{k,i+1})}{\sigma\sqrt{2(p-1)/p}}\right]^2.$$

Define

$$Q = \frac{p-1}{p}2\sigma^2 \cdot \sum_{k=1}^n (Y_k^2 - 1),$$

we have

$$Q = \|\phi_{i,i+1}\|_2^2 + \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 - 2\sum_{k=1}^n (T_{ki} - T_{k,i+1})(\theta'_{ki} - \theta'_{k,i+1}) - \frac{p-1}{p} \cdot 2n\sigma^2.$$

By Lemma 11, choosing $x = c\log p$, we have, with probability at least $1 - \frac{1}{p^c}$,

$$\|\phi_{i,i+1}\|_2^2 \leq c\sigma^2\sqrt{n\log p} + c\sigma^2\log p - \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 + 2\sum_{k=1}^n (T_{ki} - T_{k,i+1})(\theta'_{ki} - \theta'_{k,i+1}) + \frac{p-1}{p} \cdot 2n\sigma^2$$

$$\leq c\sigma^2(\sqrt{n} + \sqrt{\log p})^2 - \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 + 2\sum_{k=1}^n (T_{ki} - T_{k,i+1})(\theta'_{ki} - \theta'_{k,i+1}). \tag{3.13}$$

Note that the term

$$\sum_{k=1}^n (T_{ki} - T_{k,i+1})(\theta'_{ki} - \theta'_{k,i+1}) \sim N\left(\|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2, 2\|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 \cdot \frac{p-1}{p}\sigma^2\right)$$

It follows from the tail bound of standard Gaussian distribution $\Phi(-x) \le \frac{1}{x}\phi(x)$ that

$$P\left(\sum_{k=1}^{n} (T_{ki} - T_{k,i+1})(\theta'_{ki} - \theta'_{k,i+1}) \ge \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 + 2\sigma\sqrt{c\log p} \cdot \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2 \right) \le \frac{1}{p^c}. \quad (3.14)$$

Combining (3.13) and (3.14), we have

$$P\left(\|\phi_{i,i+1}\|_2^2 \le c'\sigma^2(\sqrt{n} + \sqrt{\log p})^2 + \|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2 + c'\sigma\sqrt{\log p}\|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2 \right) \ge 1 - \frac{C}{p^c}$$

for some constant $C, c, c' > 0$. Thus,

$$P\left(w^\top \phi_{i,i+1} \ge -\sqrt{2\delta}\|\phi_{i,i+1}\| \right)$$

$$\le \Phi\left(C\sqrt{\delta}\left[(\sqrt{n} + \sqrt{\log p})^2 + \frac{\|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2^2}{\sigma^2} + c'\sqrt{\log p}\frac{\|\Theta'_{\cdot i} - \Theta'_{\cdot i+1}\|_2}{\sigma} \right]^{1/2} + \frac{w^\top (\Theta'_{\cdot i} - \Theta'_{\cdot i+1})}{\sigma} \right) + \frac{C}{p^c}$$

$\square$

Proof of Lemma 8. The proof of Lemma 8 depends on Lemma 12 as well as the following concentration inequalities.

Lemma 14. *Suppose $X \in \mathbb{R}^{n\times p}$ is a rank-r matrix with first left singular vector $v_1 \in \mathbb{R}^n$, Z is a i.i.d. sub-Gaussian random matrix with sub-Gaussian parameter τ. $Y = X + Z$. Then there exists constants C, c such that for any $x > 0$, $\lambda_1^2(X) \ge 3\sqrt{\log(n \vee p)}$*

$$P\left(\lambda_1^2(v_1^\top Y) \le \lambda_1^2(X) + \tau^2(p - \sqrt{p\log(p\vee n)} - \log(p \vee n)) \right) \le \frac{C}{p^c}, \quad (3.15)$$

$$P(\lambda_2^2(Y) > (\lambda_2^2(X) + \tau^2 p)(1+t)) \le C\exp(Cn - c\tau^{-2}(\lambda_n^2 + \tau^2 p)t^2 \wedge t). \quad (3.16)$$

Moreover, there exists C_0, C, c such that for any $x > 0$ we have

$$P(\|\mathbb{P}_{Y^\top v_1}Y^\top v_{1\perp}\| < t) \ge 1 - C\exp(Cn - c\tau^{-2}\min(t^2, t\sqrt{\lambda_1^2(X) + \tau^2 p})) - C\exp\left\{ -c\tau^{-2}(\lambda_1^2(X) + \tau^2 p) \right\}. \quad (3.17)$$

Following a similar argument as the proof of Lemma 2, by taking $t = \sigma\sqrt{\frac{n+\log p}{\lambda_n^2 + \sigma^2 p}}$ in (3.16), since $n \lesssim p$, we have

$$P(\lambda_2^2(Y) > \lambda_2^2(\Theta') + \tau^2 p) \le \frac{C}{p^c}.$$

Similarly, by taking $t = \sigma\sqrt{n + \log p}$, we have

$$P(\|\mathbb{P}_{Y^\top w}Y^\top w_\perp\|^2 < \sigma^2(n + \log p)) \ge 1 - \frac{C}{p^c}.$$

Then Lemma 12 implies, if $\lambda_1^2(\Theta') - \lambda_2^2(\Theta') \geq C\sigma^2(n + \sqrt{np})$ for some $C > 0$,

$$P\left(|1 - (\hat{w}^\top w)^2| \leq \frac{C\sigma^2(\lambda_1^2(\Theta') + \sigma^2 p)(n + \log p)}{(\lambda_1^2(\Theta') - \lambda_2^2(\Theta'))^2} \right) \geq 1 - \frac{C}{p^c}.$$

□

Proof of Lemma 10. By SVD of $\Theta' = (\theta'_{ij}) \in \mathbb{R}^{n \times p}$, its first left singular vector

$$u_1' = \arg\max_{\|x\|_2=1} x^\top \Theta' \Theta'^\top x = \arg\max_{\|x\|_2=1} \sum_{j=1}^p \left(\sum_{i=1}^n x_i \theta'_{ij} \right)^2. \tag{3.18}$$

In the following, we show that, for any unit vector $x \in \mathbb{R}^n$,

$$\sum_{j=1}^p \left(\sum_{i=1}^n x_i \theta'_{ij} \right)^2 \leq \sum_{j=1}^p \left(\sum_{i=1}^n |x_i|\theta'_{ij} \right)^2, \tag{3.19}$$

from which we conclude that u_1' is either nonpositive or nonnegative. Toward this end, note that

$$\sum_{j=1}^p \left(\sum_{i=1}^n x_i \theta'_{ij} \right)^2 = \sum_{j=1}^p \sum_{i=1}^n x_i^2 \theta'^2_{ij} + \sum_{i \neq k} x_i x_k \left(\sum_{j=1}^p \theta'_{ij} \theta'_{kj} \right).$$

Then the inequality (3.19) follows from

$$\sum_{i \neq k} x_i x_k \left(\sum_{j=1}^p \theta'_{ij} \theta'_{kj} \right) \leq \sum_{i \neq k} |x_i x_k| \left(\sum_{j=1}^p \theta'_{ij} \theta'_{kj} \right),$$

which is true as long as $\sum_{j=1}^p \theta'_{ij} \theta'_{kj} \geq 0$ for all pairs $i \neq k$. We conclude this proof by showing that

Fact. For any nondecreasing vectors $a, b \in \mathbb{R}^n$, such that $\sum_{i=1}^n a_i = 0$. Then it follows that $a^\top b \geq 0$.

To see this, note that since a and b are both nondecreasing, there exist a constant δ such that the components of $b + \delta \cdot \mathbf{1}$ has the same sign as a. Hence the claim follows from $0 \leq a^\top(b + \delta \cdot \mathbf{1}) = a^\top b + \delta a^\top \mathbf{1} = a^\top b$. □

Proof of Lemma 14. Note that

$$\lambda_1^2(v_1^\top Y) = \|v_1^\top Y\|_2^2 = \lambda_1^2(X) + 2\lambda_1(X)v_1^\top Z u_1 + \|v_1^\top Z\|_2^2.$$

The linear term $v_1^\top Z u_1 = \sum_{i,j} v_{1i} u_{1j} Z_{ij}$ is subgaussian with parameter $c\|v_1 u_1^\top\|_F = c$ for some constant c only depending on τ. Therefore, by concentration inequality for subgaussian random

variables
$$P(v_1^\top Z u_1 \geq t) \leq \exp\left\{ - \frac{t^2}{2c\tau^2} \right\}. \tag{3.20}$$

On the other hand, the quadratic term $\|v_1^\top Z\|_2^2 = \sum_{i=1}^{p}(v_1^\top Z_i)^2$ where Z_i for $i = 1, ..., p$ are the i-th column of Z. Each of $(v_1^\top Z_i)^2$ is subexponential with mean

$$\mathbb{E}(v_1^\top Z_i)^2 = \mathbb{E}Z_{ij}^2.$$

Then concentration inequality for subexponential random variables yields

$$P\left(\left| \sum_{i=1}^{p}(v_1^\top Z_i)^2 - p\mathbb{E}Z_{ij}^2 \right| \geq t \right) \leq 2\exp\left\{ - c\min\left(\frac{t^2}{\tau^4 p}, \frac{t}{\tau^2} \right) \right\} \tag{3.21}$$

for some constant $c > 0$. By taking $t = \tau\sqrt{\log(p \vee n)}$ in (3.20), we have

$$P(v_1^\top Z u_1 \geq \tau\sqrt{\log(p \vee n)}) \leq \frac{C}{p^c}.$$

Taking $t = \tau^2\sqrt{p\log(p \vee n)}$ if $\sqrt{\log(p \vee n)} \leq \sqrt{p}$ and $t = \tau^2\log(p \vee n)$ otherwise in (3.21), we have

$$P\left(\sum_{i=1}^{p}(v_1^\top Z_i)^2 \leq \tau^2 p - \tau^2\sqrt{p\log(p \vee n)} - \tau^2\log(p \vee n) \right) \leq \frac{C}{p^c}.$$

Combining these, we have, with probability at least $1 - O(p^{-c})$,

$$\lambda_1^2(v_1^\top Y) = \lambda_1^2(X) + 2\tau\lambda_1(X)v_1^\top Z u_1 + \|v_1^\top Z\|_2^2$$
$$\geq \lambda_1^2(X) - 2\tau\lambda_1(X)\sqrt{\log(p \vee n)} + \tau^2 p - \tau^2\sqrt{p\log(p \vee n)} - \tau^2\log(p \vee n).$$

If in addition $\lambda_1(X) \geq 3\tau\sqrt{\log(n \vee p)}$, then

$$P\left(\lambda_1^2(v_1^\top Y) \leq \lambda_1^2(X) + \tau^2 p - \tau^2\sqrt{p\log(p \vee n)} - \tau^2\log(p \vee n) \right) \leq \frac{C}{p^c}.$$

For $\lambda_2^2(Y)$, note that

$$\lambda_2(Y) = \min_{\text{rank}(B)\leq 1} \|Y - B\| \leq \|Y - [v_1 \quad \mathbf{0}]^\top \cdot Y\| = \|v_{1\perp}^\top Y\|$$

for $v_{1\perp} \in \mathbb{R}^{n\times(n-1)}$. It suffices to obtain an upper bound for $\lambda_1(v_{1\perp}^\top Y)$ with high probability. Next,

$$
\begin{aligned}
\|v_{1\perp}^\top Y\|_2^2 &= \|v_{1\perp}^\top Y Y^\top v_{1\perp}\| \\
&\le \|\mathbb{E}v_{1\perp}^\top Y Y^\top v_{1\perp}\| + \|v_{1\perp}^\top Y Y^\top v_{1\perp} - \mathbb{E}v_{1\perp}^\top Y Y^\top v_{1\perp}\| \\
&= \lambda_2^2(X) + \tau^2 p + \|v_{1\perp}^\top Y Y^\top v_{1\perp} - \mathbb{E}v_{1\perp}^\top Y Y^\top v_{1\perp}\|
\end{aligned}
\tag{3.22}
$$

Define the normalization matrix $M \in \mathbb{R}^{(n-1)\times(n-1)}$ as

$$
M = \begin{bmatrix} (\lambda_2^2(X)+\tau^2 p)^{-1/2} & & & & \\ & \ddots & & & \\ & & (\lambda_r^2(X)+\tau^2 p)^{-1/2} & & \\ & & & \ddots & \\ & & & & (\lambda_n^2(X)+\tau^2 p)^{-1/2} \end{bmatrix},
$$

we have

$$
M^\top v_{1\perp}^\top \mathbb{E}Y Y^\top v_{1\perp} M = I_{n-1}.
$$

Let $Q = v_{1\perp}^\top Y Y^\top v_{1\perp} - \mathbb{E}v_{1\perp}^\top Y Y^\top v_{1\perp}$, we have

$$
\|Q\| = \|(M^{-1})^\top M^\top Q M M^{-1}\| \le \|M^\top Q M\|\|M^{-1}\|^2.
$$

By construction we have

$$
\|M^{-1}\| = (\lambda_2^2(X)+\tau^2 p)^{1/2},
$$

then

$$
\|Q\| \le (\lambda_2^2(X)+\tau^2 p)\|M^\top v_{1\perp}^\top Y Y^\top v_{1\perp} M - M^\top v_{1\perp}^\top \mathbb{E}Y Y^\top v_{1\perp} M\|.
$$

Now it suffices to obtain a concentration inequality for $\|M^\top v_{1\perp}^\top Y Y^\top v_{1\perp} M - I_{n-1}\|$. The main idea is to use the ϵ-net argument to split the spectral norm deviation to the deviations of single random variables, which can be further controlled by the Hanson-Wright inequality. Specifically, for any unit vector $u \in \mathbb{R}^{n-1}$, we have

$$
\begin{aligned}
&u^\top M^\top v_{1\perp}^\top Y Y^\top v_{1\perp} M u - u^\top I_{n-1} u \\
&= (u^\top M^\top v_{1\perp}^\top X X^\top v_{1\perp} M u - \mathbb{E}u^\top M^\top v_{1\perp}^\top X X^\top v_{1\perp} M u) \\
&\quad + 2(u^\top M^\top v_{1\perp}^\top X Z^\top v_{1\perp} M u - \mathbb{E}u^\top M^\top v_{1\perp}^\top X Z^\top v_{1\perp} M u) \\
&\quad + (u^\top M^\top v_{1\perp}^\top Z Z^\top v_{1\perp} M u - \mathbb{E}u^\top M^\top v_{1\perp}^\top Z Z^\top v_{1\perp} M u) \\
&= 2(X^\top v_{1\perp} M u)^\top Z^\top(v_{1\perp} M u) + (v_{1\perp} M u)^\top(Z Z^\top - \tau^2 p I_n)(v_{1\perp} M u).
\end{aligned}
$$

In the following, we shall bound the two terms separately.

To bound the second term, for any fixed unit vector $u \in \mathbb{R}^{n-1}$, we vectorize $Z \in \mathbb{R}^{n \times p}$ into $\text{vec}(Z) \in \mathbb{R}^{np}$ as

$$\text{vec}(Z) = (Z_{11}, Z_{21}, ..., Z_{n1}, Z_{12}, ..., Z_{n2}, ..., Z_{1p}, Z_{np})^\top.$$

We also define the block diagonal matrix

$$D = \begin{bmatrix} (v_{1\perp}Mu)(v_{1\perp}Mu)^\top & & \\ & \ddots & \\ & & (v_{1\perp}Mu)(v_{1\perp}Mu)^\top \end{bmatrix} \in \mathbb{R}^{np \times np}.$$

It then follows that

$$(v_{1\perp}Mu)^\top (ZZ^\top - \tau^2 p I_n)(v_{1\perp}Mu) = \text{vec}(Z)^\top D \cdot \text{vec}(Z) - \mathbb{E}\text{vec}(Z)^\top D \cdot \text{vec}(Z).$$

Besides,

$$\|D\| = \|(v_{1\perp}Mu)(v_{1\perp}Mu)^\top\| = \|Mu\|_2^2 \leq \|M\|^2 = (\lambda_n^2(X) + \tau^2 p)^{-1},$$

$$\|D\|_F^2 = p\|(v_{1\perp}Mu)(v_{1\perp}Mu)^\top\|_F^2 \leq p\|Mu\|_2^4 \leq p(\lambda_n^2(X) + \tau^2 p)^{-2}.$$

By Hansen-Wright inequality (Rudelson and Vershynin, 2013),

$$\begin{aligned} &P(|(v_{1\perp}Mu)^\top(ZZ^\top - \tau^2 p I_n)(v_{1\perp}Mu)| > t) \\ &= P(|\text{vec}(Z)^\top D \cdot \text{vec}(Z) - \mathbb{E}\text{vec}(Z)^\top D \cdot \text{vec}(Z)| > t) \\ &\leq 2\exp\left\{ -c\min\left(\frac{t^2(\lambda_n^2(X)+\tau^2 p)^2}{\tau^4 p}, \frac{t(\lambda_n^2(X)+\tau^2 p)}{\tau^2} \right) \right\} \end{aligned} \qquad (3.23)$$

for some $c > 0$.

Next, we bound the first term

$$\begin{aligned} (X^\top v_{1\perp}Mu)^\top Z^\top(v_{1\perp}Mu) &= \text{tr}(Z^\top(v_{1\perp}Mu)(X^\top v_{1\perp}Mu)^\top) \\ &= \text{vec}(Z)^\top \cdot \text{vec}((v_{1\perp}Mu)(X^\top v_{1\perp}Mu)^\top). \end{aligned}$$

Since

$$X^\top v_{1\perp}M = U \begin{bmatrix} 0 & \cdots & & 0 & 0 \\ \lambda_2(X)(\lambda_2^2(X)+\tau^2 p)^{-1/2} & & & & 0 \\ & & \ddots & & \vdots \\ & & & \lambda_n(X)(\lambda_n^2(X)+\tau^2 p)^{-1/2} & 0 \end{bmatrix},$$

we know $\|X^\top v_{1\perp} M\| \le 1$ and

$$\begin{aligned}
\|\mathrm{vec}((v_{1\perp} Mu)(X^\top v_{1\perp} Mu)^\top)\|_2^2 &= \|(v_{1\perp} Mu)(X^\top v_{1\perp} Mu)^\top\|_F^2 \\
&= \|v_{1\perp} Mu\|_2^2 \cdot \|X^\top v_{1\perp} Mu\|_2^2 \\
&\le \|M\|^2 \le (\lambda_n^2 + \tau^2 p)^{-1}.
\end{aligned}$$

By the concentration inequality for i.i.d. subgaussian random variables, we have

$$P(|(X^\top v_{1\perp} Mu)^\top Z^\top (v_{1\perp} Mu)| > t) \le C \exp\left(- c\frac{t^2(\lambda_n^2 + \tau^2 p)}{\tau^2} \right) \tag{3.24}$$

for some constant $C, c > 0$. Combining (3.24) and (3.23), we have, for any fixed unitary $u \in \mathbb{R}^{n-1}$,

$$P(|u^\top M^\top v_{1\perp}^\top YY^\top v_{1\perp} Mu - u^\top I_{n-1} u| > t) \le C \exp(-c\tau^{-2}(\lambda_n^2 + \tau^2 p)t^2 \wedge t)$$

for all $t > 0$. Next, we use the following lemma proved by Cai and Zhang (2018) concerning the ϵ-net argument for unit ball.

Lemma 15. *For any $p \ge 1$, denote $\mathbb{B}^p = \{x \in \mathbb{R}^p : \|x\|_2 \le 1\}$ as the p-dimensional unit ball in the Euclidean space. Suppose $K \in \mathbb{R}^{p_1 \times p_2}$ is a random matrix. Then we have for $t > 0$,*

$$P(\|K\| \ge 3t) \le 7^{p_1+p_2} \cdot \max_{u \in \mathbb{B}^{p_1}, v \in \mathbb{B}^{p_2}} P(|u^\top Kv| \ge t).$$

It then follows that

$$P(\|M^\top v_{1\perp}^\top YY^\top v_{1\perp} M - I_{n-1}\| > t) \le C \exp(Cn - c\tau^{-2}(\lambda_n^2 + \tau^2 p)t^2 \wedge t). \tag{3.25}$$

Recall (3.22), we have

$$P(\lambda_2^2(Y) > \lambda_2^2(X) + \tau^2 p + (\lambda_2^2(X) + \tau^2 p)t) \le C \exp(Cn - \tau^{-2}(\lambda_n^2 + \tau^2 p)t^2 \wedge t).$$

Finally, we consider $\|\mathbb{P}_{v_1^\top Y} v_{1\perp}^\top Y\|$. Define the constant

$$m = (\lambda_1^2(X) + \tau^2 p)^{-1/2}.$$

It then follows that

$$m^2 v_1^\top \mathbb{E} YY^\top v_1 = 1.$$

Since

$$\|\mathbb{P}_{Y^\top v_1} Y^\top v_{1\perp}\| = \|\mathbb{P}_{mY^\top v_1} Y^\top v_{1\perp}\|$$
$$= \|mY^\top v_1((mY^\top v_1)^\top(mY^\top v_1))^{-1}(mY^\top v_1)^\top Y^\top v_{1\perp}\|$$
$$\leq \|mY^\top v_1\|_2^{-1}\|(mY^\top v_1)^\top Y^\top v_{1\perp}\|.$$

In the following we analyze $\|mY^\top v_1\|_2$ and $\|(mY^\top v_1)^\top Y^\top v_{1\perp}\|$ separately.

Since

$$\|mY^\top v_1\|^2 = m^2|v_1^\top YY^\top v_1| \geq 1 - |m^2 v_1^\top YY^\top v_1 - m^2 v_1^\top \mathbb{E}YY^\top v_1|,$$

following the same argument that leads to (3.25), we have

$$P(|m^2 v_1^\top YY^\top v_1| > 1 - t) \geq 1 - C\exp(C - c\tau^{-2}(\lambda_1^2(X) + \tau^2 p)t^2 \wedge t).$$

Now set $t = 1/2$, we can choose C_0 large enough, such that $\lambda_1^2(X) \geq \tau^2 C_0$ and therefore $C - c\tau^{-2}(\lambda_1^2(X) + \tau^2 p) \leq -c'\tau^{-2}(\lambda_1^2(X) + \tau^2 p)$ for some $c, c' > 0$. In this case,

$$P(|m^2 v_1^\top YY^\top v_1| > 1/2) \geq 1 - C\exp(-c\tau^{-2}(\lambda_1^2(X) + \tau^2 p)).$$

For $\|(mY^\top v_1)^\top Y^\top v_{1\perp}\|$, note that $mv_1^\top XX^\top v_{1\perp} = 0$, we have

$$(mY^\top v_1)^\top Y^\top v_{1\perp} = mv_1^\top(X + Z)(X + Z)^\top v_{1\perp}$$
$$= mv_1^\top XZ^\top v_{1\perp} + mv_1^\top ZX^\top v_{1\perp} + mv_1^\top ZZ^\top v_{1\perp}.$$

Following similar idea of the proof of (3.25), we can show for any unit vector $u \in \mathbb{R}^{n-1}$,

$$P(|mv_1^\top XZ^\top v_{1\perp} u| > t) \leq C\exp\left(\frac{-ct^2}{\tau^2\|(v_{1\perp}u)(mv_1^\top X)\|_F^2}\right) \leq C\exp(-c\tau^{-2}t^2),$$

$$P(|mv_1^\top ZZ^\top v_{1\perp} u| > t) = P(|mv_1^\top(ZZ^\top - \tau^2 pI_n)v_{1\perp}u| > t) \leq C\exp(-c\tau^{-2}\min(t^2, \sqrt{\lambda_1^2(X) + \tau^2 p}t)).$$

By the ϵ-net argument again (Lemma 15), we have

$$P(\|(mY^\top v_1)^\top Y^\top v_{1\perp}\| > t) \leq C\exp(Cn - c\tau^{-2}\min(t^2, t\sqrt{\lambda_1^2(X) + \tau^2 p})).$$

$\square$

A   Comparison with an SVD-based Esitmator

In this section, we compare the empirical performance of our proposed estimator $\hat{\pi}$ to that of an alternative estimator $\tilde{\pi} = (\mathfrak{r}(\hat{v}_1))^{-1}$ where $\hat{v}_1$ is the first right singular vector of Y. This estimator

is closely related to $\hat{\pi}$ except that it does not centralize the rows in Y before estimating its singular subspaces. However, this normalization step is essential in order for the resulting estimator to be invariant to the unknown intercepts of the growth models. The signal matrix $\Theta = (\theta_{ij}) \in \mathbb{R}^{n \times p}$ is generated under the following two regimes:

- $S_1(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = \log(1 + \alpha j + \beta_i)$ for $1 \leq i \leq n/2$ where as $\theta_{ij} = 0$ for $n/2 < i \leq n$, $\beta_i \sim \mathrm{Unif}(1, 3)$ for all $1 \leq i \leq n$;

- $S_2(\alpha, n, p)$: For any $1 \leq j \leq p$, $\theta_{ij} = \alpha j + \beta_i$ for $i = 1$ where as $\theta_{ij} = 0$ for $2 \leq i \leq n$, $\beta_i \sim \mathrm{Unif}(1, 3)$ for all $1 \leq i \leq n$;

In each setting, we evaluate the empirical performance of each method over a range of n, p and α. Each setting is repeated for 200 times. The empirical normalized Kendall's tau is reported using boxplots, as shown in Figure 1 of our supplementary material. From Figure 6, our proposed estimator $\hat{\pi}$ performs better than $\tilde{\pi}$ in all the settings, especially in $S_2(\alpha, n, p)$ where the signals are concentrated at one row.

B Supplementary Figures and Tables

In Figure 7, the graphical representation of the weight vectors $\hat{w}$ for our proposed estimator $\hat{\pi}$, and the pseudo-weight vector $\tilde{w}$ for the estimator π_{max} based on 200 simulations under four different models in Section 6.1 of our main paper is given.

In Figure 8, the Taxonomic tree of 45 closely related species used in generating the shotgun metagenomic data used on in s Gao and Li (2018) as well as Section 6.2 of our main paper is given.

Table S.1 lists the p-values of 8 contig clusters from the Wilcoxon rank sum test of the ePTRs between the responser and non-responser groups, and the taxonomic annotations with lineage scores indicating the quality of each taxonomic classification (see Section 6.3 of our main paper).
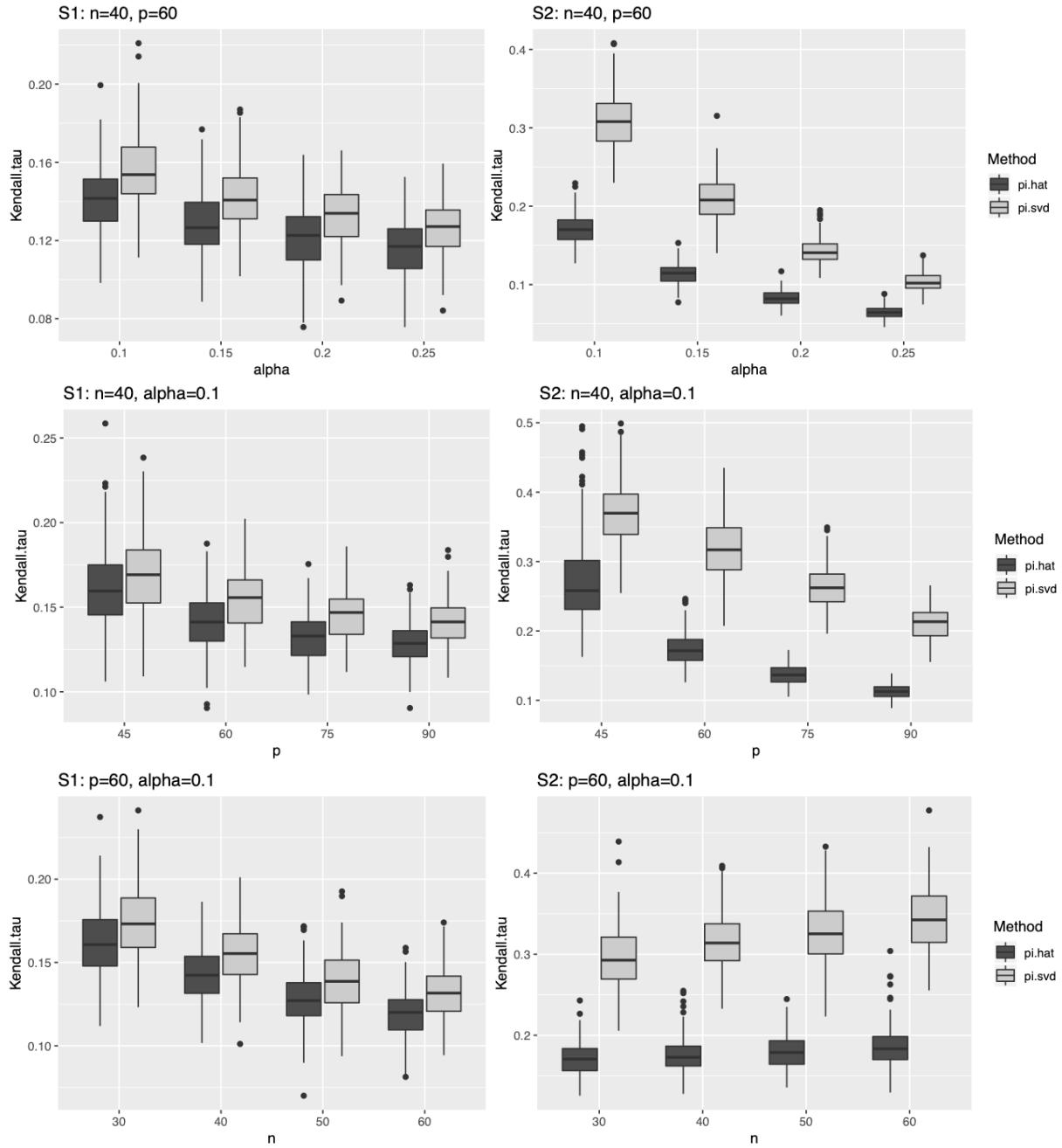
Figure 6: Boxplots of the empirical normalized Kendall's distance between the estimated permutation and true permutation under models $S_1(\alpha, p, n)$ and $S_2(\alpha, p, n)$. $\hat{\pi}$: proposed estimator; π_{svd}: estimator based on SVD.
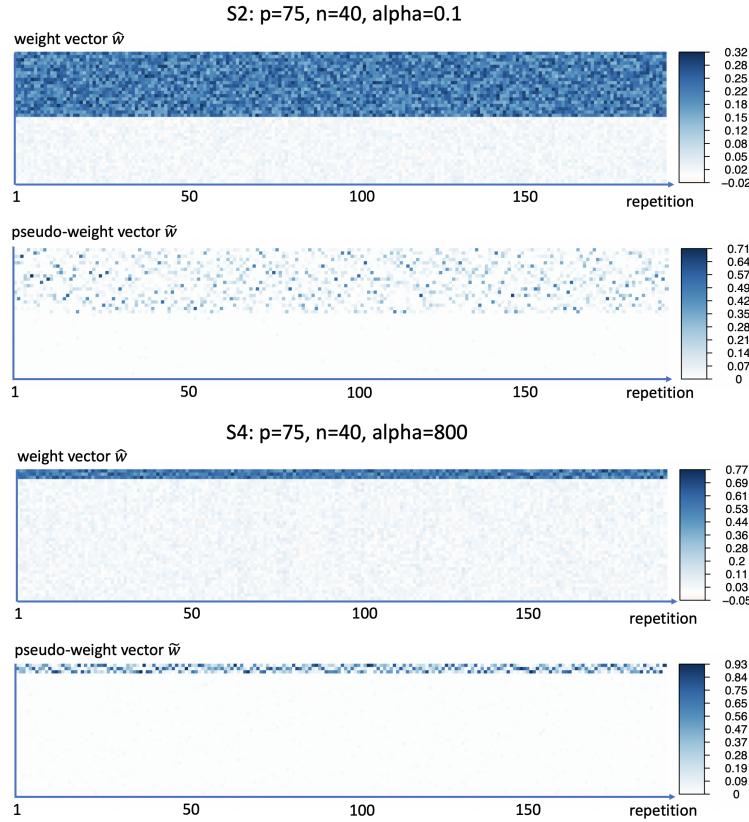
Figure 7: The graphical representation of the weight vectors $\hat{w}$ for our proposed estimator $\hat{\pi}$, and the pseudo-weight vector $\tilde{w}$ for the estimator π_{max} based on 200 simulations under four different models. Each column represents an n dimensional weight vector, and there are 200 columns in each plot.
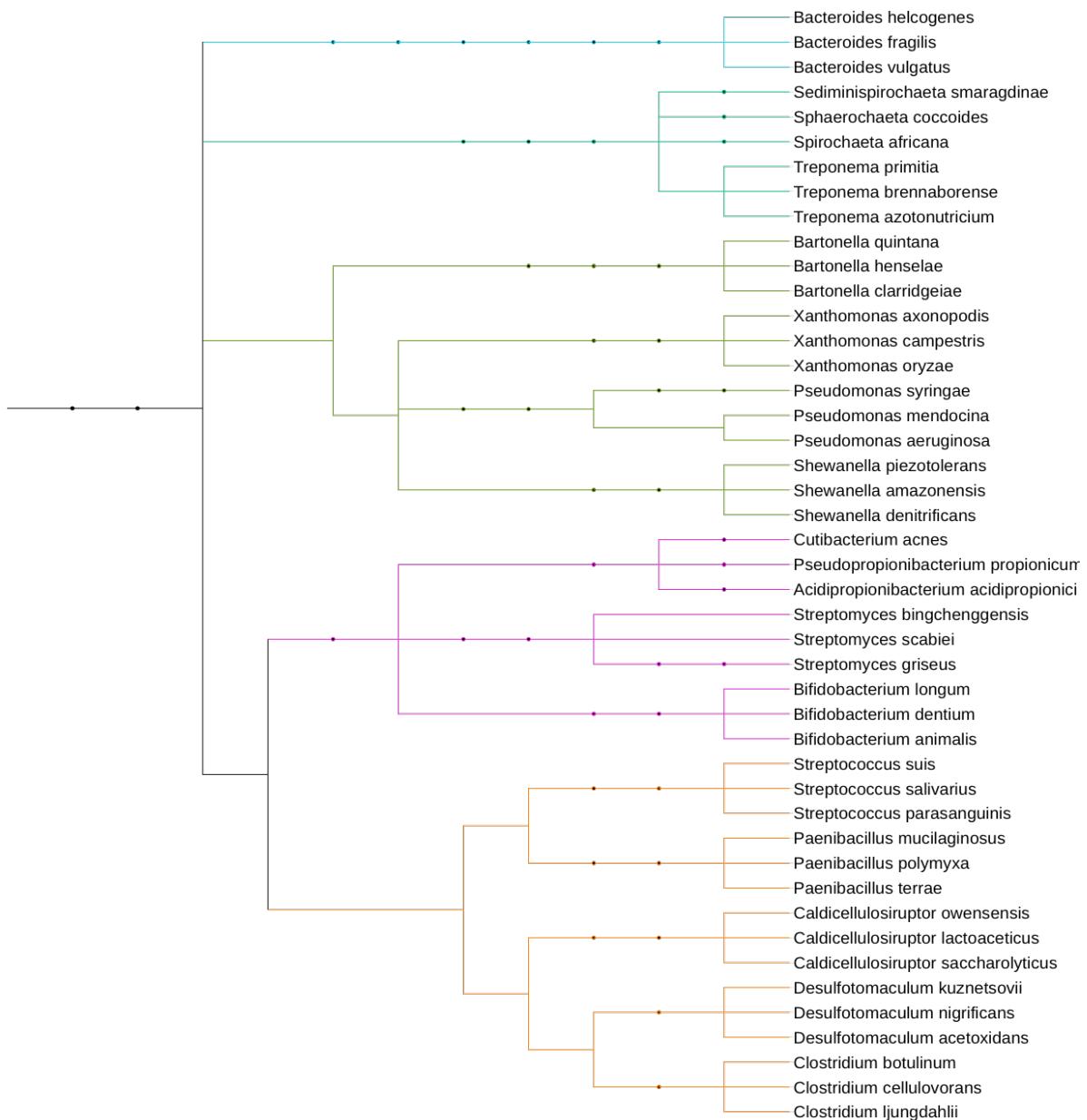
Figure 8: Taxonomic tree of 45 closely related species used in generating the shotgun metagenomic data used on in Gao and Li (2018).

Table S.1: The p-values of 8 contig clusters from the Wilcoxon rank sum test of the ePTRs between the responser and non-responser groups, and the taxonomic annotations with lineage scores indicating the quality of each taxonomic classification.

Contig Clusters (n_1, n_2)	P-values	Taxonomic Annotations with Lineage Scores
bin.004 (8,8)	0.9592	Firmicutes (phylum): 0.88; Clostridia (class): 0.78; Clostridiales (order): 0.78; Lachnospiraceae (family): 0.40
bin.007 (9,11)	0.4119	Firmicutes (phylum): 0.96; Clostridia (class): 0.92; Clostridiales (order): 0.92;
bin.016 (5,8)	0.3543	Firmicutes (phylum): 0.86; Clostridia (class): 0.74; Clostridiales (order): 0.74;
bin.017 (7,7)	0.2086	Firmicutes (phylum): 0.92; Erysipelotrichia (class): 0.47; Erysipelotrichales (order): 0.47; Erysipelotrichaceae (family): 0.47;
bin.026 (7,9)	0.0418	Firmicutes (phylum): 0.90; Clostridia (class): 0.88; Clostridiales (order): 0.88;
bin.041 (6,6)	0.2402	Firmicutes (phylum): 0.95; Clostridia (class): 0.91; Clostridiales (order): 0.91; Lachnospiraceae (family): 0.49; Roseburia (genus): 0.45;
bin.058 (8,15)	0.5063	Bacteroidetes (phylum): 0.89; Bacteroidia (class): 0.88; Bacteroidales (order): 0.88; Bacteroidaceae (family): 0.84; Bacteroides (genus): 0.84;
bin.065 (5,8)	0.0653	Bacteroidetes (phylum): 0.88; Bacteroidia (class): 0.88; Bacteroidales (order): 0.88; Bacteroidaceae (family): 0.85; Bacteroides (genus): 0.85;