Sparse equisigned PCA: Algorithms and performance bounds in the noisy rank-1 setting[*]

Arvind Prasadan and Raj Rao Nadakuditi

Department of EECS
University of Michigan
1301 Beal Avenue
Ann Arbor, MI 48109
e-mail: prasadan@umich.edu*;* rajnrao@umich.edu

Debashis Paul

Department of Statistics
University of California
One Shields Avenue
Davis, CA 95616
e-mail: debpaul@ucdavis.edu

Abstract: Singular value decomposition (SVD) based principal component analysis (PCA) breaks down in the high-dimensional and limited sample size regime below a certain critical eigen-SNR that depends on the dimensionality of the system and the number of samples. Below this critical eigen-SNR, the estimates returned by the SVD are asymptotically uncorrelated with the latent principal components. We consider a setting where the left singular vector of the underlying rank one signal matrix is assumed to be sparse and the right singular vector is assumed to be equisigned, that is, having either only nonnegative or only nonpositive entries. We consider six different algorithms for estimating the sparse principal component based on different statistical criteria and prove that by exploiting sparsity, we recover consistent estimates in the low eigen-SNR regime where the SVD fails. Our analysis reveals conditions under which a coordinate selection scheme based on a *sum-type decision statistic* outperforms schemes that utilize the ℓ_1 and ℓ_2 norm-based statistics. We derive lower bounds on the size of detectable coordinates of the principal left singular vector and utilize these lower bounds to derive lower bounds on the worst-case risk. Finally, we verify our findings with numerical simulations and a illustrate the performance with a video data where the interest is in identifying objects.

MSC 2010 subject classifications: Primary 65F50, 62H25, 62H15; secondary 62F03, 60G35, 15A18.
Keywords and phrases: Sparse PCA, random matrices, FDR, sparsistency.

Received May 2019.

Contents

1. Introduction

It is well-understood that singular value decomposition (SVD) based principal component analysis (PCA) breaks down in the high-dimensional and limited sample size regime below a certain critical eigen-SNR (eigenvalue signal-to-noise ratio) that depends on the dimensionality of the system and the number of samples [21, 5]. Several sparse PCA algorithms have been proposed in the literature (see [21, 5, 11, 25, 41, 4]) and have been shown to successfully estimate the principal components in the low eigen-SNR regime where the SVD fails.

Prior work in this area primarily considers the Gaussian signal-plus-noise model with random effects, where the signal matrix is assumed to have sparse left singular vectors, normally distributed right singular vectors, and the noise matrix is assumed to have normally distributed *i.i.d.* entries. Here, we consider the setting where the left singular vector of the rank one signal matrix is sparse *and* the right singular vector is assumed to be equisigned. We say that a vector is equisigned if its entries are all non-negative or all non-positive. This is motivated by applications such as diffusion imaging in MRI where the right singular vector represents a physical quantity (e.g. intensity as the diffusion agent is absorbed by a tissue) that is non-negative, by imaging problems such as foreground-background separation in video data [32, 39] and object detection in astronomy [35], where the data are naturally non-negative, and by problems in bioinformatics where the data are (non-negative) counts of genes [38]. When analyzing data that are non-negative, it is logical to take advantage of this property, and investigate how we may use this knowledge to do better than the (generic) alternatives. Alternatively, a practitioner may seek to use techniques that constrain or impose non-negativity to preserve interpretability of the results, e.g., non-negative matrix factorization. Additionally, we motivate the rank-1 assumption by noting that for a video with a static background, the foreground is a perturbation of a rank-1 background [27, 16]. Finally, even though we do not pursue this angle here, our framework can be extended to deal with the scenario where the signal can be viewed of a rank 1 tensor with all but one of the representors in the Kroneker product representation of the tensor is an equisigned vector.

There is precedent for and prior work on non-negative PCA, including the sparse biased PCA in [9], the sparse PCA with non-negativity priors in [31], and the work in [26]. These works differ from our work in that they impose non-negativity on the factors or the left singular vectors. In this work, we study

sparse factors with non-negative loadings; i.e., we are solving a different problem in this work.

A natural question at this juncture is the following: *how does our problem differ from that solved by Non-Negative Matrix Factorization (NNMF)?* NNMF takes a given matrix $\mathbf{X}$ and looks for non-negative matrices $\mathbf{F}$ and $\mathbf{G}$ such that $\mathbf{X} = \mathbf{F}\,\mathbf{G}^T$ [19, 40]. Ordinary NNMF has no sparsity constraints. We might impose such constraints, as is done in [18] and [24], but except in special cases, these solutions have no known theoretical guarantee of statistical performance. This problem partly stems from the fact that solutions to the corresponding optimization problems may not be unique. In contrast, our problem only constrains the right singular vectors, while the left singular vectors are free to take any sign. The work in [12] extends the NNMF framework to one wherein only one of the factors is non-negative; nevertheless, the rest of the constraints we impose are not included. The work in [42] seeks factors (left singular vectors) with disjoint supports and non-negative loadings, but this definition of sparsity does not match that from the sparse PCA literature. Hence, NNMF is not an answer to the problem we consider herein.

The main contribution of this paper is a rigorous sparsistency analysis of the various algorithms that brings into focus the various very-low eigen-SNR regimes where the new algorithms work and the SVD based methods provably fail. Additionally, a major novelty of this work is the integration of FDR-controlling (False Discovery Rate) hypothesis testing to the Sparse PCA problem.

Our analysis illustrates the situations where the sum based coordinate selection scheme dramatically outperforms the ℓ_1 and ℓ_2 [21, 5] based sparse PCA schemes. Additionally, our proposed algorithms are non-iterative, do not require the computation of the sample covariance matrix, and do not require knowledge of the sparsity level. We separate our algorithms into two groups: one where the Family-Wise Error Rate (FWER) is controlled, and another where the False Discovery Rate (FDR) is controlled. We utilize sharp tail probability bounds for relevent statistics to derive our FWER-controlling estimators [7]. For the FDR controlling estimators, we relate the problem at hand to that of the sparse normal means problem [13].

This paper is organized as follows. In Section 3, we describe three algorithms for estimating the sparse principal component that utilize a coordinate selection scheme based on the sum, ℓ_1, and ℓ_2 norm-based statistics respectively. We call our family of algorithms *SEPCA*, an abbreviation for Sparse Equisigned PCA. Section 4 proposes three FDR-controlling refinements of the sum- and ℓ_2-based algorithms in Section 3 by relating coordinate detection to the sparse normal means estimation problem. In Section 5 we show how the estimation performance is governed by the size of the smallest detectable coordinate, which we analyze in Section 6 and validate using numerical simulations in Section 7. In Section 8, we provide some geometric intuitions about the relative performance of three of our algorithms. We show that the sum statistic is potentially the most powerful, while the ℓ_1 is the least powerful. We provide some concluding remarks in Section 9.

2. Problem formulation

Let $\mathbf{X} \in \mathbb{R}^{p \times n}$ be a real-valued signal-plus-noise data matrix of the form

$$\mathbf{X} = \theta \, \mathbf{u} \, \mathbf{v}^T + \sigma \, \mathbf{G} \,. \tag{2.1}$$

The columns of the $p \times n$ data matrix $\mathbf{X}$ represent p-dimensional observations. In (2.1), $\mathbf{u}$ and $\mathbf{v}$ are the left and right singular vectors of the rank-one latent signal matrix, and have entries u_i and v_j, respectively. The entries of $\mathbf{G}$, the noise matrix, are assumed to be *i.i.d.* Gaussian random variables with mean 0 and variance $1/n$. We assume that $\mathbf{u} \in \mathbb{R}^p$ has unit norm and is sparse in the sense of small ℓ_0 norm, with $s \ll p$ non-zero entries, where $s/n \to 0$. That is, for a set $I = \{i_1, \cdots, i_s\} \subset \{1, \cdots, p\}$,

$$\begin{array}{ll} u_i \neq 0 & \text{for } i \in I, \\ u_i = 0 & \text{for } i \in I^C, \end{array} \tag{2.2}$$

where I^C denotes the complement of I. We further assume $\mathbf{v} \in \mathbb{R}^n$ to be of unit norm, deterministic, and equisigned. Given $\mathbf{X}$, our goal is to recover $\mathbf{u}$ and $\mathbf{v}$.

Note that the (i, k) entry of $\mathbf{X}$, X_{ik}, is a Gaussian random variable with mean $[\theta u_i] v_k$ and variance σ^2/n. Moreover, it follows that

$$\mathbb{E}(\mathbf{X} \, \mathbf{X}^T) = \theta^2 \, \mathbf{u} \, \mathbf{u}^T + \sigma^2 \, \mathcal{I}_p,$$

where $\mathcal{I}_p$ denotes the $p \times p$ identity matrix. The quantity $(\theta/\sigma)^2$ is, for this model, the eigen-SNR (signal-to-noise ratio).

2.1. Motivation: Breakdown of PCA/SVD

From [3], we have the following result: let $\widehat{\mathbf{u}}$ be the estimate of $\mathbf{u}$ given by the Singular Value Decomposition (SVD) of $\mathbf{X}$, and let $p(n)/n$ have limit $c \in [0, \infty]$ as n grows, with θ fixed and $\sigma = 1$. Then, with probability 1,

$$|\langle \widehat{\mathbf{u}}, \mathbf{u} \rangle|^2 \to \begin{cases} 1 - \frac{c(1+\theta^2)}{\theta^2(c+\theta^2)} & \text{if } \theta \geq c^{1/4}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

For general σ, we replace θ by θ/σ in (2.3). Hence, SVD based PCA leads to inconsistent estimates of $\mathbf{u}$ (and also for $\mathbf{v}$, which can be deduced from (2.3)) when the dimension p is comparable to or larger than the sample size n. Moreover, in the low eigen-SNR regime, the estimates break down completely. SVD does not exploit any assumed structure in $\mathbf{u}$ and $\mathbf{v}$. Consequently, (2.3) holds for arbitrary $\mathbf{u}$ and $\mathbf{v}$, including our setting where $\mathbf{u}$ is sparse and/or $\mathbf{v}$ is equisigned. Our goal, in what follows, is to derive consistent estimators for $\mathbf{u}$ and $\mathbf{v}$ that outperform the SVD by exploiting the sparsity of $\mathbf{u}$ and the equisigned nature of $\mathbf{v}$.

2.2. Problem statement

Note that we have assumed that $\mathbf{u}$ is sparse and that the sparsity s is such that s/n has limit zero. Hence, if we had oracle knowledge of the sparsity pattern I (the indices of $\mathbf{u}$ that have non-zero coordinates), restricting the matrix X to those rows indexed by I and performing the SVD on the smaller matrix would yield a consistent estimator for the non-zero elements of $\mathbf{u}$ and the vector $\mathbf{v}$. This conclusion follows from (2.3), since the value c is replaced with s/n, which has limit zero. Thus, if we derived consistent estimators of the support of $\mathbf{u}$, we have a consistent two-stage estimation procedure of the vectors $\mathbf{u}$ and $\mathbf{v}$.

Formally, we are interested in finding a procedure that estimates I by $\widehat{I}$ such that in the limit $n \to \infty$,

$$\begin{cases} \mathbb{P}\left(i \in \widehat{I}\right) \to 1 & \text{if } i \in I, \\ \mathbb{P}\left(i \in \widehat{I}\right) \to 0 & \text{if } i \notin I. \end{cases} \tag{2.4}$$

Equivalently, noting that the Hamming distance of I and $\widehat{I}$, denoted by $d_H\left(I, \widehat{I}\right)$, is given by the cardinality of their symmetric set difference,

$$d_H\left(I, \widehat{I}\right) = \left|\left(I \cup \widehat{I}\right) \setminus \left(I \cap \widehat{I}\right)\right|,$$

we want the expected Hamming distance $\mathbb{E}\, d_H\left(I, \widehat{I}\right)$ to have limit 0, which is stronger than requiring consistency in recovering the support (or sparsity pattern) of $\mathbf{u}$. However, as the work in [10, 34] indicates, this limit will not in general be zero, and will depend on the noise level, signal strength, and sparsity.

3. Proposed algorithms

We propose six different two-stage algorithms for estimating $\mathbf{u}$. The first three algorithms are designed to control the family-wise error rate (FWER), or, the probability of obtaining a false positive in the coordinate selection. The last three algorithms aim to control the false discovery rate (FDR), or, the proportion of false discoveries (coordinate detections) among all discoveries. We defer discussion of the FDR-based algorithms to Section 4.

All of the algorithms have the same basic form given in Algorithm 1. Given $\mathbf{X}$, we associate a test statistic T_i to each row of $\mathbf{X}$. The sparsity of $\mathbf{u}$ implies that the majority of the rows of $\mathbf{X}$ are purely noise, so that the majority of the T_i come from the null, noise-only distribution. Hence, based on the statistics $\{T_i\}$, we perform a form of multiple hypotheses testing procedure, and select the set $\widehat{I}$ of indices that are non-null. In this way, we can estimate the *support* of $\mathbf{u}$, thereby isolating the the rows of $\mathbf{X}$ that contain the signal. Then, taking the SVD of this submatrix (comprised of only the selected rows of $\mathbf{X}$) yields a better estimate of the non-zero coordinates in $\mathbf{u}$, as well as $\mathbf{v}$.

We begin by discussing the FWER-controlling algorithms. The work in [21] proposed a covariance thresholding method for Sparse PCA called DT-SPCA;

Algorithm 1 Variable Selection and Estimation Algorithm

Require: Threshold $\tau_{n,p}$ and form of Test Statistic T_i from Table 1

  Let $\widehat{I}$ be an empty list

  for all Rows i of $\mathbf{X}$, $1 \leq i \leq p$ **do**

    Form test statistic T_i from row i of $\mathbf{X}$

    if $T_i \geq \tau_{n,p}$ **then**

      Add i to $\widehat{I}$

    end if

  end for

  Let $[\widetilde{\mathbf{u}}, \widetilde{\theta}, \widehat{\mathbf{v}}] = \mathrm{SVD}(\mathbf{X}_{\widehat{I},:})$ be the rank-1 SVD of $\mathbf{X}$ restricted to rows in $\widehat{I} = [i_1, \cdots, i_{|\widehat{I}|}]$

  For $i_k \in \widehat{I}$, let $\widehat{u}_{i_k} = \widetilde{u}_k$; the other entries of $\widehat{\mathbf{u}}$ are set to 0.

---

this is equivalent to a coordinate selection scheme based on the ℓ_2 norm-based statistic. In our terminology and with our choice of thresholds, we label it as ℓ_2-*SEPCA*. We label the coordinate selection scheme based on the ℓ_1 norm-based statistic ℓ_1-*SEPCA*. Finally, the *sum-SEPCA* algorithm utilizes row sums of the data matrix.

3.1. *Computational complexity*

Note that the variable selection part of our procedures has a computational complexity that is $O\left(pn\right)$: the formation of the test statistic is linear in the number of columns, and the formation is repeated once per row. Noting that for a $p \times n$ matrix, the complexity of the rank-1 SVD is $O\left(1 \times pn\right)$, we find that if $\left|\widehat{I}\right|$ coordinates are selected, we have an overall complexity of $O\left(pn + \left|\widehat{I}\right|n\right) = O\left(pn\right)$ [1].

Computation of the covariance matrix has a (naive) complexity of $O\left(p^2 n\right)$, and in practice is somewhere between $O\left(p^2\right)$ and $O\left(p^3\right)$ [15]. Immediately, our methods here are faster than those requiring explicit formation of the covariance matrix [5, 21, 25, 41, 4]. Additionally, there is no iteration or convergence of any optimization problems required. Note that a semi-definite programming-based formulation is at least polynomial in the problem size: $O\left(p^4\right)$ [11] or $O\left(p^3\right)$ [4]. The ITSPCA method applied to our rank-1 setting would have a cost of $O\left(ps\right)$ per iteration [25, Sec. 4]. TPower has a similar complexity of $O\left(sp + p\right)$ per iteration [41].

3.2. *The DT-SPCA algorithm and two-stage procedures*

The DT-SPCA algorithm was proposed in [21] and later used as the first stage of the ASPCA algorithm given in [5]. The algorithm thresholds the diagonals of the matrix XX^T to perform variable selection: note that in our setting, these values are

$$\theta^2 u_i^2 + \sigma^2 \sum_{j=1}^n G_{ij}^2,$$

with expectation $\left(\theta^2 u_i^2 + \sigma^2\right)$. The DT-SPCA algorithm thresholds these diagonal values at $\sigma^2\left(1 + \gamma\sqrt{\frac{\log p}{n}}\right)$, where $\gamma > 0$, and then performs PCA on the

reduced matrix formed from the selected variables. Noting that the diagonals of XX^T are the same as the row sum-of-squares of X, we see that ℓ_2-SEPCA is essentially the same (up to choice of threshold) as DT-SPCA.

However, the innovation of [21, 5] that we carry forward is the two-stage procedure. That is, we perform some sort of testing to estimate the support of the sparse singular vector $\mathbf{u}$, and then perform an SVD on the reduced matrix. As we will see in what follows, there is flexibility in the choice of testing or support estimation method.

3.3. Statement of thresholds

We shall choose the thresholds $\tau_{n,p}$ for the coordinate selection scheme so that in the noise-only case,

$$\mathbb{P}\left(\max_{1 \leq i \leq p} T_i \geq \tau_{n,p}\right) \leq \frac{1}{ep} \to 0, \tag{3.1}$$

where e is Euler's number, or the base of the natural logarithm. This choice ensures that the probability of a false positive tends to zero as $p \to \infty$. That is, the FWER is asymptotically zero and is bounded by $1/ep$ in the finite-dimensional case. Note that the constraint used to control the FWER is simply that the distribution of the noise is log-concave. In the Gaussian case, we obtain the specific expressions given summarized in Table 1; however, with knowledge of the moments $\mathbb{E}T_i$ and Var T_i, we can repeat our analysis and find thresholds for the ℓ_1 and ℓ_2-SEPCA algorithms with *any* log-concave noise distribution. The thresholds are summarized in Table 1.

TABLE 1
Test Statistics and Thresholds for Algorithm (1)

Algorithm	Statistic T_i	Threshold $\tau_{n,p}$
ℓ_1-SEPCA	$\frac{1}{\sqrt{n}} \sum_{k=1}^n |X_{i,k}|$	$\sigma \left(\sqrt{\frac{2}{\pi}} + C_1 \frac{\log ep}{\sqrt{n}}\right)$
ℓ_2-SEPCA	$\sum_{k=1}^n X_{i,k}^2$	$\sigma^2 \left(1 + C_2 \frac{\log ep}{\sqrt{n}}\right)$
sum-SEPCA	$\frac{1}{\sqrt{n}} |\sum_{k=1}^n X_{i,k}|$	$\sigma C_U \sqrt{\frac{\log p}{n}}$

See (3.2) and (3.6) for definitions of the constants C_2, C_1, and C_U.

Remark Note that we impose strong control over the FWER and seek to reject individual null hypotheses, instead of weak control and considering the global null hypothesis as in [4].

3.4. FWER thresholds

3.4.1. ℓ_2- and ℓ_1-SEPCA

In the noise-only cases, the statistics for ℓ_2- and ℓ_1-SEPCA are distributed as scaled χ_n^2 and sums of half-normal, respectively. Both of these quantities are

log-concave random variables, so we may apply the result in [23] to set the threshold $\tau_{n,p}$ in both cases.

Defining K to be some absolute constant (we may use $K = e$, as in [6]), we define the constants

$$C_2 = \sqrt{2}K \text{ and } C_1 = K\sqrt{(1 - 2/\pi)}. \tag{3.2}$$

3.4.2. sum-SEPCA

From Proposition 4.4 of [8], we obtain that the threshold for sum-SEPCA is given by

$$\tau_{n,p} = \frac{\sigma}{\sqrt{n}} \left(\sqrt{2\log p} + \frac{1}{U(p)} \left(\frac{1}{3}\log ep + \sqrt{\log ep} \right) + \delta_p \right). \tag{3.3}$$

In (3.3), we have that

$$U(p) = \sqrt{2}\,\mathrm{Erf}^{-1}\left(1 - \frac{1}{p}\right) \text{ and } \delta_p \asymp \frac{\pi^2}{12}\left(\log p\right)^{-3/2}, \tag{3.4}$$

where Erf denotes the *error function*, or alternatively, the cumulative distribution function of a standard Gaussian random variable is given by

$$\Phi(x) = \frac{1}{2}\left(1 + \mathrm{Erf}\left(\frac{x}{\sqrt{2}}\right)\right). \tag{3.5}$$

Moreover, $\tau_{n,p} \leq \sigma C_U \sqrt{\frac{\log p}{n}}$ for some constant C_U. For a fixed value of p, choosing

$$\kappa_U \geq \frac{\sqrt{2}}{U(p)}\left(3 + \sqrt{\log p}\right) > 1 \text{ and } C_U = \sqrt{2} + \frac{\kappa_U}{3\sqrt{2}} \tag{3.6}$$

is sufficient. The choice of $1/ep$ is the largest bound justified by Proposition 4.4 of [8], so we have calibrated all of our algorithms to the same constant factor times $1/p$. The thresholds are summarized in Table 1.

3.5. Estimation of the noise variance, σ^2

In this work, we assume that the noise variance σ^2 is known; however, in general, estimation of σ^2 may not be straightforward [29]. Recently proposed procedures such as those proposed in [29, 30, 37] could be employed to estimate the noise variance, and we point the interested reader to these references for more theoretical background on the problem. We note that in most applications, including the video example we consider, one can obtain a relatively sparse representation of the object in a multiscale basis such as a wavelet basis [20, Sec. 7.5]. Under such circumstances, under the assumed additive, isotropic noise model, we can easily obtain a consistent estimate of σ^2 by utilizing the inherent sparsity of the

signal, especially in finer scales. This can be done, for example, by computing the variance of the wavelet coefficients in the finest scale [20, Sec. 7.5]. One can obtain a more robust estimate by taking the median absolute deviation of the coefficients about their median and then by multiplying its square with a known scale factor (assuming normality) [30, 20].

4. Controlling the false discovery rate

So far, we have controlled the probability of a false alarms when detecting coordinates. However, there are two relevant observations to make. First, under the Gaussian noise, rank-1, and equisigned assumptions, the vector of test statistics $\{T_i\}$ in the sum-SEPCA algorithm looks like a sparse vector plus Gaussian noise (or a vector of χ_n^2-variates with varying non-centralities, in the ℓ_2-SEPCA algorithm). Secondly, controlling the false discovery rate, that is, the proportion of rejected nulls that are false positives, can lead to increased detection power relative to controlling the false positive rate. We hence look at FDR-controlling tests for the *Sparse Normal Means* problem.

That is, given a vector of test statistics (as before), we replace the thresholding and selection in Algorithm 1 with an FDR-controlling selection procedure. We summarize this change in Algorithm 2. There are three procedures we consider. The first two are known as Higher Criticism, and directly extend the sum- and ℓ_2-SEPCA algorithms [13, 14]. The third is a method for detection in the sparse normal means problem that comes out of complexity-penalized estimation theory for linear inverse problems [22].

Algorithm 2 FDR-Controlling Variable Selection and Estimation Algorithm

Require: Test Statistic T_i from Table 1 and Selection Procedure
  Let $\widehat{I}$ be an empty list
  for all Rows i of $\mathbf{X}$, $1 \leq i \leq p$ **do**
    Form test statistic T_i from row i of $\mathbf{X}$
  end for
  Perform an FDR-Controlling selection procedure, and add the selected indices to $\widehat{I}$
  Let $[\widetilde{\mathbf{u}}, \widetilde{\theta}, \widehat{\mathbf{v}}] = \mathrm{SVD}(\mathbf{X}_{\widehat{I},:})$ be the rank-1 SVD of $\mathbf{X}$ restricted to rows in $\widehat{I} = [i_1, \cdots, i_{|\widehat{I}|}]$
  For $i_k \in \widehat{I}$, let $\widehat{u}_{i_k} = \widetilde{u}_k$; the other entries of $\widehat{\mathbf{u}}$ are set to 0.

4.1. Higher criticism

4.1.1. Formulation

Assume we have p independent tests of the form

$$
\begin{aligned}
H_{o,i}: &\quad W_i \sim \mathcal{N}\left(0,1\right), \\
H_{1,i}: &\quad W_i \sim \mathcal{N}\left(\mu_i, 1\right),
\end{aligned}
\tag{4.1}
$$

and assume that at most $p^{1-\beta}$ of the p hypotheses are truly non-null, for some $\beta \in (1/2, 1)$. Further assume that the non-null means have magnitude

$$\mu_i = \mu_p = \sqrt{2r \log p},$$

for $r \in (0, 1)$. Here, the means will correspond to the coordinate size. Note that the expected maximum of p standard Gaussian random variables is upper bounded by $\sqrt{2 \log p}$, with the bound being asymptotically sharp.

If we let $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(p)}$ be the sorted p-values of the individual tests, we may define the Higher Criticism statistic:

$$HC_p = \max_{i:1/p \leq p_{(i)} \leq 1/2} \frac{\sqrt{p}\left(i/p - p_{(i)}\right)}{\sqrt{p_{(i)}(1 - p_{(i)})}}. \tag{4.2}$$

Rejecting the global null hypothesis (that there are no non-null coordinates) when $HC_p > \sqrt{2 \log \log p}(1 + o(1))$ leads to asymptotically full power when r is greater than some decision boundary ρ, and that under the global null,

$$\frac{HC_p}{\sqrt{2 \log \log p}} \to 1 \tag{4.3}$$

in probability as $n, p \to \infty$. The function ρ depends on the sparsity index β, and as [14] indicate:

$$\rho(\beta) = \begin{cases} \beta - 1/2 & \text{when } \beta \in (1/2, 3/4), \\ \left(1 - \sqrt{1 - \beta}\right)^2 & \text{when } \beta \in (3/4, 1). \end{cases} \tag{4.4}$$

If we replace the normal distribution with a χ_n^2 distribution, the same results hold for tests of the form

$$\begin{aligned} H_{o,i} : \quad & W_i \sim \chi_n^2, \\ H_{1,i} : \quad & W_i \sim \chi_n^2(\delta), \end{aligned} \tag{4.5}$$

where δ is a non-centrality parameter and we consider $r \in (0, 1)$ such that $\delta = 2r \log p$.

Remark While Higher Criticism is typically formulated for the case of identical non-null means or parameters (all of the non-zero μ_i are identical), this constraint is not mandatory [2, 17]. Indeed, the results hold without modification for the Gaussian model with non-null means of size $\mu_i = \alpha_i \sqrt{2 \log p}$, where α_i is a non-negative random variable with the property that $\mathbb{P}(\alpha_i \leq \sqrt{r}) = 1$ and $\mathbb{P}(\alpha_i > \sqrt{r} - \epsilon) > 0$ for all $\epsilon > 0$ [17]. The case of a χ_n^2 distribution is similar.

As a point of interest, the test in (4.1) can be extended to (and potentially strengthened in) the case where the p tests are correlated, i.e., when the additive Gaussian noise has a non-identity covariance [17].

4.1.2. Application to our problem

Recall that for the sum-SEPCA algorithm, we formed a vector of row-sums. That is, in the equisigned setting, taking sums across the rows of $\mathbf{X}$, we obtain a vector $\mathbf{y}$ where $y_i = \mu_i + \sigma z_i$, with $\mu_i = (\theta u_i)\| \mathbf{v} \|_1$: this situation is exactly that of a sparse mean vector embedded in Gaussian noise. Similarly, taking sums of squares across the rows of $\mathbf{X}$ (as in the ℓ_2-SEPCA algorithm) yields scaled χ^2_n distributed random variables, of which only a few have non-zero non-centrality parameters.

With knowledge of the noise distribution, we may compute the p-values of each row statistic: these p-values are used to form the Higher Criticism statistic (4.2). As in [14], we may adapt the original global testing problem to a selection problem. For each p-value $p_{(i)}$, we have a value $HC_{p,i}$ of the higher criticism statistic (the value that is maximized in (4.2)). Rejecting each null hypothesis (that the coordinate of the corresponding row is zero) when $HC_{p,i}$ is larger than the threshold $\sqrt{2\log\log p}$ is a variable selection procedure. We refer to the procedure based on the sum statistic as HC-sum-SEPCA and that based on the sum of squares statistic as HC-ℓ_2-SEPCA. Importantly, we note that the form of the decision boundary ρ is identical to the global testing case, and that applying Higher Criticism to our row statistics is a viable global testing procedure [14].

4.2. FDR-SEPCA

In this section, we give an summary of the algorithm for uncorrelated noise and defer the general case and details to Appendix D. We continue in the same vein as in the previous section on Higher Criticism.

We note that in the equisigned, rank-1 setting, coordinate selection is equivalent to the estimation of a sparse mean vector. Let $y_i = \mu_i + \sigma z_i$, where $i \in \{1, \cdots, p\}$ and the vector $\mathbf{z}$ of the z_i is normally distributed with mean 0 and covariance $\mathcal{I}_p$. The mean vector $\boldsymbol{\mu}$ of the μ_i is assumed to be sparse; the goal is to estimate $\boldsymbol{\mu}$. Taking sums across the rows of $\mathbf{X}$, we obtain a vector $\mathbf{y}$ where $y_i = \mu_i + \sigma z_i$, with $\mu_i = (\theta u_i)\| \mathbf{v} \|_1$. Hence, we are in the same setting as in the previous section.

The following penalized least squares formulation, taken from [22], yields an estimator for μ:

$$\widehat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \| \mathbf{y} - \boldsymbol{\mu} \|_2^2 + \sigma^2 \mathrm{pen}\left(\| \boldsymbol{\mu} \|_0\right), \tag{4.6}$$

where $\mathrm{pen}(k)$ is defined as

$$\mathrm{pen}(k) = \zeta k \left(1 + \sqrt{2\log(\nu p/k)}\right)^2, \tag{4.7}$$

with $\zeta > 1$; we may take $\zeta = 1 + o(1)$. The parameter ν is no smaller than e. We define $\| \boldsymbol{\mu} \|_0$ to be the number of non-zero coordinates of μ.

The solution to (4.6) is given by hard-thresholding. Let $|y|_{(i)}$ be the ith order statistic of $|y_i|$, namely $|y|_{(1)} \geq \cdots \geq |y|_{(p)}$. Then if

$$\widehat{k} = \arg\min_{k \geq 0} \sum_{i > k} |y|_{(i)}^2 + \sigma^2 \mathrm{pen}(k), \tag{4.8}$$

defining

$$t_k^2 = \mathrm{pen}(k) - \mathrm{pen}(k-1), \tag{4.9}$$

the solution is to hard threshold at $t_{\widehat{k}}$.

In this set-up, we have that

$$t_k \approx \sqrt{\zeta}(1 + \sqrt{2\log(\nu p/k)}).$$

We provide a precise quantification of t_k in Appendix D.

Hence, by computing t_k and performing hard thresholding of the row sums, we can perform coordinate selection. Once again, this procedure replaces the test statistic/thresholding in Algorithm 1.

5. Estimation error and smallest detectable coordinate

As we will see, our theorems discuss the "detectability" of the coordinates u_i of $\mathbf{u}$. However, it is common in the sparse PCA literature to discuss lower bounds for the risk (estimation error) [21, 5, 25]. In what follows, we will show that these two notions are equivalent.

We define the L^2 estimation error for a principal component estimator as

$$L(\widehat{\mathbf{u}}, \mathbf{u}) = \|\mathbf{u} - \mathrm{sign}(\langle \mathbf{u}, \widehat{\mathbf{u}}\rangle)\widehat{\mathbf{u}}\|_2^2. \tag{5.1}$$

The quantity in (5.1) is upper bounded by 2; this bound is attained when $\mathbf{u}$ and $\widehat{\mathbf{u}}$ are unit norm and mutually orthogonal. Following [5], we want to compute a lower bound for the maximum expected loss for the s-sparse vectors $\mathbf{u}$ (in the sense of ℓ_0 sparsity) defined as

$$\sup_{\mathbf{u} \in \mathbb{S}^{p-1} : \|\mathbf{u}\|_0 \leq s} \mathbb{E}L\left(\widehat{\mathbf{u}}, \mathbf{u}\right), \tag{5.2}$$

where $\mathbb{S}^{p-1}$ denotes the unit sphere in $\mathbb{R}^p$. Let $\widehat{I}$ be some index set of coordinates selected by an algorithm of the form given in Algorithm (1). We may take $\langle \mathbf{u}, \widehat{\mathbf{u}}\rangle$ to be non-negative, and decompose the loss as

$$\|\mathbf{u} - \widehat{\mathbf{u}}\|_2^2 = \underbrace{\|\mathbf{u}_{\widehat{I}} - \widehat{\mathbf{u}}\|_2^2}_{\substack{\text{Estimation Error from} \\ \text{detected coordinates}}} + \underbrace{\|\mathbf{u}_{\widehat{I}^c}\|_2^2}_{\substack{\text{Error from} \\ \text{missed coordinates}}} \geq \|\mathbf{u}_{\widehat{I}^c}\|_2^2. \tag{5.3}$$

Equation (5.3) shows that the loss is lower-bounded by the squared sum of the missed coordinates. Indeed, it is a natural consequence of the result in [3] that

if the sparsity s grows slower than does n, and we have a consistent estimate of the support of $\mathbf{u}$, the estimation error will asymptotically be small. Essentially, we are estimating the singular vectors of an $s \times n$ matrix instead of a $p \times n$ matrix, so that if the ratio s/n has limit zero, our estimates will be consistent (see (2.3) and [3]). This suggests the following strategy for lower-bounding (5.2): we want to construct a non-trivial 'worst-case' sparse vector. That is, we want a vector $\mathbf{u}$ that has a non-trivial loss (less than 2), is sparse (fewer than s non-zero coordinates), and has maximal error from missed coordinates. To ensure a non-trivial loss, we set the first coordinate u_1 to be large, *i.e.*, $u_1 = \sqrt{1 - r^2}$, where $r = o(1)$. To ensure sparsity, we set $u_2, \cdots, u_{m+1}$ to be non-zero for some $m \leq s - 1$, with the subsequent coordinates of u set to 0. Then, the expected loss has the lower bound

$$
\begin{aligned}
\mathbb{E}L(\mathbf{u}, \widehat{\mathbf{u}}) &\geq \sum_{k=1}^{p} |u_k|^2 \mathbb{P}\left(\text{Not Selecting Coordinate k}\right) \\
&\geq \sum_{k=2}^{m+1} |u_k|^2 \mathbb{P}\left(\text{Not Selecting Coordinate k}\right),
\end{aligned}
\tag{5.4}
$$

since u_1 is detected with probability approaching 1 and u_k is zero for $k > m+1$. Now, let u_2 through u_{m+1} all have value $r/\sqrt{m}$, so that we may simplify the lower bound to

$$
\mathbb{E}L(\mathbf{u}, \widehat{\mathbf{u}}) \geq r^2 \mathbb{P}\left(\text{Not Selecting Coordinate k}\right).
\tag{5.5}
$$

If coordinates of size $r/\sqrt{m}$ are not detected with a probability approaching 1, r^2 is a lower-bound on the risk. This construction shows that specifying the sizes of coordinates that are not detected with probability approaching 1 is equivalent to specifying a worst-case risk lower bound. Note that the value of r^2 depends on the specific algorithm and estimator, and that this is not a general or universal bound. Rather, the purpose of this construction is to show the equivalence between the two perspectives (a lower bound and detectable coordinate size).

Consequently, in what follows we focus on the smallest detectable and largest undetectable coordinates because they directly shed light on the attainable estimation error. The details of the risk calculations and extensions to approximate sparsity are deferred to Appendix C, where we summarize our findings in Theorem 3.

6. Main results

The following theorem characterizes consistent support recovery conditions. These results are the analogue of the 'sparsistency' guarantees found in the LASSO and ℓ_1-norm minimization literature [33]. Throughout, $\widehat{I}$ denotes the set of coordinates selected by the coordinate selection scheme.

Theorem 1. *For the model specified in (2.1) and (2.2) and the algorithms specified in Table 1, assume that $p(n), n \to \infty$, $s(n)/n \to 0$, and $\log p(n) = o(n)$. Let $\epsilon \in (0,1)$. We have that*

a. For $i \in I^c$,

$$\max_{i \in I^c} \mathbb{P}\left(i \in \widehat{I}\right) \to 0,$$

b. For $i \in I$,

$$\min_{i \in I \; : \; |\theta u_i| > \beta_{crit}(1+\epsilon)} \mathbb{P}\left(i \in \widehat{I}\right) \to 1,$$

$$\max_{i \in I \; : \; |\theta u_i| < \beta_{crit}(1-\epsilon)} \mathbb{P}\left(i \in \widehat{I}\right) \to 0.$$

Here

$$\beta_{crit} = \begin{cases} \sigma C_U \frac{\sqrt{\log p}}{|\sum_k v_k|} & \text{for sum-SEPCA,} \\ \sigma \sqrt{C_2} \sqrt{\frac{\log ep}{\sqrt{n}}} & \text{for } \ell_2\text{-SEPCA,} \\ \sigma t_{\ell_1} & \text{for } \ell_1\text{-SEPCA,} \end{cases} \tag{6.1}$$

and t_{ℓ_1} satisfies the relation

$$\left(\sqrt{\frac{2}{\pi}} + C_1 \frac{\log ep}{\sqrt{n}}\right) = \frac{1}{n}\sqrt{\frac{2}{\pi}}\left[\sum_k \exp\left(-\left(\sqrt{n}\frac{(t_{\ell_1})v_k}{\sqrt{2}}\right)^2\right) + \right.$$
$$\left. \sqrt{\pi}\sum_k \left(\sqrt{n}\frac{(t_{\ell_1})v_k}{\sqrt{2}}\right) Erf\left(\sqrt{n}\frac{(t_{\ell_1})v_k}{\sqrt{2}}\right)\right].$$

We defer the proof to Appendix A.

Theorem 1 identifies a phase transition in the ability of the algorithms to accurately estimate the support of $\mathbf{u}$. Note that the analysis brings into sharp focus the dependence of β_{crit} on $\mathbf{v}$ for the ℓ_1- and sum-SEPCA algorithms, but not the ℓ_2-SEPCA algorithm. Consequently, we can expect the algorithms to perform differently depending on the structure of the underlying $\mathbf{v}$. It is important to note that the sparsity s of $\mathbf{u}$ is not a parameter in the thresholds and results.

It is also important to note that ℓ_2-SEPCA and ℓ_1-SEPCA do not rely on the equisigned character of $\mathbf{v}$. However, it is clear that the sum-SEPCA algorithm explicitly depends on the equisigned assumption.

6.1. Hamming loss

It is also possible to state the above results in terms of the Hamming loss for the support of $\mathbf{u}$, and prove consistency of the coordinate selection scheme by assuming that all the nonzero coordinates of $\mathbf{u}$ lie above a critical threshold. A detailed decision-theoretic analysis of variable selection under a sequence model with *i.i.d.* noise and with respect to the Hamming loss has recently been

carried out by [10]. Recall that the Hamming loss measures the number of elements in two sets that are different, so that here the loss between the true support I and the estimated support $\widehat{I}$ would be the size of the symmetric set difference of I and $\widehat{I}$. Let $d_H\left(I, \widehat{I}\right)$ denote the Hamming loss and assume that whatever algorithm we are using has a threshold β_{crit}. Then, for any $\epsilon \in (0, 1)$, we may write

$$
\begin{aligned}
\mathbb{E}\, d_H\left(I, \widehat{I}\right) \;=&\; \sum_{i \in I : |\theta u_i| > \beta_{crit}(1+\epsilon)} \mathbb{P}\left(i \notin \widehat{I}\right) + \sum_{i \in I : |\theta u_i| < \beta_{crit}(1-\epsilon)} \mathbb{P}\left(i \notin \widehat{I}\right) \\
&+ \sum_{i \in I : (1-\epsilon) \le |\theta u_i|/\beta_{crit} \le (1+\epsilon)} \mathbb{P}\left(i \notin \widehat{I}\right) + \sum_{i \notin I} \mathbb{P}\left(i \in \widehat{I}\right). \quad (6.2)
\end{aligned}
$$

We can then restate the results on coordinate selection in terms of the Hamming loss under a more restricted setting that assumes an exact form of sparsity of the vector **u**.

Corollary 1. *For the model specified in (2.1) and (2.2), and an algorithm specified in Table 1, assume that the conditions of Theorem 1 hold. Let the support I of* **u** *be estimated by $\widehat{I}$. Moreover, assume that the algorithm has a threshold β_{crit} (given in (6.1)) such that for a small, fixed $\epsilon_0 > 0$, the set*

$$
I_0 := \{j : |\theta u_j| > \beta_{crit}(1 + \epsilon_0)\} \tag{6.3}
$$

equals the set I. Then the expected Hamming loss satisfies

$$
\mathbb{E}\, d_H\left(I, \widehat{I}\right)/|I| \to 0. \tag{6.4}
$$

The proof of the corollary, given in Appendix A.1, follows from applying Theorem 1, with a more detailed enumeration of the sets and the inclusion probabilities, to each term of (6.2).

6.2. FDR-based algorithms

We may summarize the coordinate selection properties of the FDR refinements as follows:

Theorem 2. *For the model specified in (2.1) and (2.2) and the three FDR-controlling algorithms summarized in Algorithm 2, assume that $p(n), n \to \infty$, $s(n)/n \to 0$, and $\log p(n) = o(n)$. Let $\epsilon \in (0, 1)$. We have that*

a. For all three algorithms and $i \in I^c$,

$$
\max_{i \in I^c} \;\mathbb{P}\left(i \in \widehat{I}\right) \;\to 0,
$$

b. For the Higher Criticism-based algorithms and $i \in I$,

$$
\min_{i \in I \,:\, |\theta u_i| > \beta_{crit}(1+\epsilon)} \;\mathbb{P}\left(i \in \widehat{I}\right) \;\to 1,
$$

$$
\max_{i \in I \,:\, |\theta u_i| < \beta_{crit}(1-\epsilon)} \;\mathbb{P}\left(i \in \widehat{I}\right) \;\to 0.
$$

c. *For the FDR-SEPCA algorithm, uniformly over $i \in I$,*
   if $|\theta u_i| > \beta_{crit}(1 + \epsilon)$, coordinate i is selected;
   if $|\theta u_i| < \beta_{crit}(1 - \epsilon)$, coordinate i is not selected
   with probability tending to 1.

Here

$$\beta_{crit} = \begin{cases} \sigma\sqrt{\rho(\beta)}\frac{\sqrt{2\log p}}{\|\mathbf{v}\|_1} & \text{for HC-sum-SEPCA,} \\ \sigma\rho(\beta)\frac{2\log p}{\sqrt{n}} & \text{for HC-}\ell_2\text{-SEPCA,} \\ \sigma\left(1 - o(1)\right)\sqrt{\zeta}\frac{1+\sqrt{2\log(\nu p/\widehat{k})}}{\|\mathbf{v}\|_1} & \text{for FDR-SEPCA,} \end{cases} \qquad (6.5)$$

where $\zeta > 1$, $\nu > e$, and the FDR-SEPCA algorithm detects $\widehat{k}$ coordinates.

We defer the proof to Appendix B.

Once again, we see that the structure of the underlying $\mathbf{v}$ plays a role in the performance of the sum-based algorithms, but not for the ℓ_2-based HC-ℓ_2-SEPCA algorithm. Unlike in the FWER-controlling cases, the sparsity of $\mathbf{u}$ plays a (small) role here, via the constant $\rho(\beta)$ for the Higher Criticism-based methods and via $\widehat{k}$ for FDR-SEPCA. Moreover, ℓ_2-HC-SEPCA, like ℓ_2-SEPCA, does not make use of the equisigned nature of $\mathbf{v}$.

6.3. Higher ranks

In this work, we restrict our focus to the rank-1, equisigned setting. A natural question is are our results extensible to the higher rank setting?

The first point is concerned with the right singular vectors. To preserve orthogonality, we would need equisigned right-singular vectors $\mathbf{v}_i$ with disjoint supports. The second point is concerned with the left singular vectors. Our algorithms are based on thresholding row-statistics: it is possible that the union of supports of several sparse vectors is a relatively large set. The FWER-controlling algorithms (by design) are not sensitive to the increased supports, but the FDR-controlling algorithms are sensitive to this. Indeed, the decision boundaries for the FDR algorithms explicitly depend on the sparsity levels. Third point, once again, is concerned with the left singular vectors. It is possible that a sum-based statistic suffers from cancellations that decrease the size of the row-statistic. For example, in a rank-2 setting, if $\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \end{bmatrix}^T$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \end{bmatrix}^T$ and $\|\mathbf{v}_1\|_1$ and $\|\mathbf{v}_2\|_1$ have similar values and are both non-negative, the row-sum of the second row will be small. Note, however, that the ℓ_2-norm based methods do not suffer from this issue.

7. Simulations

To illustrate the relative powers of the six algorithms, we compute the theoretical limits on the sizes of detectable coordinates as a function of n. We use a unit-
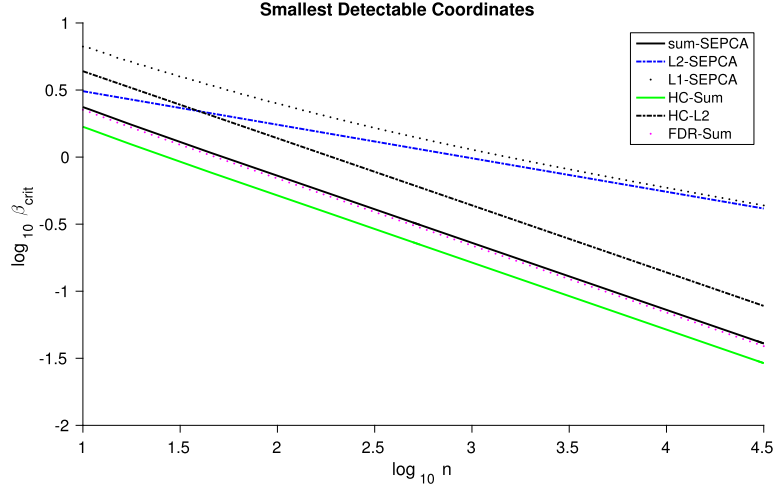
FIG 1. *This plot shows β_{crit} for all six algorithms for the* **v** *described in (7.1).*

norm, equisigned **v** such that

$$v_k \propto \exp\left(-5\frac{k}{n}\right)\left|\sin\left(4\frac{k}{n}\right)\right| \text{ for } 1 \leq k \leq n. \tag{7.1}$$

This choice of **v** has a 'rise and fall' sort of behavior, and is motivated by physical signals, e.g., chemical reactions or nerve signals in the brain. The value of β_{crit} is shown in Figure 1; for this choice of **v**, it is clear that the sum-SEPCA dramaticaly outperforms the other SEPCA variants in terms of size of the smallest detectable component. The FDR-SEPCA algorithm has similar performance to sum-SEPCA, and the HC-sum-SEPCA algorithm has the strongest performance.

In Figure 2, we plot the estimation error as a function of n and θ for all six algorithms. We also include results for the SVD and competing algorithms TPower [41] and ITSPCA [25]. In the simulations, we fix $p = 1000$ and vary n, since the dependence in p in the thresholds is logarithmic, whereas that in n is not. The left singular vector **u** is chosen to be the vector with 1 in the first coordinate and 0 elsewhere. We fix the noise variance σ^2 at 1, so that θ^2 is the eigen-SNR. The results should be interpreted as follows. For the particular **v** chosen here, we expect HC-sum-SEPCA to have the lowest detectable limit, and ℓ_1-SEPCA to have the largest. This behavior is confirmed. Moreover, the sum-based algorithms offer a slight strengthening of both ITSPCA and TPower. Importantly, note that the sum-based algorithms explicitly take advantage of the equisigned nature of **v**: that is, algorithms that explicitly use the equisigned property outperform algorithms that do not (the ℓ_2 and ℓ_1 algorithms, as well as ITSPCA and TPower).

We repeat our simulations for a $\mathbf{u} \in \mathbb{R}^p$ with $\sqrt{p}$ non-zero coordinates (of equal size) and the same **v**, as seen in Figure 3. We find similar conclusions as in
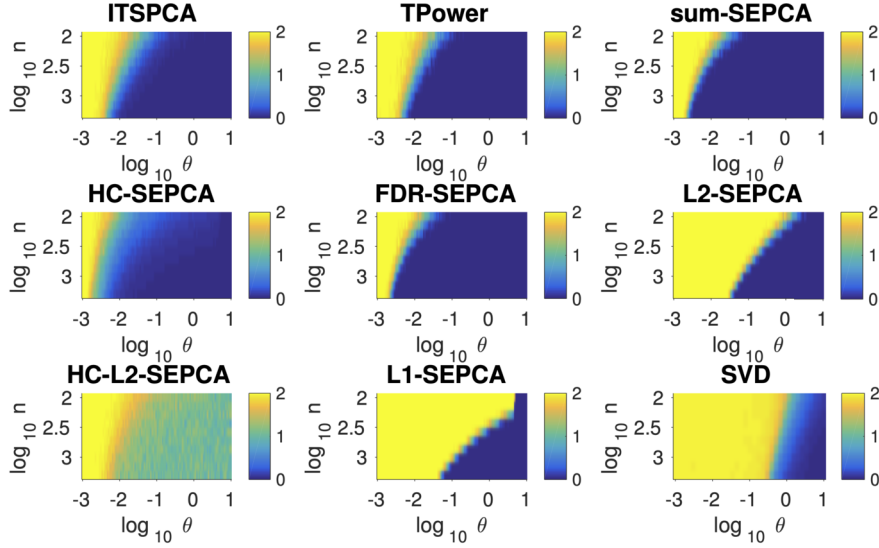
FIG 2. *The plots show the empirical estimation error for all six algorithms for the* **u** *with one non-zero coordinate and the* **v** *described in* (7.1). *We include results from TPower, ITSPCA and the SVD for comparison.*

Figure 2, where the sum-based algorithms offer a strengthening over ITSPCA and TPower; the ℓ_2-norm based algorithms do not perform well. Note that a sparsity of $\sqrt{p}$ is at the limit/valid edge for the higher criticism-based methods, but that these methods still perform well.

7.1. Comments on the FDR-controlling procedures

The Higher Criticism for the χ_n^2-variates 'pushes back' the phase transition between detecting nothing and something to a lower value of θ relative to the ℓ_2-SEPCA algorithm, but is still less powerful than any of the sum-based algorithms. Moreover, even above the phase transition, the ℓ_2-SEPCA algorithm may be preferable, as the error is increased by unacceptably many false positives.

The Higher Criticism procedure for the sum statistic has the lowest phase transition point and hence the highest power. Its transition is more gradual than the penalized FDR thresholding procedure and sum-SEPCA, which have roughly the same performance in this simulation.

7.2. An example where ℓ_2-based algorithms outperform sum-based algorithms

Sum-SEPCA has a β_{crit} that depends on **v**. Looking at the form in (6.1), if $\| \mathbf{v} \|_1$ is smaller than $n^{1/4}$, we would expect ℓ_2-SEPCA to detect a smaller coordinate
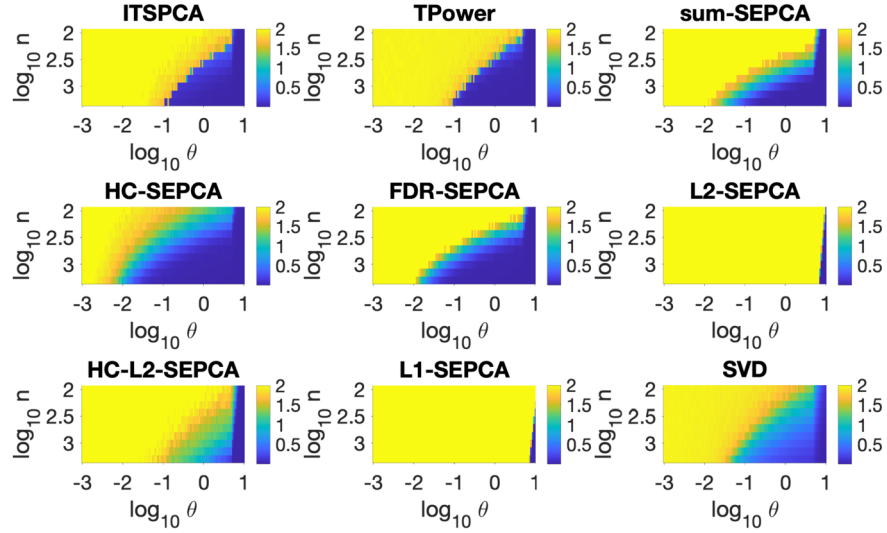
FIG 3. *The plots show the empirical estimation error for all six algorithms for the* **u** *with* $\sqrt{p}$ *non-zero coordinates and the* **v** *described in (7.1). We include results from TPower, ITSPCA and the SVD for comparison.*

size. Vectors with smaller coordinates have a smaller ℓ_1-norm, i.e., one that is closer to their ℓ_2-norm. Hence, if we choose

$$v_k \propto \frac{1}{k^2} \text{ for } 1 \leq k \leq n, \tag{7.2}$$

we expect sum-SEPCA to have worse performance relative to ℓ_2-SEPCA. Figures 4 and 5 confirm this expectation. The FDR refinements perform poorly. It should be noted, however, that TPower and ITSPCA retain their performance. This choice of **v** effectively corresponds to a very small value of n: the majority of coordinates are tiny in size and buried beneath noise regardless of the value of θ. If we 'corrected' the scenario and used a smaller n and a subset of **v**, we would be in a situation closer to that given in (7.1).

7.3. A video data example

We conclude our sequence of examples with a real data study. This example is motivated by the problem of foreground-background separation in videos. Consider a grayscale video of stars twinkling against a black background [36]. Our goal is to estimate the locations of the stars: by reshaping the video, we may treat each frame as a vector and hence treat the video as a sparse matrix. Only a few locations have a star and are hence non-zero. The scale of the video pixels is between 0 and 255. We examine the top-left 72×64 pixels for 89 frames, as shown in Figure 6a. In Figure 6b, we plot the singular values of the video matrix. The first singular value stands out strongly against the rest, and at
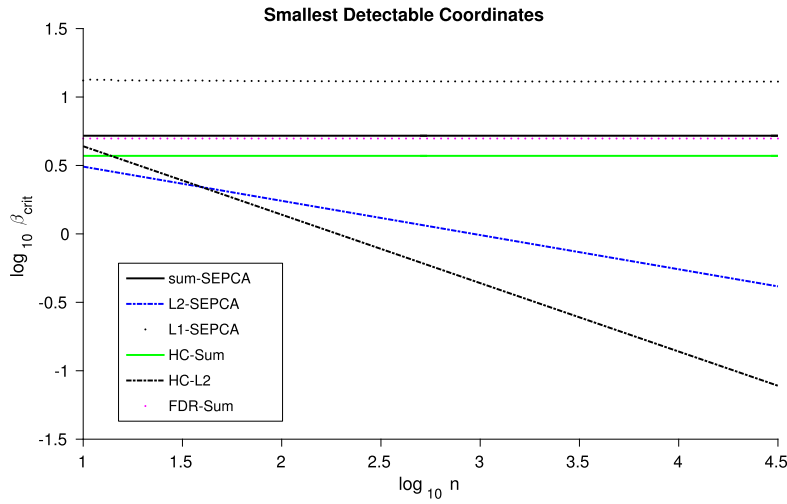
FIG 4. *This plot shows β_{crit} for all six algorithms for the* **v** *described in (7.2).*
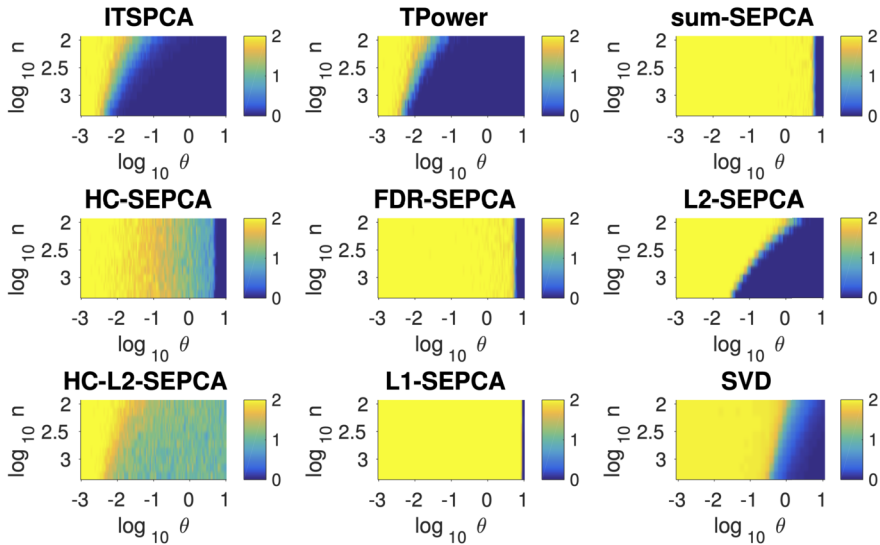

FIG 5. *The plots show the empirical estimation error for all six algorithms for the* **u** *and* **v** *described in (7.2). We include results from TPower, ITSPCA and the SVD for comparison.*

most two more singular values are well-separated from the bulk. This structure suggests that our rank-1 based approach is well suited to this problem.

We add Gaussian noise of variance σ^2 and study the True Positive Rates (TPR) and False Discovery Rates (FDR) across all algorithms and across different values of σ. In Figure 7, we show the results of our simulations. In terms

(a) The image shows the mean intensity of pixels from the top-left 72×64 pixels for 89 frames. White indicates the presence of a star.
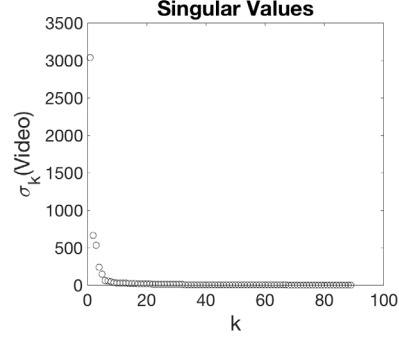
(b) The plots shows the singular values of the video data. The spacing suggest a low-rank-plus-noise structure.

FIG 6. *Video Example Figures.*

of the TPR, everything other than the SVD has a similar performance, while the test-statistic SEPCA-based algorithms enjoy the best performance in terms of the FDR. In Figure 8 we zoom in on the top-right three stars and show how the algorithms perform as noise increases. Here, we see that the behavior alluded to in the TPR/FDR results actually occurs in the video.

8. A geometric view: Which algorithm to use?

We have stated detectability results for each algorithm in Section 6 and provided a numerical verification and comparison in Section 7. In this section, we wish to analytically compare the algorithms. In particular, we have seen that the right singular vector $\mathbf{v}$ plays a critical role in the detectability and estimability of $\mathbf{u}$, and we will characterize this behavior carefully.

In this section, will use the following notational convenience: we absorb (θu_i) into $\mathbf{v} \in \mathbb{R}^n$, and write the detectability of coordinates in terms of $\mathbf{v}$. That is, if $\mathbf{v}^T$ is a row of $\mathbf{X}$, we specify when that row is selected. Moreover, we take $\sigma = 1$ for simplicity.

There are two 'classes' of detectability: in terms of $\|\mathbf{v}\|_1$ and in terms of $\|\mathbf{v}\|_2$. The sum-, HC-sum, and FDR-SEPCA algorithms select a coordinate if $|\sum_k v_k| = \|\mathbf{v}\|_1$ is large enough for a $\mathbf{v}$ in the orthant with all non-negative or all non-positive coordinates. Geometrically, the vector $\mathbf{v}$ is selected if it is 'outside' a hyperplane with a normal vector proportional to the vector of all 1s. The ℓ_1-SEPCA algorithm is similar, as it selects a coordinate when $\|\mathbf{v}\|_1$ is large
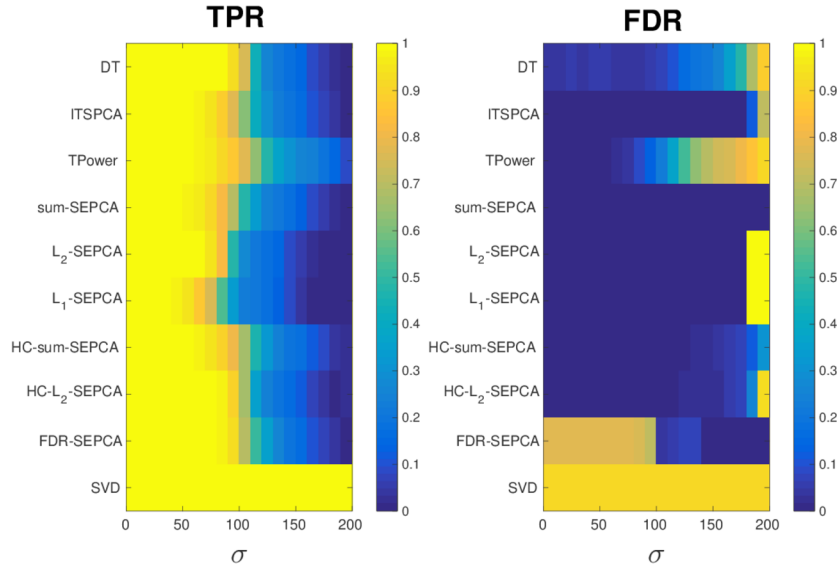
FIG 7. *The left plot shows the True Positive Rate of the various algorithms as a function of the noise level σ. The right plot shows the False Discovery Rates.*

enough, or if $\mathbf{v}$ lies outside an ℓ_1-ball of some radius. The connection between the previous three algorithms and ℓ_1-SEPCA comes from noting that the faces of an ℓ_1-ball are sections of hyperplanes with normal vectors proportional to a vector of ± 1s. Finally, the ℓ_2- and HC-ℓ_2-SEPCA algorithms select a coordinate when $\| \mathbf{v} \|_2$ is large enough. I.e., when $\mathbf{v}$ lies outside some ℓ_2-ball.

Our goal in this section is to derive comparisons between the six algorithms. Specifically, for a given vector $\mathbf{v}$, which algorithm will have the greatest detection ability (we are, for the moment, only concerned with maximizing power)? Note that when $\mathbf{v}$ has a large norm, it does not matter which algorithm is used. Questions only arise when $\| \mathbf{v} \|_1$ or $\| \mathbf{v} \|_2$ are relatively small and are close to the thresholds.

8.1. Intersection of a hyperplane and a hypersphere

We may think of the ℓ_1 ball as a hyperplane when restricted to a single orthant. If a hypersphere of radius r intersects a hyperplane with a normal vector proportional to the vector of all ± 1s and minimum distance to the origin of $r - h$, a hyperspherical cap of height h is formed: see Figure 9 for a simple illustration. Geometrically, a right triangle is formed, with hypotenuse r and leg $r - h$. Hence, the angle between the center of the cap and the edge is:

$$\theta_{lim} = \cos^{-1} \frac{r - h}{r}. \tag{8.1}$$

(a) $\sigma = 0$

(b) $\sigma = 0.3$
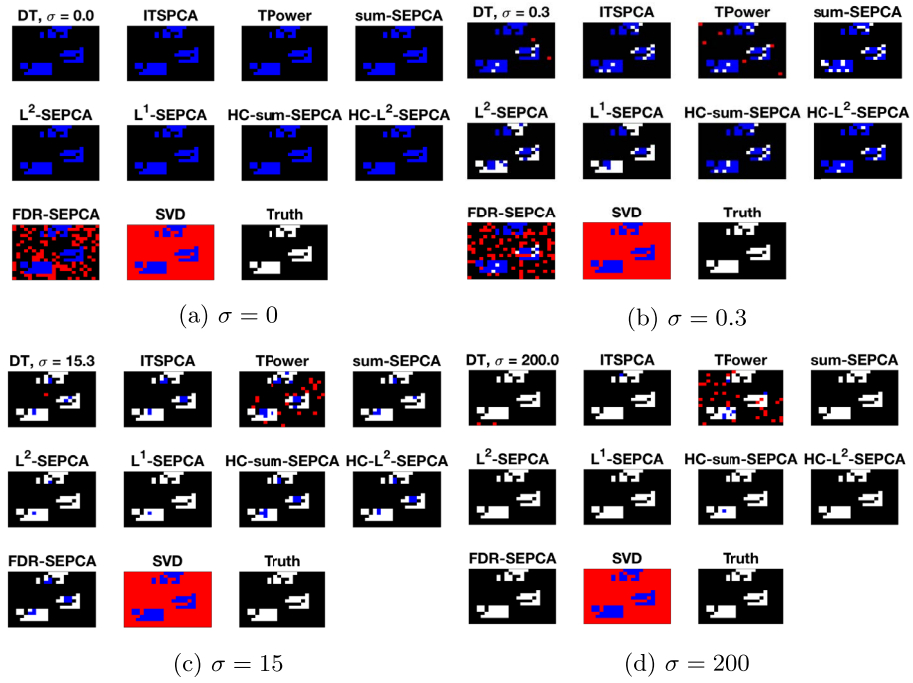
(c) $\sigma = 15$

(d) $\sigma = 200$

FIG 8. *A zoomed-in view of the three top-right stars in the video example. White indicates a false negative (missed star), Red a false positive (a guessed pixel where there was nothing), and Blue a true positive (correctly identified pixel).*
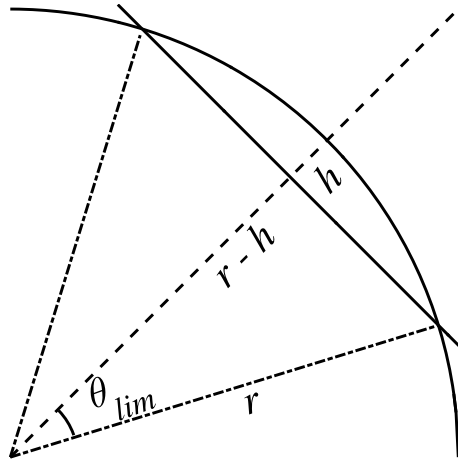

FIG 9. *A spherical cap in $\mathbb{R}^2$.*

<div align="center">

TABLE 2
A summary of the six algorithms.

</div>

Algorithm	Threshold	Geometric Quantity
sum	$\| \mathbf{v} \|_1 \geq C_U \sqrt{\log p}$	$r - h = C_U \sqrt{\frac{\log p}{n}}$
HC-sum	$\| \mathbf{v} \|_1 \geq \sqrt{2\rho(\beta) \log p}$	$r - h = \sqrt{2\rho(\beta) \frac{\log p}{n}}$
FDR	$\| \mathbf{v} \|_1 \geq (1 - o(1)) \sqrt{\zeta} \left(1 + \sqrt{2 \log \left(\nu p / \widehat{k} \right)} \right)$	$r - h = \dfrac{(1 - o(1)) \sqrt{\zeta} \left(1 + \sqrt{2 \log \left(\nu p / \widehat{k} \right)} \right)}{\sqrt{n}}$
ℓ_1	$\| \mathbf{v} \|_1 \geq t_{\ell_1}$	$r - h = t_{\ell_1}$
ℓ_2	$\| \mathbf{v} \|_2 \geq \sqrt{C_2} \sqrt{\frac{\log ep}{\sqrt{n}}}$	$r = \sqrt{C_2} \sqrt{\frac{\log ep}{\sqrt{n}}}$
HC-ℓ_2	$\| \mathbf{v} \|_2 \geq 2\rho(\beta) \frac{2 \log p}{\sqrt{n}}$	$r = 2\rho(\beta) \frac{2 \log p}{\sqrt{n}}$

It is sufficient to guarantee that

$$0 \leq \frac{r - h}{r} \leq 1$$

for the hyperspherical cap to exist. Moreover, a vector $\mathbf{v}$ has a direction contained inside the cap when the angle between $\mathbf{v}$ and the vector of ± 1 in the orthant containing $\mathbf{v}$ is smaller than θ_{lim}. In other words, defining the angle for a vector $\mathbf{v}$ as

$$\theta(\mathbf{v}) = \cos^{-1} \frac{\| \mathbf{v} \|_1}{\| \mathbf{v} \|_2 \sqrt{n}}, \tag{8.2}$$

we need $\theta(\mathbf{v}) \leq \theta_{lim}$.

8.2. Comparison: ℓ_2-based versus sum-based statistics

We begin with a summary of the performance of each individual algorithm in Table 2. We first compare ℓ_2-SEPCA and then compare HC-ℓ_2-SEPCA with sum-, HC-sum, FDR-SEPCA in Tables 3 and 4. In our comparisons, we consider when the hyperspherical cap exists and give the angle of the cap. These are routine calculations, so we omit the details. We also omit ℓ_1-SEPCA from our comparisons, as we lack a closed-form expression for t_{ℓ_1}.

Note that the existence of this cap is a proxy for the equivalent statement that there exist vectors for which the sum-based algorithms are more powerful than the ℓ_2-based algorithm. While this existence is not the same as attributing uniformly greater power to the sum-based algorithms relative to the ℓ_2-based algorithm, the cap *not* existing is equivalent to the ℓ_2-based algorithm having uniformly greater power.

Essentially, we observe that for n and p sufficiently large, the cap will exist. Moreover, for $\mathbf{v}$ that is sufficiently dense ($\| \mathbf{v} \|_1$ is sufficiently large), θ_{lim} will lie inside the cap. Hence, in these situations we would prefer a sum-based algorithm over an ℓ_2-based algorithm.

TABLE 3
The relative performance of ℓ_2-SEPCA.

Algorithm	$\cos\theta_{lim}$	Cap exists if
sum	$C_U/\sqrt{C_2}\sqrt{\frac{\log p}{(1+\log p)\sqrt{n}}}$	$n \geq C_U^4/C_2^2, p \geq 1$
HC-sum	$\sqrt{2\rho(\beta)/C_2}\sqrt{\frac{\log p}{(1+\log p)\sqrt{n}}}$	$n \geq 1, p \geq 1$
FDR-sum	$\frac{1+\sqrt{2\log\nu p/\widehat{k}}}{\sqrt{C_2}\sqrt{n}\log\nu p}$	$p \geq 11, n \geq 1$

TABLE 4
The relative performance of HC-ℓ_2-SEPCA.

Algorithm	$\cos\theta_{lim}$	Cap exists if
sum	$\frac{C_U}{2\rho(\beta)\sqrt{\log p}}$	$p \geq \exp\left(\frac{C_U^2}{4\rho(\beta)^2}\right)$
HC-sum	$[2\rho(\beta)\log p]^{-1}$	$p \geq \exp\left(\frac{1}{2\rho(\beta)}\right)$
FDR-sum	$\frac{1+\sqrt{2\log\nu p/\widehat{k}}}{2\rho(\beta)\log p}$	$\log p \geq$ $\frac{1}{4\rho(\beta)^2}\left(1+2\rho(\beta)\right.$ $\left.+\sqrt{8\rho(\beta)^2+4\rho(\beta)+1}\right)$

8.3. HC-ℓ_2-SEPCA versus ℓ_2-SEPCA

Now, we consider when HC-ℓ_2-SEPCA is more powerful than ℓ_2-SEPCA. The ratio of the radii is given by

$$\frac{2\rho(\beta)}{\sqrt{C_2}n^{1/4}}\frac{\log p}{\sqrt{1+\log p}}. \tag{8.3}$$

If this ratio is smaller than 1, HC-ℓ_2-SEPCA is more powerful than ℓ_2-SEPCA. Note that the quantity

$$\frac{2\sqrt{2}}{\sqrt{C_2}}\sqrt{\frac{\log p}{\sqrt{n}}}$$

is an upper bound for (8.3), so that if

$$\frac{\log p}{\sqrt{n}} < \frac{e}{4\sqrt{2}},$$

the original ratio is smaller than 1 and HC-ℓ_2-SEPCA is preferable to ℓ_2-SEPCA.

8.4. Comparing the sum-based algorithms

Finally, we compare sum-, HC-sum-, and FDR-SEPCA. First, the ratio of the thresholds for HC-sum- and sum-SEPCA is

$$\frac{\sqrt{2\rho(\beta)}}{C_U}. \tag{8.4}$$

Noting that $\rho(\beta) \leq 1$ and that $C_U \geq \sqrt{2} + 1/3\sqrt{2}$, it is clear that this ratio is always smaller than 1 so that HC-sum-SEPCA is a strict improvement on sum-SEPCA.

Next, we compute the ratio of the thresholds for FDR- and sum-SEPCA:

$$\frac{1 + \sqrt{2\log \nu p/\widehat{k}}}{C_U \sqrt{\log p}}. \tag{8.5}$$

Using the lower bound on C_U, we find that if $\widehat{k} \geq 11$ (and $p \geq \widehat{k}$, naturally), FDR-SEPCA is always more powerful than sum-SEPCA. For smaller values of $\widehat{k}$, for sufficiently large values of p, the ratio will be smaller than 1.

Lastly, we compare FDR-SEPCA to HC-sum-SEPCA, wherein the ratio of the thresholds is (FDR to HC-sum):

$$\frac{1 + \sqrt{2\log \nu p/\widehat{k}}}{\sqrt{2\rho(\beta)}\sqrt{\log p}}. \tag{8.6}$$

Because of involvement of $\rho(\beta)$, this quantity is hard to analyze. If in an oracle manner, FDR-SEPCA obtained $\widehat{k}$ correctly as $p^{1-\beta}$, we would find that this ratio is always larger than 1 for $p > 1$. That is if $\widehat{k}$ assumes the the correct value, HC-sum-SEPCA is more powerful than FDR-SEPCA. Alternatively, we can note that $\rho(\beta) \in (0,1]$ and ask when the ratio is larger than 1. Based on the ratio above, we can see that in the following scenarios

$$\begin{cases} \widehat{k} = 1 & \text{and} \quad p > 1, \\ 2 \leq \widehat{k} \leq 18 & \text{and} \quad p \geq \widehat{k} \,(\text{always}), \\ \widehat{k} \geq 19 & \text{and} \quad \log p \geq \frac{1}{8}\left(4(\log \widehat{k})^2 - 4\log \widehat{k} + 1\right), \end{cases} \tag{8.7}$$

HC-sum-SEPCA is more powerful than FDR-SEPCA.

To summarize, we prefer the FDR-controlling alternatives to sum-SEPCA, but depending on the output of FDR-SEPCA, HC-sum-SEPCA may be more powerful. However, as the simulations in Section 7 revealed (see Figure 2), the number of false positives with HC-sum-SEPCA may be higher than with FDR-SEPCA.

8.5. Overall message

We have seen that for n and p sufficiently large and $\mathbf{v}$ that is sufficiently dense (in the sense of $\| \mathbf{v} \|_1$ being large), a sum-based statistic and algorithm leads to better performance. This is expected behavior, as by using a sum-based method, we are taking advantage of the equisigned nature of $\mathbf{v}$. Moreover, within the class of sum-based algorithms, controlling the FDR leads to greater power, as expected. It is difficult to clearly identity which of HC-sum- and FDR-SEPCA will have the greatest power, and the end result may come down to a practitioner's tolerance for false discoveries.

9. Conclusions

We have considered the setting where the left singular vector of the underlying rank one signal matrix plus noise data matrix is assumed to be sparse and the right singular vector is assumed to be equisigned. We have proposed six different SEPCA algorithms for estimating the sparse principal component based on different decision statistics and provided sparsistency conditions for the same. Our analysis reveals conditions where a coordinate selection scheme based on a sum-based decision statistic outperforms schemes that utilize the ℓ_1 and ℓ_2 decision statistics. Thereby, the proposed algorithm outperforms known schemes such as *diagonal thresholded PCA* [21] in terms of estimation of the singular vectors associated with the rank-1 component. We have derived lower bounds on the size of detectable coordinates of the principal left singular vector, utilized these lower bounds to derive lower bounds on the worst-case risk and verified our findings with numerical simulations. Finally, we have discussed the results of our simulations analytically, by providing a geometric interpretation of the differences in power among the algorithms.

We note that while we have stated our results for Gaussian noise with identity covariance, we can extend the FWER-controlling results to any log-concave noise distribution, and the FDR-controlling procedures to Gaussian noise with certain non-identity covariances. Additionally, another way to view this work is that it proposes a two-stage procedure/framework for sparse PCA based around hypothesis testing of statistics associated to each row. Some natural extensions would be the inclusion or consideration of other testing frameworks, e.g., that in [28], where knowledge of the size of coordinates is taken into account.

Appendix A: Proof of Theorem 1

a. Note that
$$\mathbb{P}\left(T_i \geq \tau\right) \leq \mathbb{P}\left(\max_{j \in I^c} T_j \geq \tau\right)$$

for $i \in I^c$. Taking the maximum over the left-hand side and noting that the right-hand side has limit zero yields the result. This follows from (3.1). ☐

b. We consider when true positives occur with probability approaching 1. We want to find the smallest coordinate (θu_i) such that the following probability approaches 1:

$$\mathbb{P}\left(T_i > \tau_{n,p}\right) = \mathbb{P}\left(\frac{T_i - \mathbb{E}T_i}{\sqrt{\operatorname{Var} T_i}} > \frac{\tau_{n,p} - \mathbb{E}T_i}{\sqrt{\operatorname{Var} T_i}}\right). \tag{A.1}$$

Note that if $(\tau_{n,p} - \mathbb{E}T_i)$ is negative and not tending to zero as n grows, and if the variance of T_i decays to zero as n grows, the quantity

$$\frac{\tau_{n,p} - \mathbb{E}T_i}{\sqrt{\operatorname{Var} T_i}} \tag{A.2}$$

tends toward negative infinity. Hence, we will specify conditions so that Var T_i decays to zero as n grows and then compute when a coordinate is detectable by considering when $\tau_{n,p}$ is strictly less than $\mathbb{E}T_i$. For brevity, we omit the computations in solving $\tau_{n,p} < \mathbb{E}T_i$ for $|\theta u_i|$ and present verifications that the variance of T_i has limit 0. These results show that above the *decision boundary*, we have uniform detection.

In sum-SEPCA, T_i is a Gaussian random variable with mean $\frac{(\theta u_i)}{\sqrt{n}} \sum_k v_k$ and variance $\frac{\sigma^2}{n}$. Since σ does not grow with n, Var T_i always decays to zero.

In ℓ_2-SEPCA, T_i has

$$\mathbb{E}T_i = (\theta u_i)^2 + \sigma^2$$

and

$$\text{Var } T_i = \frac{2\sigma^2}{n} \left(\sigma^2 + 2 (\theta u_i)^2 \right).$$

Since σ and θ are fixed, the variance always decays to 0.

Let $x_{i,k} = \left(\sqrt{n} \frac{v_k(\theta u_i)}{\sigma} \right)$. In ℓ_1-SEPCA, T_i has

$$
\begin{aligned}
\text{Var } T_i = {} & \frac{\sigma^2}{n^2} \sum_k x_{i,k}^2 \left(1 - \left(\text{Erf } \left(\frac{x_{i,k}}{\sqrt{2}} \right) \right)^2 \right) \\
& + \frac{\sigma^2}{n} \left(1 - \frac{2}{n\pi} \sum_k \exp \left(-x_{i,k}^2 \right) \right) \\
& - 2\sqrt{\frac{2}{\pi}} \frac{\sigma^2}{n^2} \sum_k x_{i,k} \exp \left(-\frac{x_{i,k}^2}{2} \right) \text{Erf } \left(\frac{x_{i,k}}{\sqrt{2}} \right),
\end{aligned}
$$

which is less than or equal to

$$\frac{(\theta u_i)^2}{n} \sum_k v_k^2 + \frac{\sigma^2}{n} + 2\sqrt{\frac{2}{\pi}} \frac{\sigma}{n\sqrt{n}} | (\theta u_i) | \sum_k |v_k|. \tag{A.3}$$

Since $\| \mathbf{v} \|_2 = 1$, the variance of T_i has limit 0. Because we cannot solve the inequality $\tau_{n,p} < \mathbb{E}T_i$ analytically, we leave the bound in the form given previously. □

In the proof above, note that if $(\tau_{n,p} - \mathbb{E}T_i)$ is positive and not tending to zero as n grows, the quantity in (A.2) tends to positive infinity when the variance decays to zero. Hence, modifying the proof by solving $\tau_{n,p} > \mathbb{E}T_i$ for $|\theta u_i|$ yields when a coordinate is not detectable with probability approaching 1: i.e., when $|\theta u_i|$ is smaller than the values given in (6.1). □

A.1. Proof of Corollary 1

The first three terms of (6.2) are characterized by Theorem 1(b) and by noting that $I_0 = I$ in the corollary. The last term can be characterized as follows. In

particular, for an algorithm with row test statistics T_i and a threshold τ from Table 1, we may write

$$\sum_{i \notin I} \mathbb{P}\left(i \in \widehat{I}\right) = \sum_{i \notin I} \mathbb{P}\left(T_i \geq \tau\right). \tag{A.4}$$

For the ℓ_2- and ℓ_1-SEPCA algorithms, we may reindex the T_i according to their order statistics $T_{(i)}$, where

$$\left|T_{(1)}\right| \geq \left|T_{(2)}\right| \geq \cdots \geq \left|T_{(p-s)}\right|,$$

and write

$$\sum_{i \notin I} \mathbb{P}\left(i \in \widehat{I}\right) = \sum_{i=1}^{p-s} \mathbb{P}\left(T_{(i)} \geq \tau\right). \tag{A.5}$$

Note that there are $p - s$ null entries. For the ℓ_2- and ℓ_1-SEPCA algorithms, we have that (as a consequence of [23, Thm. 3])

$$\mathbb{P}\left(T_{(i)} \geq \tau\right) \leq \left(\frac{1}{ep}\right)^{\sqrt{i}},$$

so that

$$\sum_{i \notin I} \mathbb{P}\left(i \in \widehat{I}\right) = \sum_{i=1}^{p-s} \mathbb{P}\left(T_{(i)} \geq \tau\right) \leq \sum_{i=1}^{p-s} \left(\frac{1}{ep}\right)^{\sqrt{i}}. \tag{A.6}$$

The right-hand side of (A.6) has limit zero, as needed.

For the sum-SEPCA algorithm, from (3.6), it follows that there exists a non-zero constant $\epsilon > 0$ such that the threshold τ satisfies

$$\tau = \sigma\left(\sqrt{2} + \epsilon\right)\sqrt{\frac{\log p}{n}}.$$

In particular, from (3.6), we have that

$$\epsilon = \frac{1 + \sqrt{\log p}/3}{\sqrt{2}\mathrm{Erf}^{-1}\left(1 - 1/p\right)},$$

where for $p > 1$,

$$\frac{1}{3\sqrt{2}} < \epsilon < 2.$$

Hence, for any T_i such that $i \notin I$,

$$\mathbb{P}\left(T_i \geq \tau\right) \leq \exp\left(-\frac{\left(\sqrt{2} + \epsilon\right)^2}{2}\log p\right) = p^{-\left(1 + \sqrt{2}\epsilon + \epsilon^2/2\right)},$$

where we have used a Gaussian tail bound [7, Sec. 2.3]. Then,

$$\sum_{i \notin I} \mathbb{P}\left(i \in \widehat{I}\right) = \sum_{i \notin I} \mathbb{P}\left(T_i \geq \tau\right)$$

$$\leq (p - s) \, p^{-\left(1 + \sqrt{2}\epsilon + \epsilon^2/2\right)} \tag{A.7}$$

$$\leq p^{-\left(\sqrt{2}\epsilon + \epsilon^2/2\right)}.$$

Since ϵ is larger than $1/3\sqrt{2}$, the right-hand side of (A.7) is upper bounded by $p^{-13/36}$, which has limit zero, as desired.

Appendix B: Proof of Theorem 2

B.1. Size of detectable coordinates

B.1.1. Sum: HC-SEPCA

If $\mathbf{v}$ is equisigned, summing across the rows of $\mathbf{X}$ yields a normally distributed quantity with mean $(\theta u_i) \| \mathbf{v} \|_1$ and variance σ^2. Dividing by σ and adopting the notation of HC, we have that under the alternative hypothesis, $\mu_i = \sqrt{2r \log p}$, so that

$$r = \left(\frac{|\theta u_i| \| \mathbf{v} \|_1}{\sqrt{2 \log p}}\right)^2.$$

Rearranging the inequality $r > \rho(\beta)$ yields

$$|\theta u_i| > \sigma \sqrt{\rho(\beta)} \frac{\sqrt{2 \log p}}{\| \mathbf{v} \|_1}. \tag{B.1}$$

Note that sum-SEPCA can detect coordinates of size

$$|\theta u_i| > \sigma C_U \frac{\sqrt{\log p}}{\| \mathbf{v} \|_1}. \tag{B.2}$$

However, C_U is strictly larger than $\sqrt{2} + 1/(3\sqrt{2})$. Thus, using HC yields a threshold of the same order, but with a strictly smaller scaling.

B.1.2. Sum of squares: HC-ℓ_2-SEPCA

If we sum the squares of the entries of rows of $\mathbf{X}$, abusing notation slightly and using $\mathcal{N}(\mu, \sigma^2)$ to indicate a Gaussian random variable with mean μ and variance σ^2, the statistic for the ith coordinate is of the form

$$\sum_{k=1}^{n} \left(\frac{\sigma}{\sqrt{n}} \mathcal{N}\left(\frac{\theta u_i}{\sigma} v_k \sqrt{n}, 1\right)\right)^2.$$

Assuming oracular knowledge of σ, the statistic

$$\frac{n}{\sigma^2} \sum_k X_{ik}^2$$

places us in the setting of (4.5). The non-centrality parameter δ is given by

$$\delta = \sqrt{\sum_{k=1}^n \left(\frac{\theta u_i}{\sigma} v_k \sqrt{n}\right)^2} = \left|\frac{\theta u_i}{\sigma}\right| \sqrt{n}.$$

Setting $\delta = 2r \log p$ and solving $r > \rho(\beta)$ yields

$$|\theta u_i| > \sigma \rho(\beta) \frac{2 \log p}{\sqrt{n}}. \tag{B.3}$$

We have that ℓ_2-SEPCA can detect coordinates with

$$|\theta u_i| > \sigma \sqrt{e\sqrt{2}} \sqrt{\frac{1 + \log p}{\sqrt{n}}}. \tag{B.4}$$

Using HC offers a significant improvement over ℓ_2-SEPCA. However, we also expect HC with the χ_n^2 statistic to have a smaller detectable coordinate: $\|\mathbf{v}\|_1 \leq \sqrt{n}$, so that for fixed β and p, the threshold in (B.1) is asymptotically larger than that in (B.3) (but potentially of the same order). This result is strange in context of the non-FDR results. In any case, HC improves on ℓ_2-SEPCA.

B.1.3. FDR-SEPCA

Recall that taking sums across the rows of $\mathbf{X}$, we obtain a vector $\mathbf{y}$ where $y_i = \mu_i + \sigma z_i$, with $\mu_i = (\theta u_i)\|\mathbf{v}\|_1$. Moreover, we have noted that

$$t_k \approx \sqrt{\zeta}(1 + \sqrt{2 \log(\nu p/k)}),$$

where t_k is the level at which $\mathbf{y}$ is thresholded. It follows that, entries of $\mathbf{y}$ that are of size at least

$$y_i > (1 - o(1))\sqrt{\zeta}\sigma\left(1 + \sqrt{2 \log(\nu p/\widehat{k})}\right)$$

are selected, or, since $\mu_i = (\theta u_i)\|\mathbf{v}\|_1$ (when v is equisigned), if we select $\widehat{k}$ coordinates, we expect to detect

$$\begin{aligned} |\theta u_i| &> (1 - o(1))\sqrt{\zeta}\sigma \frac{\left(1 + \sqrt{2 \log(\nu p/\widehat{k})}\right)}{\|v\|_1} \\ &= O\left(\sigma \frac{\sqrt{2 \log\left(\nu p/\widehat{k}\right)}}{\|\mathbf{v}\|_1}\right). \end{aligned} \tag{B.5}$$

Relative to HC and sum-SEPCA, the gain here is found when there are many smaller coordinates of $\mathbf{u}$ and $\widehat{k}$ is large.

B.2. Proofs for the higher criticism-based methods

a. From (2.8) in [17],

$$\mathbb{P}\left(T_i \geq \tau\right) \leq \mathbb{P}\left(\max_{j \in I^c} T_j \geq \tau\right)$$

has limit zero. □

b. Let $I_1 \subseteq I$ be the set of coordinates with signal larger than the detection limit ($i \in I$ such that $|\theta u_i| > \beta_{crit}(1+\epsilon)$), and let $I_2 \subseteq I$ contain the rest of the coordinates ($i \in I$ such that $|\theta u_i| < \beta_{crit}(1-\epsilon)$). By Theorem 1 in [2], the asymptotic power for detecting signals below the detection limit is one, and that for signals below the limit is zero. Hence, for $i \in I_1$,

$$\min_{i \in I \,:\, |\theta u_i| > \beta_{crit}(1+\epsilon)} \mathbb{P}\left(i \text{ selected}\right) \to 1,$$

and for $i \in I_2$,

$$\max_{i \in I \,:\, |\theta u_i| < \beta_{crit}(1-\epsilon)} \mathbb{P}\left(i \text{ selected}\right) \to 0.$$

As with Theorem 1, we omit the computation of β_{crit}, as it follows from the discussion in Section 4.1. □

B.3. FDR-SEPCA

The details of these computations are in Appendix D.1, so we will summarize the properties here.

a. The choice of $\nu = 2^{1/\omega}$ controls the FDR at level ω [22]. Choosing $\omega = \omega(p) \to 0$ as $p \to \infty$ leads to an asymptotic FDR of zero. I.e., for $i \in I^c$,

$$\max_{i \in I^c} \quad \mathbb{P}\left(i \text{ selected}\right) \to 0. \quad \square$$

b. Noting that the consistency of estimating the mean vector $\boldsymbol{\mu} = (\theta\|v\|_1)\,\mathbf{u}$ encompasses the estimation of the support of $\mathbf{u}$, risk bounds for the estimation of $\boldsymbol{\mu}$ yield the result. To be precise, if the expected risk $\mathbb{E}\|\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}\|_2^2 \leq B$ for some bound B, we expect to detect coordinates of size larger than B and to not detect those smaller than B. □

Appendix C: Risk bounds under ℓ_q sparsity

In this section, we simultaneously generalize our setting to approximate sparsity, and specify the risk lower-bounds. We omit the ℓ_1-SEPCA algorithm from consideration.

Let $\mathbf{u} \in \mathbb{R}^p$ have unit ℓ_2-norm and belong to an ℓ_q ball with radius C for $q \in (0, 2]$. I.e.,

$$\sum_{i=1}^{p} |u_i|^q \leq C^q. \tag{C.1}$$

When $q = 0$, we replace C^q with s, the level of 'hard' sparsity. We have the following result:

Theorem 3. *Let*

$$L(\widehat{\mathbf{u}}, \mathbf{u}) = \|\mathbf{u} - sign(\langle \mathbf{u}, \widehat{\mathbf{u}} \rangle)\widehat{\mathbf{u}}\|_2^2 \tag{C.2}$$

be the risk of the estimator $\widehat{\mathbf{u}}$ of $\mathbf{u}$, where $\mathbf{u}$ is as specified in (2.1) and the estimators are the six algorithms that we have previously described. Then,

a. sum-, HC-sum, and FDR-SEPCA have expected risks lower-bounded by

$$\mathbb{E}L(\widehat{\mathbf{u}}, \mathbf{u}) \geq O\left([C^q - 1]\|\mathbf{v}\|_1^{-(2-q)}\right). \tag{C.3}$$

b. ℓ_2-SEPCA has a risk lower-bounded by

$$\mathbb{E}L(\widehat{\mathbf{u}}, \mathbf{u}) \geq O\left([C^q - 1]n^{-\frac{1}{2}(1-q/2)}\right). \tag{C.4}$$

c. HC-ℓ_2-SEPCA has a risk lower-bounded by

$$\mathbb{E}L(\widehat{\mathbf{u}}, \mathbf{u}) \geq O\left([C^q - 1]n^{-(1-q/2)}\right). \tag{C.5}$$

The rest of this section contains the proof of Theorem 3.

C.1. Proof of Theorem 3

We construct a 'worst-case' sparse $\mathbf{u}$. Note that $C^q \geq 1$ necessarily, and that if $C \geq p^{1-q/2}$, every unit norm vector is in the ℓ_q ball. Hence, we take $C \in [1, p^{1-q/2})$.

Let θ and σ be fixed. We want a sparse vector with several coordinates guaranteed to be missed (the probability of not detecting them is asymptotically 1). For this vector $\mathbf{u}$ to be sparse and for the loss to not be 1, set u_1 to be $\sqrt{1 - r_n^2}$, where $r_n^2 = o(1)$, and take $u_2, \cdots, u_{m_n+1}$ to be $r_n/\sqrt{m_n}$. The other coordinates of u are 0, so that $\mathbf{u}$ has unit ℓ_2-norm.

We assume that u_1 is detected with probability 1 as $n \to \infty$, and want to set $u_2, \cdots, u_{m_n+1}$ so that the expected loss is lower bounded by:

$$\begin{aligned}
\mathbb{E}L(\mathbf{u}, \widehat{\mathbf{u}}) &\geq \sum_{k=1}^{p} |u_k|^2 \mathbb{P}(\text{Not Selecting Coordinate k}) \\
&\geq \sum_{k=2}^{m_n+1} |u_k|^2 \mathbb{P}(\text{Not Selecting Coordinate k}).
\end{aligned} \tag{C.6}$$

If coordinates of size $\frac{r_n}{\sqrt{m_n}}$ are not detected, the expected loss is lower bounded by r_n^2.

Let $m_n = \lfloor m \rfloor$ where

$$m = \delta n^\phi r_n^\psi \|v\|^\eta.$$

Note that we have not specified the norm used in $\|v\|$: we will choose the norm at the very end of the calculation. Let

$$r_n = [C^q - 1]^\alpha n^{\beta + \gamma q} \|v\|^\kappa,$$

so that,

$$\frac{r_n}{\sqrt{m_n}} \approx \frac{r_n}{\sqrt{m}} = \frac{1}{\delta} n^{-\phi/2} \|v\|^{\kappa - \eta/2} r_n^{1 - \psi/2}.$$

We will choose $\delta, \phi, \eta, \alpha, \beta, \gamma, \kappa, \psi$ so that the ℓ_q sparsity constraint is met and the lower bound r_n^2 is maximized. The sparsity constraint requires:

$$\sum_{i=1}^{p} |u_i|^q = (1 - r_n)^{q/2} + m_n^{1-q/2} r_n^q \leq 1 + m_n^{1-q/2} r_n^q \leq C^q. \qquad \text{(C.7)}$$

First, we will assume (for now) that $r_n = o(1)$ and that via other parameters we may control the scaling of the coordinate sizes; hence, we set $\psi = 2$. Then,

$$r_n^q m^{1-q/2} = \delta^{1-q/2} n^{2(\beta + \gamma q) + (1-q/2)\phi} [C^q - 1]^{2\alpha} \|v\|^{2\kappa + \eta(1-q/2)}. \qquad \text{(C.8)}$$

We need this quantity to be smaller than $C^q - 1$. To eliminate the n dependence, we set $\beta = \frac{-\phi}{2}$ and $\gamma = \frac{\phi}{4}$. We choose $\alpha = \frac{1}{2}$ to match powers of $[C^q - 1]$ on both sides of the inequality. Defining another parameter ρ, let $\delta = \rho \|\mathbf{v}\|^{-\eta}$. Then, the inequality is

$$\rho^{1-q/2} \|\mathbf{v}\|^{2\kappa} [C^q - 1] \leq [C^q - 1].$$

Choosing $\rho \leq \|\mathbf{v}\|^{-2\kappa/(1-q/2)}$ is enough.

With these choices of parameters,

$$r_n = \sqrt{[C^q - 1]} n^{-\frac{1}{2}\phi(1-q/2)} \|\mathbf{v}\|^\kappa,$$

and

$$m = \rho n^\phi r_n^2,$$

so that

$$\frac{r_n}{\sqrt{m}} = \frac{1}{\sqrt{\rho}} n^{-\phi/2}.$$

Noting that

$$\frac{1}{\sqrt{\rho}} \geq \|\mathbf{v}\|^{\kappa/(1-q/2)},$$

choosing $\rho = \|\mathbf{v}\|^{-2\kappa/(1-q/2)}$ leads to the smallest possible choice of coordinate.

In summary:

$$r_n = \sqrt{[C^q - 1]} n^{-\frac{1}{2}\phi(1-q/2)} \|\mathbf{v}\|^\kappa, \qquad \text{(C.9)}$$

$$r_n^2 = [C^q - 1] n^{-\phi(1-q/2)} \|\mathbf{v}\|^{2\kappa}, \qquad \text{(C.10)}$$

$$m = \|\mathbf{v}\|^{-2\kappa/(1-q/2)} n^\phi r_n^2, \qquad \text{(C.11)}$$

and

$$\frac{r_n}{\sqrt{m}} = \| \mathbf{v} \|^{\kappa/(1-q/2)} n^{-\phi/2}. \tag{C.12}$$

So, for a given algorithm, it remains to choose ϕ and κ so that the worst-case risk is lower-bounded by r_n^2. Sum-SEPCA misses coordinates of size $O\left(\frac{\sqrt{\log p}}{\| \mathbf{v} \|_1}\right)$ and ℓ_2-SEPCA misses coordinates of size $O\left(\frac{\sqrt{\log p}}{n^{1/4}\| \mathbf{v} \|_2}\right)$. For sum-SEPCA, $\kappa = \frac{q-2}{2}$, and for ℓ_2-SEPCA, κ is irrelevant, as $\| \mathbf{v} \|_2 = 1$. Sum-SEPCA uses $\phi = 0$ and ℓ_2-SEPCA uses $\phi = \frac{1}{2}$. Hence, sum-SEPCA has a risk lower-bounded by

$$O\left([C^q - 1]\| \mathbf{v} \|_1^{-(2-q)}\right). \tag{C.13}$$

Noting that $\| \mathbf{v} \|_2 = 1$, ℓ_2-SEPCA has

$$O\left([C^q - 1]n^{-\frac{1}{2}(1-q/2)}\| \mathbf{v} \|_2^{-(1-q/2)}\right) = $$
$$O\left([C^q - 1]n^{-\frac{1}{2}(1-q/2)}\right). \tag{C.14}$$

In the ℓ_0 case, i.e., when $\mathbf{u}$ has no more than s non-zero entries, the preceding analysis goes through with C^q replaced by s and q set to zero.

C.1.1. FDR algorithms

For HC-sum-SEPCA, the β_{crit} is of the same order as that for sum-SEPCA. Similarly, for FDR-SEPCA, if $\widehat{k}$ is much smaller than p, β_{crit} is of roughly the same order. Hence, these two algorithms have the same risk bound as sum-SEPCA. For HC-ℓ_2-SEPCA, $\kappa = 0$ and $\phi = 1$. The risk is therefore lower-bounded by

$$O\left([C^q - 1]n^{-(1-q/2)}\right). \tag{C.15}$$

Appendix D: FDR-SEPCA: Further details

Let $y_i = \mu_i + \sigma z_i$, where $i \in \{1, \cdots, p\}$, the vector $\mathbf{z}$ of the z_i is normally distributed with mean 0 and covariance Σ, and Σ satisfies

$$\xi_o \mathcal{I}_p \leq \Sigma \leq \xi_1 \mathcal{I}_p.$$

Here, ξ_0 is the smallest eigenvalue of Σ and ξ_1 is the largest. The mean vector $\boldsymbol{\mu}$ of the μ_i is assumed to be sparse; the goal is to estimate $\boldsymbol{\mu}$. The following penalized least squares formulation yields an estimator for $\boldsymbol{\mu}$:

$$\widehat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \| \mathbf{y} - \boldsymbol{\mu} \|_2^2 + \sigma^2 \text{pen}\left(\| \boldsymbol{\mu} \|_0\right), \tag{D.1}$$

where $\text{pen}(k)$ is defined as

$$\text{pen}(k) = \xi_1 \zeta k \left(1 + \sqrt{2L_{p,k}}\right)^2, \tag{D.2}$$

with $\zeta > 1$ and

$$L_{p,k} = (1 + 2\beta)\log(\nu p/k). \tag{D.3}$$

The parameter β may be set to 0 here, and ν is chosen to be no smaller than $e^{1/(1+2\beta)}$. We define $\|\boldsymbol{\mu}\|_0$ to be number of non-zero coordinates of $\boldsymbol{\mu}$.

The solution to (D.1) is given by hard-thresholding. Let $|y|_{(i)}$ be the ith order statistic of $|y_i|$, namely $|y|_{(1)} \geq \cdots \geq |y|_{(p)}$. Then if

$$\widehat{k} = \arg\min_{k \geq 0} \sum_{i > k} |y|_{(i)}^2 + \sigma^2 \text{pen}(k), \tag{D.4}$$

defining

$$t_k^2 = \text{pen}(k) - \text{pen}(k-1), \tag{D.5}$$

the solution is to hard threshold at $t_{\widehat{k}}$.

In this set-up, we have that

$$t_k \approx \lambda_{p,k} = \sqrt{\xi_1 \zeta}(1 + \sqrt{2L_{p,k}}),$$

with $|t_k - \lambda_{p,k}| < c/\lambda_k$. More precisely, Lemma 11.7 of [20] says that

$$\lambda_{p,k} - \frac{4\zeta b}{\lambda_{p,k}} \leq t_k \leq \lambda_{p,k}.$$

When $\nu \geq e^2$, we may take $b = (1+2\beta)$. In any case, if $k = o(n)$, $\lambda_{p,k} \asymp \sqrt{\log p}$. Hence, entries of $\mathbf{y}$ that are of size at least

$$y_i > (1 - o(1))\sqrt{\xi_1 \zeta}\sigma\left(1 + \sqrt{2\log(\nu p/\widehat{k})}\right)$$

are selected, or, since $\mu_i = (\theta u_i)\|\mathbf{v}\|_1$ (when $\mathbf{v}$ is equisigned), if we select $\widehat{k}$ coordinates, we expect to detect

$$\begin{aligned}
|\theta u_i| &> (1 - o(1))\sqrt{\xi_1 \zeta}\sigma\frac{\left(1 + \sqrt{2\log(\nu p/\widehat{k})}\right)}{\|\mathbf{v}\|_1} \\
&= O\left(\sigma\sqrt{\xi_1}\frac{\sqrt{2\log\left(\nu p/\widehat{k}\right)}}{\|\mathbf{v}\|_1}\right).
\end{aligned} \tag{D.6}$$

D.1. Risk behavior

Recalling that (D.1) solves a penalized least squares problem for $\widehat{\mathbf{y}}$ close to $\mathbf{y}$, we may discuss the statistical behavior of this estimator. The following discussion follows and reproduces that in [22]

First, note that for $\beta = 0$, the parameter ν directly controls the FDR (where a false positive corresponds to selecting a zero coordinate in $\mathbf{y}$): a choice of $\nu = 2^{1/\omega}$ for $\omega \in (0,1)$ bounds the FDR at a level ω.

Second, the expected risk, $\mathbb{E}\|\mathbf{y} - \widehat{\mathbf{y}}\|_2^2$, is bounded as follows. By Proposition 4.1 in [22],

$$\mathbb{E}\|\mathbf{y} - \widehat{\mathbf{y}}\|_2^2 \leq D\left[2M_p'\xi_1\sigma^2 + \mathcal{R}(\mathbf{y}, \sigma)\right], \tag{D.7}$$

where D is a constant $2\zeta(\zeta + 1)^3(\zeta - 1)^{-3} = \Theta(1)$, we assume that $\xi_1 = 1$, and $0 \leq M_p' \leq C_\beta p^{-2\beta}\nu^{-1}$, for some $C_\beta > 0$. Since $\beta = 0$, $M_p' = O(1/\nu) = O(\omega)$, if we control the FDR at level ω.

The second term in (D.7) is the ideal risk, or, the infimum of the penalized least squares objective. If $\mathbf{y}$ belongs to an ℓ_q ball with radius C and $0 < q < 2$, and we define

$$r_{p,q}(C) = \begin{cases} C^2 & \text{if } C \leq \sqrt{1 + \log p}, \\ C^q[1 + \log(p/C^q)]^{1-q/2} & \text{if } \sqrt{1 + \log p} \leq C \leq p^{1/q}, \\ p & \text{if } C \geq p^{1/q}, \end{cases} \tag{D.8}$$

the ideal risk is bounded as

$$\sup_{\mathbf{y} \in \mathbb{R}^p : \sum_i |y_i|^q \leq C^q} \mathcal{R}(\mathbf{y}, \sigma) \leq c(\log \nu)\sigma^2 r_{p,q}(C/\sigma), \tag{D.9}$$

for some $c > 0$. The supplementary results in [22] yield that $\mathcal{R}(\mathbf{y}, \sigma)$ is bounded by $C^2 \log \nu$, and by C^2 when $C \leq \sqrt{1 + \log p}$.

As in Appendix C, we may replace q with 0 and C^q with s in the case of hard sparsity with s non-zero coordinates. Doing so leads to the bound:

$$\mathbb{E}\|\mathbf{y} - \widehat{\mathbf{y}}\|_2^2 \leq s\sigma \log \nu \log \frac{\sigma p\nu}{s} + \sigma_1 \sigma^2 \frac{2}{\nu}. \tag{D.10}$$

Note that we have recovered the factor of $\log \nu p/s$ in β_{crit}.

Acknowledgments

References

[1] ALLEN-ZHU, Z. and LI, Y. (2016). LazySVD: even faster SVD decomposition yet without agonizing pain. In *Advances in Neural Information Processing Systems* 974–982.

[2] ARIAS-CASTRO, E., CANDÈS, E. J. and PLAN, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *Annals of Statistics* 2533–2556. MR2906877

[3] BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2012). The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis* **111** 120–135. MR2944410

[4] BERTHET, Q., RIGOLLET, P. et al. (2013). Optimal detection of sparse principal components in high dimension. *Annals of Statistics* **41** 1780–1815. MR3127849

[5] BIRNBAUM, A., JOHNSTONE, I. M., NADLER, B. and PAUL, D. (2013). Minimax bounds for sparse PCA with noisy high-dimensional data. *Annals of Statistics* **41** 1055. MR3113803

[6] BOBKOV, S. G. and NAZAROV, F. L. (2003). On convex bodies and log-concave probability measures with unconditional basis. In *Geometric aspects of functional analysis* 53–69. Springer. MR2083388

[7] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence.* Oxford University Press. MR3185193

[8] BOUCHERON, S. and THOMAS, M. (2012). Concentration inequalities for order statistics. *Electronic Communications in Probability* **17** no. 51, 1–12. MR2994876

[9] BRENNAN, M., BRESLER, G. and HULEIHEL, W. (2018). Reducibility and computational lower bounds for problems with planted sparse structure. *arXiv preprint* arXiv:1806.07508.

[10] BUTUCEA, C., NDAOUD, M., STEPANOVA, N. A., TSYBAKOV, A. B. et al. (2018). Variable selection with Hamming loss. *The Annals of Statistics* **46** 1837–1875. MR3845003

[11] D'ASPREMONT, A., EL GHAOUI, L., JORDAN, M. I. and LANCK-RIET, G. R. (2007). A direct formulation for sparse PCA using semidefinite programming. *SIAM Review* **49** 434–448. MR2353806

[12] DING, C. H., LI, T. and JORDAN, M. I. (2010). Convex and semi-nonnegative matrix factorizations. *IEEE Transactions on Pattern analysis and Machine Intelligence* **32** 45–55.

[13] DONOHO, D. and JIN, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Annals of Statistics* 962–994. MR2065195

[14] DONOHO, D., JIN, J. et al. (2015). Higher criticism for large-scale inference, especially for rare and weak effects. *Statistical Science* **30** 1–25. MR3317751

[15] GALL, F. L. and URRUTIA, F. (2018). Improved rectangular matrix multiplication using powers of the Coppersmith-Winograd tensor. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* 1029–1046. SIAM. MR3775855

[16] GAO, C., MOORE, B. E. and NADAKUDITI, R. R. (2017). Augmented robust PCA for foreground-background separation on noisy, moving camera video. In *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* 1240–1244. IEEE.

[17] HALL, P., JIN, J. et al. (2010). Innovated higher criticism for detecting sparse signals in correlated noise. *Annals of Statistics* **38** 1686–1732. MR2662357

[18] HOYER, P. O. (2004). Non-negative matrix factorization with sparse-

ness constraints. *Journal of Machine Learning Research* **5** 1457–1469. MR2248024

[19] Huang, K., Sidiropoulos, N. D. and Swami, A. (2014). Non-negative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition. *IEEE Transactions on Signal Processing* **62** 211–224. MR3187989

[20] Johnstone, I. M. (2017). Gaussian estimation: Sequence and wavelet models. *Unpublished manuscript.* http://statweb.stanford.edu/~imj/GE_08_09_17.pdf.

[21] Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association* **104** 682. MR2751448

[22] Johnstone, I. M. and Paul, D. (2014). Adaptation in some linear inverse problems. *Stat* **3** 187–199.

[23] Latała, R. (2011). Order statistics and concentration of norms for log-concave vectors. *Journal of Functional Analysis* **261** 681–696. MR2799576

[24] Liu, H., Wu, Z., Li, X., Cai, D. and Huang, T. S. (2012). Constrained nonnegative matrix factorization for image representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** 1299–1311.

[25] Ma, Z. et al. (2013). Sparse principal component analysis and iterative thresholding. *Annals of Statistics* **41** 772–801. MR3099121

[26] Montanari, A. and Richard, E. (2015). Non-negative principal component analysis: Message passing algorithms and sharp asymptotics. *IEEE Transactions on Information Theory* **62** 1458–1484. MR3472260

[27] Moore, B., Gao, C. and Nadakuditi, R. R. (2019). Panoramic robust PCA for foreground-background separation on noisy, free-motion camera video. *IEEE Transactions on Computational Imaging.*

[28] Ndaoud, M. (2018). Interplay of minimax estimation and minimax support recovery under sparsity. *arXiv preprint arXiv:1810.05478*. MR3932863

[29] Passemier, D., Li, Z. and Yao, J. (2017). On estimation of the noise variance in high dimensional probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **79** 51–67. MR3597964

[30] Pastor, D. and Socheleau, F.-X. (2012). Robust estimation of noise standard deviation in presence of signals with unknown distributions and occurrences. *IEEE transactions on Signal Processing* **60** 1545–1555. MR2934179

[31] Perry, A., Wein, A. S., Bandeira, A. S., Moitra, A. et al. (2018). Optimality and sub-optimality of PCA I: Spiked random matrix models. *The Annals of Statistics* **46** 2416–2451. MR3845022

[32] Piccardi, M. (2004). Background subtraction techniques: A review. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)* **4** 3099–3104. IEEE.

[33] Ravikumar, P., Wainwright, M. J. and Lafferty, J. D. (2010). High-dimensional Ising model selection using ℓ_1-regularized logistic regression. *Annals of Statistics* **38** 1287–1319. MR2662343

[34] REEVES, G. and GASTPAR, M. (2008). Sampling bounds for sparse support recovery in the presence of noise. In *2008 IEEE International Symposium on Information Theory* 2187–2191. IEEE.

[35] REN, B., PUEYO, L., ZHU, G. B., DEBES, J. and DUCHÊNE, G. (2018). Non-negative matrix factorization: robust extraction of extended structures. *The Astrophysical Journal* **852** 104.

[36] ROSS, P. Stars. https://archive.org/details/Stars_2D. Accessed: 2016 November 16.

[37] SOCHELEAU, F.-X. and PASTOR, D. (2014). Testing the energy of random signals in a known subspace: An optimal invariant approach. *IEEE Signal Processing Letters* **21** 1182–1186.

[38] TASLAMAN, L. and NILSSON, B. (2012). A framework for regularized non-negative matrix factorization, with application to the analysis of gene expression data. *PloS one* **7** e46331.

[39] VASWANI, N., BOUWMANS, T., JAVED, S. and NARAYANAMURTHY, P. (2018). Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery. *IEEE signal processing magazine* **35** 32–55. MR3923185

[40] WANG, Y.-X. and ZHANG, Y.-J. (2013). Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering* **25** 1336–1353.

[41] YUAN, X.-T. and ZHANG, T. (2013). Truncated power method for sparse eigenvalue problems. *Journal of Machine Learning Research* **14** 899–925. MR3063614

[42] ZASS, R. and SHASHUA, A. (2007). Nonnegative sparse PCA. In *Advances in neural information processing systems* 1561–1568.