The Annals of Statistics
2015, Vol. 43, No. 3, 1300–1322
DOI: 10.1214/15-AOS1310

DO SEMIDEFINITE RELAXATIONS SOLVE SPARSE PCA UP TO THE INFORMATION LIMIT?

BY ROBERT KRAUTHGAMER[1,*], BOAZ NADLER[2,*] AND DAN VILENCHIK[3,†]

Weizmann Institute of Science and Ben-Gurion University†*

Estimating the leading principal components of data, assuming they are sparse, is a central task in modern high-dimensional statistics. Many algorithms were developed for this sparse PCA problem, from simple diagonal thresholding to sophisticated semidefinite programming (SDP) methods. A key theoretical question is under what conditions can such algorithms recover the sparse principal components? We study this question for a single-spike model with an ℓ_0-sparse eigenvector, in the asymptotic regime as dimension p and sample size n both tend to infinity. Amini and Wainwright [*Ann. Statist.* **37** (2009) 2877–2921] proved that for sparsity levels $k \geq \Omega(n/\log p)$, no algorithm, efficient or not, can reliably recover the sparse eigenvector. In contrast, for $k \leq O(\sqrt{n/\log p})$, diagonal thresholding is consistent. It was further conjectured that an SDP approach may close this gap between computational and information limits. We prove that when $k \geq \Omega(\sqrt{n})$, the proposed SDP approach, at least in its standard usage, cannot recover the sparse spike. In fact, we conjecture that in the single-spike model, no computationally-efficient algorithm can recover a spike of ℓ_0-sparsity $k \geq \Omega(\sqrt{n})$. Finally, we present empirical results suggesting that up to sparsity levels $k = O(\sqrt{n})$, recovery is possible by a simple covariance thresholding algorithm.

1. Introduction. Principal components analysis (PCA) is a popular technique for dimension reduction that has a wide range of applications involving multivariate data, in both science and engineering; see, for example, [4, 22]. The first principal component (PC) of a p-dimensional random variable $\mathbf{x} = (x_1, \dots, x_p)$ is the direction in which the variance of $\mathbf{x}$ is maximal, or equivalently, the leading eigenvector of its population covariance matrix $\Sigma = \mathbb{E}[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T]$ where $\mu = \mathbb{E}[\mathbf{x}]$. In practice, one typically does not have explicit access to Σ, but rather is given n samples from $\mathbf{x}$, from which one computes the sample covariance matrix $\hat{\Sigma}$ and its leading eigenvectors.

Received September 2014; revised January 2015.

[1]Supported in part by a US–Israel BSF Grant #2010418, ISF Grant #897/13 and by the Citi Foundation.

[2]Supported in part by the Citi Foundation.

[3]Research done in part while a postdoctoral fellow at the Weizmann Institute.

MSC2010 subject classifications. Primary 62H25; secondary 62F12.

Key words and phrases. Principal component analysis, spectral analysis, spiked covariance ensembles, sparsity, high-dimensional statistics, convex relaxation, semidefinite programming, Wishart ensembles, random matrices, integrality gap.

In contemporary applications where variables are plentiful (large p) but samples are relatively scarce (small n), PCA suffers from two major limitations: (1) the principal components are typically a linear combination of all variables, which hinders their interpretation and subsequent use, and (2) while PCA is consistent in the classical setting (p is fixed and $n \to \infty$) [4, 31], it is generally inconsistent in high-dimensions. Indeed, as shown, for example, in [8, 13, 20, 21, 32], when p is comparable to, or significantly larger than n, the sample covariance matrix $\hat{\Sigma}$ may be a poor approximation to the population's covariance matrix Σ, and its leading eigenvectors may be far from the population's principal components.

To address the first drawback, one can consider a *sparse* PCA problem, in which for some appropriate parameter k, we search for a direction with at most k nonzero coefficients and with maximal variance. Formally, the ℓ_0-sparse PCA problem is defined by

$$\mathcal{L}_0(\Sigma) = \operatorname{argmax}\{\mathbf{w}^T \Sigma \mathbf{w} : \|\mathbf{w}\|_2 = 1, \|\mathbf{w}\|_0 \le k\}. \tag{1.1}$$

We note that other notions of sparsity were considered in the literature, for example, a population covariance matrix that has only a few large eigenvalues, whose corresponding eigenvectors are sparse in ℓ_q-norm for $q \in (0, 2)$ [9, 10, 21, 28].

While standard (nonrestricted) PCA can be efficiently solved by computing the eigenvectors of a symmetric matrix, sparse PCA is a difficult combinatorial problem, and in fact solving $\mathcal{L}_0(\Sigma)$ is NP-hard.[4] Nevertheless, various computationally efficient approaches were developed to deal with the problem. These include greedy or nonconvex optimization procedures [23, 40], methods based on ℓ_1-regularization [12, 30, 39, 42], regularized singular-value-decomposition [34], an augmented Lagrangian method [27], a simple diagonal thresholding (DT) algorithm [21], and sophisticated semidefinite programming (SDP) methods [11]. The latter approach, and in particular its ability to recover an ℓ_0-sparse PC, are the focus of the current paper.

SDP-based algorithm. We study the following concrete SDP relaxation of (1.1), which was suggested by d'Aspremont et al. [12]:

$$\operatorname{argmax}\{\langle \hat{\Sigma}, X\rangle : X \in \mathcal{S}_+^p, \operatorname{tr}(X) = 1, \|X\|_S \le k\}, \tag{1.2}$$

where for two matrices $X, Y \in \mathbb{R}^{p\times p}$ we denote by $\langle X, Y\rangle = \sum_{i,j} X_{ij}Y_{ij} = \operatorname{tr}(X^T Y)$ their Frobenius inner-product, $\|X\|_S = \sum_{i,j} |X_{ij}|$ is the "absolute-sum norm," and $\mathcal{S}_+^p = \{X \in \mathbb{R}^{p\times p} : X = X^T, X \succeq 0\}$ is the cone of symmetric positive semidefinite (PSD) matrices. As SDP (1.2) returns a symmetric matrix rather than a vector, d'Aspremont et al. [12] suggested to output its leading eigenvector as an estimate for the first sparse-PC. This algorithm is summarized as follows.

Single-spike input model. We examine Algorithm 1 under the single-spike multivariate Gaussian model introduced in [20], where the samples $\mathbf{x}_i$ are of the form

$$\mathbf{x}_i = \sqrt{\beta} u_i \mathbf{z} + \boldsymbol{\xi}_i, \qquad i = 1, \dots, n. \tag{1.3}$$

[4]This claim follows from [29] and [33], but can also proved by a direct reduction from the k-clique problem in a p-vertex graph, and considering $\Sigma = A + pI$ where A is the graph's adjacency matrix.

Algorithm 1: SDP-estimator

input : (mean-centered) vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$, sparsity parameter k
output: vector $\hat{\mathbf{z}} \in \mathbb{R}^p$

let $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T$
compute a solution $X \in \mathbb{R}^{p \times p}$ of SDP (1.2)
let $\hat{\mathbf{z}}$ be the leading (unit-length) eigenvector of X

Here, the parameter $\beta > 0$ is the *signal strength*, $\mathbf{z} \in \mathbb{R}^p$ is the *planted spike* assumed to be a k-sparse unit-length vector, $\boldsymbol{\xi}_i \in \mathbb{R}^p$ is a noise vector whose entries are all i.i.d. $N(0, 1)$ and $u_i \sim N(0, 1)$. Furthermore, all the u_i's and $\boldsymbol{\xi}_i$'s are independent of each other. The corresponding population covariance matrix is

$$\Sigma = \beta \mathbf{z}\mathbf{z}^T + \mathbf{I}_p, \tag{1.4}$$

and its largest eigenvalue is $1 + \beta$, with associated eigenvector $\mathbf{z}$. We consider throughout the scenario $(n, p, k) \to \infty$, and mention additional assumptions (e.g., β is fixed or p/n tends to $c > 0$) as needed.

Information versus computational limits. Amini and Wainwright [3] studied this single-spike input model, under the additional assumption that the nonzero entries of $\mathbf{z}$ are exactly of the form $\pm 1/\sqrt{k}$, which represents the hardest type of k-sparse vectors. They proved that up to sparsity level $k = O(\kappa_{n,p})$ where $\kappa_{n,p} = \sqrt{n/\log p}$, Algorithm 1 outputs a vector $\hat{\mathbf{z}}$ whose support coincides with that of $\mathbf{z}$;[5] they further showed, using a simple second moment calculation, that up to the same order of sparsity level $k = O(\kappa_{n,p})$, the diagonal thresholding algorithm [21] also recovers the support of $\mathbf{z}$ and fails whenever $k/\kappa_{n,p} \to \infty$. In contrast, Amini and Wainwright [3] showed that for $k = \Omega(\kappa_{n,p}^2)$,[6] *every* method [including exhaustive search over all $\binom{p}{k}$ subsets of size k] will err with probability at least $1/2$. In fact, even the simpler task of *detecting* the presence of a spike is not possible for this range of parameters, as recently proved in [6, 7]. For further results including minimax rates, under more general sparsity models, see [10, 37, 38].

The following question thus remained open: Does Algorithm 1, which is more sophisticated and computationally heavy, outperform the simple DT algorithm? Specifically, are there intermediate sparsity levels $\kappa_{n,p} < k < \kappa_{n,p}^2$ (such that $k/\kappa_{n,p} \to \infty$, and ignoring multiplicative constants) for which $\hat{\mathbf{z}}$ still approximates $\mathbf{z}$ in some useful sense? While not answering this question, Amini and Wainwright

[5]For technical reasons, their proof requires the additional condition $k = O(\log p)$, which they conjecture can be removed.

[6]We write $f = \Omega(g)$ if $f(n) \geq Cg(n)$ for some absolute positive constant C and all sufficiently large n. Similarly, $f = \Theta(g)$ means $C_1 g(n) \leq f(n) \leq C_2 g(n)$.

proved that for sparsity level up to $k = O(\kappa_{n,p}^2)$, *if* the solution to (1.2) remains rank one, *then* the support of $\hat{\mathbf{z}}$ coincides with that of $\mathbf{z}$. They then suggested that for this ℓ_0-sparse PCA problem, the information and computational limits coincide, both are equal $\Theta(\kappa_{n,p}^2)$, and Algorithm 1 is optimal. In their words, "under the rank-one condition, the SDP is in fact statistically optimal, that is, it requires only the necessary number of samples (up to a constant factor) to succeed" [3], page 2880.

Our results, formally stated below, prove that unfortunately this is *not* the case—in fact, when k slightly exceeds $\kappa_{n,p}$, namely $k = \Omega(\kappa_{n,p}\sqrt{\log p}) = \Omega(\sqrt{n})$, the solution X of SDP (1.2) *does not* have rank one and is not close to $\mathbf{z}\mathbf{z}^T$. Furthermore, if X has a low rank, then the output $\hat{\mathbf{z}}$ of Algorithm 1 is at best weakly correlated with $\mathbf{z}$. In Section 3 we present empirical simulation results showing that indeed Algorithm 1 and DT perform similarly.

Given that the SDP algorithm does not seem to significantly improve over DT under the single spike model, the following question arises: Is there a simple algorithm which outperforms both? Motivated by the work of Bickel and Levina [8], we suggest a light-weight greedy algorithm called Covariance Thresholding (CT), which can be seen as a generalization of Diagonal Thresholding. We provide experimental results suggesting that CT is consistent for $k = O(\sqrt{n})$; see Section 3 for details. Recently, following our work, Deshpande and Montanari [16] rigorously proved that a variant of our CT algorithm indeed asymptotically recovers the support of $\mathbf{z}$ up to these sparsity levels. Finally, we note that despite our results, there are other settings, such as estimating sparse eigenvectors of correlation matrices, where SDP-based methods are provably better than diagonal thresholding, possibly even achieving the relevant minimax rates [26, 41].

1.1. *Our results.* We consider the single-spike model defined in (1.3) in high-dimensional settings whereby $(n, p, k) \to \infty$ and $p/n^{\alpha} \to c$ for positive constants $c, \alpha \geq 1$. We further assume that the k-sparse vector $\mathbf{z}$ has k nonzero entries of the form $\pm 1/\sqrt{k}$. In what follows, we denote by $\operatorname{supp}(\mathbf{x})$ the set $\{i : \mathbf{x}_i \neq 0\}$. In the analysis, we assume without loss of generality that the nonzero coordinates of the spike $\mathbf{z}$ are exactly its first k coordinates, that is, $\operatorname{supp}(\mathbf{z}) = \{1, 2, \dots, k\}$.

For the case $\alpha = 1$, that is, $p/n \to c$, we focus on weak signal strengths $\beta \leq \sqrt{\frac{p}{n}}$, whereas when $\alpha > 1$, the signal strength may grow to infinity provided it still satisfies $\beta \leq \sqrt{\frac{p}{n}}$; see assumption (b) below. The reason is that when $\alpha = 1$ and $\beta > \sqrt{\frac{p}{n}}$, as the next theorem shows, recovering the support of $\mathbf{z}$ is computationally easy, almost up to the information limit. As before, we let $\kappa_{n,p} = \sqrt{n/\log p}$.

THEOREM 1.1 (Strong signal). *Fix $c > 1$ and $\beta > \sqrt{c}$, and let $(n, p, k) \to \infty$ such that $p/n \to c$ and $k/\kappa_{n,p}^2 \to 0$. Let $\hat{\mathbf{w}}_1$ be the leading eigenvector of $\hat{\Sigma}$, and denote by $\operatorname{supp}_k(\hat{\mathbf{w}}_1)$ its k largest entries in absolute value. Then $\operatorname{supp}_k(\hat{\mathbf{w}}_1) = \operatorname{supp}(\mathbf{z})$ with probability tending to one as $(n, p, k) \to \infty$.*

Our next results, stated in the three theorems below, refer to the following assumptions:

(a) Fix positive $c, \alpha \geq 1$, and let $(n, p, k) \to \infty$ such that $p/n^\alpha \to c$.
(b) The signal strength, either fixed or growing with n, p, satisfies $\beta \leq \sqrt{\frac{p}{n}}$.
(c) The sparsity level k satisfies $k \geq 2p/\sqrt{n}$, and $k/p \to 0$.

We next analyze the quality of the output $\hat{\mathbf{z}}$ of Algorithm 1, as measured by its cosine-similarity to the planted spike $\mathbf{z}$.

THEOREM 1.2 (Cosine similarity). *Assume* (a)–(c). *Then there exists* $\varepsilon = \varepsilon(n) \to 0$, *such that if* X *is a solution of SDP* (1.2), *and* λ_1 *is its largest eigenvalue, then with probability tending to one as* $(n, p, k) \to \infty$, *the output* $\hat{\mathbf{z}}$ *of Algorithm* 1 *satisfies*

$$|\langle \hat{\mathbf{z}}, \mathbf{z}\rangle|^2 \leq \frac{23}{\lambda_1}\sqrt{\frac{n}{p}}(1 + \sqrt{\beta}) + \frac{\varepsilon}{\lambda_1}. \tag{1.5}$$

The following corollary of Theorem 1.2 shows that the SDP solution is far from $\mathbf{z}\mathbf{z}^T$. For a matrix A we denote its spectral norm by $\|A\| = \sqrt{\lambda_{\max}(AA^T)}$.

COROLLARY 1.3. *Assume* (a)–(c), *and further that* $p \geq 150^4 n$. *Let* X *be a solution of SDP* (1.2). *Then* $\|X - \mathbf{z}\mathbf{z}^T\| \geq \frac{1}{3}$ *with probability tending to one as* $(n, p, k) \to \infty$.

PROOF. Assume for contradiction that the matrix $Y = X - \mathbf{z}\mathbf{z}^T$ has a small spectral norm $\eta_1 = \|Y\| < 1/3$. Using Weyl's inequality [35], $\|X\| \geq \|\mathbf{z}\mathbf{z}^T\| - \|Y\|$. Since $\|\mathbf{z}\|_2 = 1$, the largest eigenvalue of X is thus lower bounded by $\lambda_1 \geq 1 - 1/3 = 2/3$. Let $\hat{\mathbf{z}}$ be a (unit-length) eigenvector of X corresponding to this largest eigenvalue λ_1. Recalling the variational definition of the largest eigenvector of a matrix, we obtain

$$\tfrac{2}{3} \leq \lambda_1 = \hat{\mathbf{z}}^T X \hat{\mathbf{z}} = \hat{\mathbf{z}}^T (Y + \mathbf{z}\mathbf{z}^T)\hat{\mathbf{z}} = \hat{\mathbf{z}}^T Y \hat{\mathbf{z}} + \hat{\mathbf{z}}^T \mathbf{z}\mathbf{z}^T \hat{\mathbf{z}}. \tag{1.6}$$

Using our assumption, $\hat{\mathbf{z}}^T Y \hat{\mathbf{z}} \leq \|Y\| = \eta_1 \leq 1/3$. By Theorem 1.2

$$\hat{\mathbf{z}}^T \mathbf{z}\mathbf{z}^T \hat{\mathbf{z}} = |\langle \hat{\mathbf{z}}, \mathbf{z}\rangle|^2 \leq \frac{23}{\lambda_1}\sqrt{\frac{n}{p}}(1 + \sqrt{\beta}) + \frac{\varepsilon}{\lambda_1}. \tag{1.7}$$

Plugging $p/n = 150^4$, $\beta \leq \sqrt{p/n} = 150^2$ and $\lambda_1 \geq 2/3$ into equation (1.7) gives that its right-hand side is at most $0.2315 + \frac{3}{2}\varepsilon$. Since by Theorem 1.2, $\varepsilon = \varepsilon(n) \to 0$ as $n \to \infty$, (1.7) is strictly smaller than 1/3 for a sufficiently large n. Combining (1.6) and (1.7) we arrive at the following contradictory set of inequalities:

$$\tfrac{2}{3} \leq \lambda_1 = \hat{\mathbf{z}}^T Y \hat{\mathbf{z}} + \hat{\mathbf{z}}^T \mathbf{z}^T \mathbf{z}\hat{\mathbf{z}} < \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{2}{3}. \qquad \square$$

Note that the constant 23 appearing in equation (1.5), and consequently the factor 150^4 in the corollary, are not necessarily optimal. Both may be further reduced at the expense of more involved proofs.

Further note that if $\lambda_1(X)$ is bounded away from zero as $p, n \to \infty$, then for $\alpha > 1$, equation (1.5) implies that $\langle \hat{\mathbf{z}}, \mathbf{z} \rangle \to 0$. Namely, in this case the output of Algorithm 1 is nearly orthogonal to $\mathbf{z}$. Such an empirical behavior of λ_1 was observed in our experimental results; see Figure 4.

We prove Theorem 1.2 using the next result, which itself may be of interest as it bounds the value of SDP (1.2). Recall that the SDP solution is highly nonlinear in its inputs, and therefore no closed-form explicit expression is known for the solution X or the SDP value $\langle \hat{\Sigma}, X \rangle$.

THEOREM 1.4 (SDP value). *Assume* (a)–(c). *Then there exists $\zeta = \zeta(n) \to 0$ such that with probability tending to one as $(n, p, k) \to \infty$, every solution X of SDP* (1.2) *satisfies*

$$(1.8) \qquad (1 - \zeta)\left(1 + \frac{p}{n}\right) \le \langle \hat{\Sigma}, X \rangle \le (1 + \zeta)\left(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta}\right)^2.$$

For $\alpha > 1$, the ratio between the upper and lower bounds in (1.8) is at most $1 + O(\zeta + \sqrt{\frac{n}{p}}(1 + \sqrt{\beta}))$ and tends to one as $p, n \to \infty$.

For the important regime $\alpha = 1$, we can use Theorem 1.4 to sharpen our conclusion from Theorem 1.2 and show that with probability tending to one, not only $X \neq \mathbf{z}\mathbf{z}^T$, but X is not even rank one. We arrive at this conclusion by combining Theorem 1.4 with the next theorem.

THEOREM 1.5. *Assume* (a)–(c), *and in addition $\alpha = 1$, $c > 20$ and $k/(p/\log^2 p) \to 0$. Then with probability tending to one as $(n, p, k) \to \infty$, every rank-one matrix $Y = \mathbf{y}\mathbf{y}^T$ that is feasible for SDP* (1.2) *satisfies*

$$(1.9) \qquad \langle \hat{\Sigma}, Y \rangle \le \frac{8}{9} \cdot \frac{p}{n}.$$

To see that the solution X of SDP (1.2) is indeed not rank one, we compare the upper bound in (1.9) with the (larger) lower bound in (1.8), namely, $\langle \hat{\Sigma}, Y \rangle \le \frac{8}{9} \cdot \frac{p}{n} < (1 - \zeta)(1 + \frac{p}{n}) \le \langle \hat{\Sigma}, X \rangle$.[7]

In conclusion, Theorems 1.2–1.5 suggest that the standard SDP-based approach (provided by Algorithm 1) is not significantly more effective than the simpler, light-weight diagonal thresholding. In particular, for weak signal strengths, Algorithm 1 does not yield a rank-one solution and hence cannot provably solve sparse

[7] We remark that another lower bound $\langle \hat{\Sigma}, X \rangle \ge 1 + \beta$ was proved in [6], Proposition 6.1, in a setting similar to Theorem 1.4, but we cannot use it to derive $\langle \hat{\Sigma}, Y \rangle < \langle \hat{\Sigma}, X \rangle$ because $\frac{8}{9}\frac{p}{n}$ could be larger than $1 + \beta$.

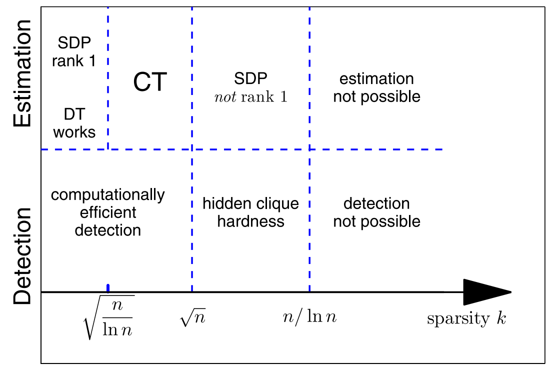

FIG. 1. *State of the art for detection and estimation in the single-spike model under various regimes of l_0-sparsity* (*assuming $n \approx p$ and omitting constant factors*).

PCA up to the information limit, as previously suspected. Our conclusion is in line with a recent, independently obtained result of Berthet and Rigollet [7], which asserts that the existence of a polynomial-time computable statistic for reliably *detecting* the presence of a single spike of ℓ_0-sparsity k for $k/\sqrt{n} \to \infty$, implies a polynomial-time algorithm for reliably detecting the presence of a planted clique of size k', for $k'/\sqrt{n} \to 0$, in an otherwise random graph $G(n, 1/2)$. The latter problem, known as the *hidden clique problem* in the computer science literature, is believed to be a computationally hard task, and polynomial-time algorithms known to date can only find a planted clique whose size k' is at least of order $\sqrt{n}$ [1, 2, 14, 15, 18, 19]. Furthermore, Wang et al. [38] showed that under the hidden clique hardness assumption, in certain sparsity regimes no randomized polynomial time algorithms can estimate the leading spiked eigenvector with optimal rate.

Our result differs from [7] in several respects. First, our results are unconditional; that is, Theorems 1.2–1.5 are not based on any computational hardness assumptions, and thus remain valid even if future developments will yield a polynomial-time algorithm for finding a hidden clique of size $n^{0.49}$. Second, our focus is on estimation and not on detection, which in general are different problems.

We summarize in Figure 1 the picture emerging from the results of Amini and Wainwright [3], Berthet and Rigollet [7], Deshpande and Montanari [16] and our work. Based on these results and the fact that even a sophisticated SDP-based algorithm fails to estimate $\mathbf{z}$ for $k \geq \sqrt{n}$, we conclude with the following conjecture.

CONJECTURE 1.6. *In the single-spike model with $p/n \to c$, fixed signal strength $\beta \leq \sqrt{p/n}$ and ℓ_0-sparsity $k = n^{0.5+\varepsilon}$ for fixed $\varepsilon > 0$, no polynomial-time algorithm can recover the support of $\mathbf{z}$ with probability tending to one as* $(n, p, k) \to \infty$.

Algorithm 2: Covariance thresholding

input : vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$, threshold t, sparsity level k
output: subset $S \subseteq [p]$ of cardinality k

1 compute $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T$

2 compute $T \in \mathbb{R}^{p \times p}$ by thresholding the entries of $\hat{\Sigma}$, namely,

$$T_{ij} = \begin{cases} \hat{\Sigma}_{ij}, & \text{if } |\hat{\Sigma}_{ij}| > t; \\ 0, & \text{otherwise} \end{cases}$$

3 let $\mathbf{w} \in \mathbb{R}^p$ be the leading eigenvector of T

4 let $S \subseteq [p]$ contain the k coordinates of largest absolute value in $\mathbf{w}$

Organization. In Section 2 we describe our covariance thresholding algorithm, followed by experimental results in Section 3. In Section 4 we give a short proof of Theorem 1.1. In Section 5 we assert preliminary facts that will be later used in the proofs of Theorem 1.2 in Section 6, Theorem 1.4 in Section 7 and Theorem 1.5 in Section 8.

2. Covariance thresholding algorithm. Motivated by the work of Bickel and Levina [8], we suggest Algorithm 2 for the ℓ_0-sparse PCA problem, which we call *covariance thresholding*, or CT for short.

We present some intuition as to why we expect this algorithm to work. From the definition of $\hat{\Sigma}$ in (1.3), it follows easily that the off-diagonal noise entries have expected value zero and standard deviation $1/\sqrt{n}$, while for signal entries the expected value is $\pm\beta/k$ with s.d. $C(\beta)/\sqrt{n}$. Consider, for example, a signal strength $\beta = 1$, sparsity $k \le \sqrt{n}/10$ (where 10 is rather arbitrary), and choose $t = 5/\sqrt{n}$. Then for a noise entry to survive thresholding, it must deviate from its mean by 5 s.d. and an analogous deviation for a signal entry to be zeroed out. Both events happen with small constant probability; hence most noise entries are zeroed and a constant fraction of signal entries survive. In fact, when $k = O(\sqrt{n/\log p})$ one can easily show that CT, similar to DT, recovers the support of $\mathbf{z}$. Recently, Deshpande and Montanari [16] proved that a variant of our algorithm is consistent up to sparsity levels $k = O(\sqrt{n})$. Their proof method is not directly applicable to our algorithm, but simulation results, detailed below, suggest that our algorithm is also able to recover the correct support up to $k = O(\sqrt{n})$. Hence, covariance thresholding is thus far the only algorithm, with polynomial run-time, that can *provably* recover the support up to sparsity levels $k = O(\sqrt{n})$.

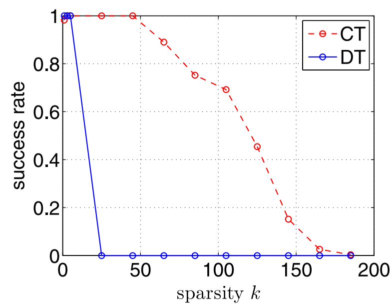

FIG. 2. *Performance of DT vs. CT, $n = p = 5000$. y-axis is the success rate averaged over* 500 *runs, with signal strength $\beta = 2$, and CT parameterized with threshold $t = 3/(2k)$.*

3. Simulation results.

3.1. *Covariance thresholding versus diagonal thresholding.* We compare a few algorithms under the following setup. We generate n i.i.d. samples $\mathbf{x}_i$ from the single-spike model (1.3) with a spike $\mathbf{z}$ of the form $\mathbf{z} = (\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{k}}, \dots, \frac{1}{\sqrt{k}}, 0, 0, \dots, 0)$. We assume the sparsity level k is a priori known, and say that an execution of an algorithm is *successful* if it returns the support of $\mathbf{z}$ exactly, that is, if the output is the set $\{1, \dots, k\}$. The *success rate* of an algorithm in M independent executions is the number of times it is successful divided by M. In each experiment we fix $n = p$ and for various values of k we measure the success rate averaged over $M = 500$ independent executions. Figure 2 compares the performance of our CT algorithm to DT. It is evident from this figure that in our setting, CT outperforms DT. Figure 3 shows the success rate of CT as a function of the sparsity level k scaled by $\sqrt{n}$, plotted for five different values of n. These results reinforce our prediction that CT works up to sparsity levels proportional to $\sqrt{n}$ (perhaps even slightly more).

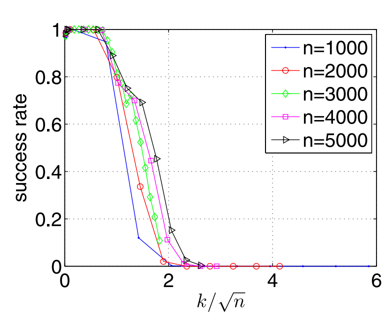

FIG. 3. *Performance of CT in proportion to $k/\sqrt{n}$ (depicted for different k and n). y-axis is the success rate averaged over* 500 *runs, with signal strength $\beta = 2$, and CT parameterized with threshold $t = 3/(2k)$.*

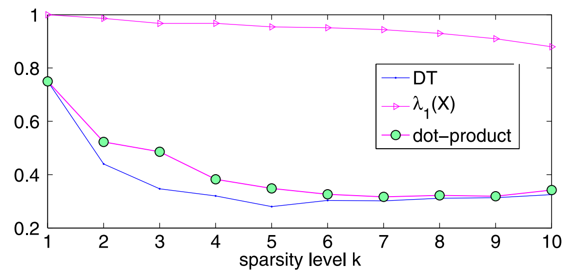

FIG. 4. *Comparison of DT and SDP* (*Algorithm* 1) *for* $n = p = 50$, $\beta = 0.8$, *averaged over* 100 *runs. The blue dots represent the success rate of DT. The filled circles are the average of the dot product* $|\langle \hat{\mathbf{z}}, \mathbf{z}\rangle|$ *of the SDP leading eigenvector, whereas the triangles are the largest eigenvalue of the SDP solution*, $\lambda_1(X)$.

3.2. *SDP* (*Algorithm* 1) *versus diagonal thresholding*. We run Algorithm 1 with parameters $n = p = 50$ and $\beta = 0.8$, averaging over $M = 100$ runs. We solve the SDP in line 2 of Algorithm 1 using SeDuMi 1.2.1 [36]. Figure 4 plots the dot-product (in absolute value) between $\hat{\mathbf{z}}$, the output of Algorithm 1 and the planted spike $\mathbf{z}$. As expected, the dot-product gets smaller as the sparsity k increases. For comparison, the figure plots also the recovery rate of DT, which also deteriorates as k increases. The figure also shows the largest eigenvalue of the SDP solution X; we remark that this value is rather close to one, even when the output of Algorithm 1 is far from $\mathbf{z}$, and is certainly bounded away from 0, as assumed in the discussion following Theorem 1.2.

4. Proof of Theorem 1.1 (Strong signal). Let $\hat{\mathbf{w}}_1$ be the leading eigenvector of $\hat{\Sigma}$, and write it as a linear combination of the spike $\mathbf{z}$ and some unit vector $\mathbf{a} \perp \mathbf{z}$, namely, $\hat{\mathbf{w}}_1 = g\mathbf{z} + \sqrt{1-g^2}\mathbf{a}$. We may assume $g \in [0, 1]$ by negating $\hat{\mathbf{w}}_1$, if necessary. According to [13], Theorem 4, for our setting of $\beta > \sqrt{c}$,

$$g = g(\beta) \xrightarrow{\text{a.s.}} \sqrt{(\beta^2 - c)/(\beta^2 + \beta c)} \qquad \text{as } n \to \infty. \tag{4.1}$$

Furthermore, according to Debashis ([13], Theorem 6), the vector $\mathbf{a} \in \mathbb{R}^p$ is distributed uniformly on the unit sphere of dimension $p - 1$ of vectors in $\mathbb{R}^p$ orthogonal to $\mathbf{z}$. Using this fact we prove below the following property of the entries of $\mathbf{a}$.

LEMMA 4.1. *With probability tending to one*, *all entries of* $\mathbf{a}$ *are bounded in absolute value by* $h\sqrt{\frac{\log p}{p}}$ *for a suitable constant* $h > 0$.

Lemma 4.1 implies that with probability tending to one, for all $i \in [1, k]$ we have $|(\hat{\mathbf{w}}_1)_i| \geq \frac{g}{\sqrt{k}} - \sqrt{1-g^2} \cdot h\sqrt{\frac{\log p}{p}}$, and for all $i \in [k+1, p]$ we have $|(\hat{\mathbf{w}}_1)_i| \leq$

$\sqrt{1-g^2} \cdot h\sqrt{\frac{\log p}{p}}$. To correctly identify the support of $\mathbf{z}$, it suffices to require a gap between signal and nonsignal coordinates, namely,

$$\frac{g}{\sqrt{k}} > 2h\sqrt{1-g^2}\sqrt{\frac{\log p}{p}}.$$

Solving for k and using (4.1), this inequality holds whenever $k < h' p/\log p$ for suitable $h' = h'(\beta) > 0$, which in turn holds with probability tending to one, because our assumption $k/\kappa_{n,p}^2 = k/(n/\log p) \to 0$ implies $k/(p/\log p) \to 0$. This completes the proof of Theorem 1.1.

PROOF OF LEMMA 4.1. Let $\{\mathbf{s}_1, \dots, \mathbf{s}_{p-1}\}$ be an orthonormal basis for the subspace of vectors in $\mathbb{R}^p$ orthogonal to $\mathbf{z}$. Since $\mathbf{a} = (a_1, \dots, a_p)$ is uniformly distributed in this subspace, it can be represented as $\mathbf{a} = \frac{1}{\|\boldsymbol{\xi}\|}\sum_{i=1}^{p-1} \xi_i \mathbf{s}_i$, where $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_{p-1})$ is a vector of i.i.d. standard Gaussians.

Fix a coordinate $i \in \{1, \dots, p\}$, and write its corresponding standard basis vector as $\mathbf{e}_i = \zeta_i \mathbf{z} + \sqrt{1-\zeta_i^2}\tilde{\mathbf{e}}_i$ for a unit vector $\tilde{\mathbf{e}}_i \perp \mathbf{z}$ and $\zeta_i \in [-1, 1]$. Then $a_i = \mathbf{a}^T \mathbf{e}_i = \mathbf{a}^T(\zeta_i \mathbf{z} + \sqrt{1-\zeta_i^2}\tilde{\mathbf{e}}_i) = \sqrt{1-\zeta_i^2}\mathbf{a}^T\tilde{\mathbf{e}}_i$, which implies $|a_i| \le |\mathbf{a}^T\tilde{\mathbf{e}}_i|$. Since $\mathbf{a}$ and $\tilde{\mathbf{e}}_i$ are both unit vectors in $\operatorname{span}\{\mathbf{s}_1, \dots, \mathbf{s}_{p-1}\}$, our task reduces to estimating the inner-product between the uniformly distributed *random* vector $\mathbf{a}$ on the $(p-1)$-dimensional unit sphere and a *fixed* vector $\tilde{\mathbf{e}}_i$ on the sphere. Since $\mathbf{a}$ is random, we may replace $\tilde{\mathbf{e}}_i$ with another fixed vector, say $\mathbf{s}_1$. Namely, $\mathbf{a}^T\tilde{\mathbf{e}}_i$ has the same distribution as $\mathbf{a}^T\mathbf{s}_1 = \xi_1/\|\boldsymbol{\xi}\|$. Standard tail bounds for the Gaussian and χ^2 distributions (note that $\|\boldsymbol{\xi}\|^2 \sim \chi_{p-1}^2$) imply that $\frac{|\xi_1|}{\|\boldsymbol{\xi}\|} \le h\sqrt{\frac{\log p}{p}}$ with probability at least $1 - 1/p^4$, for a suitable constant $h > 0$. The lemma follows by a union bound over all p coordinates of $\mathbf{a}$. $\square$

5. Preliminaries. In this section we record a few standard results that will be used later in the proofs. The first is a large deviation result for a Chi-square random variable.

LEMMA 5.1 ([24]). *Let $X \sim \chi_n^2$. For all $x \ge 0$,*

$$\Pr[X \ge n + 2\sqrt{nx} + x] \le e^{-x} \quad \textit{and} \quad \Pr[X \le n - 2\sqrt{nx}] \le e^{-x}.$$

The second lemma records a well-known argument about the inner-product of two high-dimensional Gaussians.

LEMMA 5.2. *Let $\{x_i, y_i\}_{i=1}^n$ be standard i.i.d. Gaussian random variables. Then $\sum_{i=1}^n x_i y_i$ is distributed like the product of two independent random variables $\|\mathbf{x}\| \cdot \tilde{y}$, where $\mathbf{x} = (x_1, \dots, x_n)$, $\|\mathbf{x}\|^2 \sim \chi_n^2$ and $\tilde{y}$ is a standard Gaussian.*

PROOF. For every fixed realization of $\mathbf{x}$, we have $x_i y_i \sim N(0, x_i^2)$ and by the independence of the y_i's,

$$\sum_{i=1}^{n} x_i y_i \sim N(0, \|\mathbf{x}\|^2) = \|\mathbf{x}\| \cdot N(0,1) := \|\mathbf{x}\| \cdot \tilde{y}.$$

The lemma follows by observing that $\|\mathbf{x}\|^2 \sim \chi_n^2$. □

The next proposition establishes an upper bound on $\lambda_{\max}(\hat{\Sigma})$, the maximal eigenvalue of the sample covariance matrix $\hat{\Sigma}$, in the single-spike model, in two regimes: (i) $p/n^\alpha \to c$ for positive $c, \alpha \geq 1$ and (ii) $p/n \to 0$. The spectrum of the covariance matrix has been studied extensively in the literature. Specifically, both Baik and Silverstein [5], Theorem 1.2 and Johnstone [20], Theorem 1.1, provide the limiting behavior of $\lambda_{\max}(\hat{\Sigma})$ for $p/n \to c \geq 1$ (i.e., $\alpha = 1$). The regime of a fixed p with $n \to \infty$ which implies $p/n \to 0$ was analyzed in [22], Chapter 3, for example. Since we could not locate a reference for the case $p/n^\alpha \to c$ and $\alpha > 1$, or for $p/n \to 0$ and p not necessarily fixed, we provide the following proposition. The proof uses standard arguments and is given in Section 9.

PROPOSITION 5.3. *Let $\hat{\Sigma}$ be a $p \times p$ sample covariance matrix of n samples in the k-sparse single-spike model with signal strength $\beta > 0$, arbitrary k and either*: (i) $p/n \to 0$ *or* (ii) $p/n^\alpha \to c$ *for positive constants* $c, \alpha \geq 1$. *Then there exists an $\varepsilon = \varepsilon(n) \to 0$ such that with probability tending to one as $n \to \infty$,*

$$\lambda_{\max}(\hat{\Sigma}) \leq (1+\varepsilon)\left(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta}\right)^2. \tag{5.1}$$

COROLLARY 5.4. *Let $\hat{\Sigma}$ be a $p \times p$ sample covariance matrix of n samples and a k-sparse spike $\mathbf{z}$ with signal strength $\beta > 0$. Further assume that $k/n \to 0$. Then there exists an $\varepsilon = \varepsilon(n) \to 0$ such that with probability tending to one as $n \to \infty$, for every rank-one trace-one $p \times p$ matrix $Y = \mathbf{y}\mathbf{y}^T$ with* $\operatorname{supp}(\mathbf{y}) \subseteq \operatorname{supp}(\mathbf{z})$,

$$\langle \hat{\Sigma}, Y \rangle \leq (1+\varepsilon)(1+\sqrt{\beta})^2.$$

PROOF. Consider $\sup_Y \langle \hat{\Sigma}, Y \rangle$ where Y ranges over all matrices Y as stated above. For each such $Y = \mathbf{y}\mathbf{y}^T$, we have $\|\mathbf{y}\|^2 = \sum_i \mathbf{y}_i^2 = \sum_i Y_{ii} = \operatorname{tr}(Y) = 1$. Let $\mathbf{y}_\mathbf{z} \in \mathbb{R}^k$ be the projection of $\mathbf{y} \in \mathbb{R}^p$ on the coordinates of $\operatorname{supp}(\mathbf{z})$, then $\|\mathbf{y}_\mathbf{z}\| = \|\mathbf{y}\| = 1$. Similarly, let $\hat{\Sigma}_\mathbf{z}$ be the $k \times k$ submatrix of $\hat{\Sigma}$ corresponding to $\operatorname{supp}(\mathbf{z})$, namely, restricting it to the first k rows and first k columns. Observe that we can write

$$\langle \hat{\Sigma}, Y \rangle = \operatorname{tr}(\hat{\Sigma}\mathbf{y}\mathbf{y}^T) = \mathbf{y}^T \hat{\Sigma} \mathbf{y} = \mathbf{y}_\mathbf{z}^T \hat{\Sigma}_\mathbf{z} \mathbf{y}_\mathbf{z} \leq \lambda_{\max}(\hat{\Sigma}_\mathbf{z}),$$

hence $\sup_Y \langle \hat{\Sigma}, Y \rangle \leq \lambda_{\max}(\hat{\Sigma}_\mathbf{z})$. Now the desired upper bound on $\lambda_{\max}(\hat{\Sigma}_\mathbf{z})$ follows using the fact $k/n \to 0$ from Proposition 5.3, that is, plugging $p = k$ into (5.1). □

Our next proposition estimates $\mathrm{tr}(\hat{\Sigma})$ and $\mathrm{tr}(\hat{\Sigma}^2)$ for the case $\beta = 0$ (no signal). These estimates were derived in [25], Proposition 1, for example, but again only for $\alpha = 1$. For lack of reference we reprove it for $\alpha \geq 1$ in Section 9.

PROPOSITION 5.5. *Let $\hat{\Sigma}$ be a $p \times p$ sample covariance matrix of n multivariate Gaussian observations whose population covariance matrix is the identity. Assume that $(\log p)/n \to 0$ as $n, p \to \infty$. Then there exists an $\varepsilon = \varepsilon(n) \to 0$ such that with probability tending to one as $n \to \infty$,*

$$(1-\varepsilon)p \leq \mathrm{tr}(\hat{\Sigma}) \leq (1+\varepsilon)p,$$
$$(1-\varepsilon)p\left(1+\frac{p}{n}\right) \leq \mathrm{tr}(\hat{\Sigma}^2) \leq (1+\varepsilon)p\left(1+\frac{p}{n}\right).$$

6. Proof of Theorem 1.2 (Cosine similarity). Let X be a solution to SDP (1.2), with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ and a corresponding orthonormal set of eigenvectors $\hat{\mathbf{z}} = \mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^p$. We can then write $X = \sum_{i=1}^p \lambda_i \mathbf{v}_i \mathbf{v}_i^T$, and by linearity of the Frobenius inner-product, $\langle \hat{\Sigma}, X \rangle = \sum_{i=1}^p \lambda_i \mathbf{v}_i^T \hat{\Sigma} \mathbf{v}_i$. Using the simple observations $\mathbf{v}_i^T \hat{\Sigma} \mathbf{v}_i \leq \lambda_{\max}(\hat{\Sigma})$ (by the variational characterization of eigenvalues) and $\sum_i \lambda_i = \mathrm{tr}(X) = 1$, we get

$$\langle \hat{\Sigma}, X \rangle = \sum_{i=1}^p \lambda_i \mathbf{v}_i^T \hat{\Sigma} \mathbf{v}_i \leq \lambda_1 \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}} + (1-\lambda_1) \cdot \lambda_{\max}(\hat{\Sigma}). \tag{6.1}$$

Let us first provide a high-level description of the proof idea. We can bound $\langle \hat{\Sigma}, X \rangle$ from below (using Theorem 1.4, which we prove in Section 7, and as mentioned earlier is used here) and $\lambda_{\max}(\hat{\Sigma})$ from above (using Proposition 5.3) both by roughly $\frac{p}{n}$. Now suppose λ_1 is not too small; then on the right-hand side of (6.1), a large contribution must come from the first term $\lambda_1 \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}}$. But the quadratic form $\hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}}$ has small value in the direction $\hat{\mathbf{z}} = \mathbf{z}$ (using Corollary 5.4), and thus $\hat{\mathbf{z}}$ and $\mathbf{z}$ cannot be too close to each other.

We now proceed to the detailed proof, starting with a lower bound on $\lambda_1 \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}}$. Assume henceforth that the high-probability event asserted by Theorem 1.4 indeed occurs; namely, inequality (1.8) holds. Similarly Corollary 5.4 implies that inequality (5.1) holds. Plugging these two bounds into (6.1) and using $\lambda_1 > 0$, we get

$$\begin{aligned}\lambda_1 \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}} &\geq (1-\zeta)\frac{p}{n} - (1-\lambda_1)(1+\varepsilon)\left(1+\sqrt{\frac{p}{n}}+\sqrt{\beta}\right)^2 \\ &\geq \frac{p}{n}(\lambda_1 - \zeta - \varepsilon) - (1+\varepsilon)\left(2\sqrt{\frac{p}{n}}+1+\sqrt{\beta}\right)(1+\sqrt{\beta}).\end{aligned}$$

Observe that $\sqrt{\frac{p}{n}} \geq \frac{1}{1+\varepsilon_1}$ for suitable $\varepsilon_1 \to 0$ and sufficiently large n, p. In addition, assumption (b) yields that $\beta \leq \sqrt{\frac{p}{n}} \leq (1+\varepsilon_1)\frac{p}{n}$. For suitable $\varepsilon_2 = O(\zeta + \varepsilon +$

ε_1), we get

$$\lambda_1 \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}} \geq \frac{p}{n}(\lambda_1 - \varepsilon_2) - (4 + \varepsilon_2)\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}). \tag{6.2}$$

Next, we analyze the quadratic form $\hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}}$ in terms of $\gamma = \langle \hat{\mathbf{z}}, \mathbf{z} \rangle \in [-1, 1]$. Write $\hat{\mathbf{z}} = \gamma \mathbf{z} + \sqrt{1 - \gamma^2}\mathbf{s}$, where $\mathbf{s}$ is a unit vector orthogonal to $\mathbf{z}$, and recall that our goal is to upper bound γ^2. Using Cauchy–Schwarz and the triangle inequality,

$$\begin{aligned} \hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}} &\leq \|\hat{\mathbf{z}}\| \cdot \|\hat{\Sigma} \hat{\mathbf{z}}\| = 1 \cdot \|\hat{\Sigma}(\gamma \mathbf{z} + \sqrt{1 - \gamma^2}\mathbf{s})\| \\ &\leq |\gamma| \cdot \|\hat{\Sigma} \mathbf{z}\| + \sqrt{1 - \gamma^2}\|\hat{\Sigma} \mathbf{s}\|. \end{aligned} \tag{6.3}$$

Since $\hat{\Sigma}$ is PSD, it can be written as $\hat{\Sigma} = B^T B$ for some matrix B whose spectral norm is $\|B\| = \|B^T\| = \sqrt{\lambda_{\max}(\hat{\Sigma})}$. Assume henceforth that the high-probability event asserted by Corollary 5.4 indeed occurs, and we have $\|B\mathbf{z}\|^2 = \mathbf{z}^T \hat{\Sigma} \mathbf{z} = \langle \hat{\Sigma}, \mathbf{z}\mathbf{z}^T \rangle \leq (1 + \varepsilon_3)(1 + \sqrt{\beta})^2$ for suitable $\varepsilon_3 \to 0$. Using Proposition 5.3 similarly yields $\|B\|^2 = \lambda_{\max}(\hat{\Sigma}) \leq (1 + \varepsilon_4)(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta})^2$ for suitable $\varepsilon_4 \to 0$. Together, for suitable $\varepsilon_5 = O(\varepsilon_3 + \varepsilon_4 + \varepsilon_1)$,

$$\|\hat{\Sigma} \mathbf{z}\| \leq \|B^T\| \cdot \|B\mathbf{z}\| \leq (3 + \varepsilon_5)\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}),$$

and similarly

$$\|\hat{\Sigma} \mathbf{s}\| \leq \lambda_{\max}(\hat{\Sigma}) \cdot \|\mathbf{s}\| \leq \frac{p}{n}(1 + \varepsilon_4) + (4 + \varepsilon_5)\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}).$$

Plugging these into (6.3) and using $|\gamma| \leq 1$ and $\sqrt{1 - \gamma^2} \leq 1 - \frac{\gamma^2}{2} \leq 1$, we have

$$\hat{\mathbf{z}}^T \hat{\Sigma} \hat{\mathbf{z}} \leq \left(1 - \frac{\gamma^2}{2} + \varepsilon_4\right)\frac{p}{n} + (7 + 2\varepsilon_5)\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}). \tag{6.4}$$

Now combining this upper bound (6.4) with our lower bound 6.2 (after dividing by λ_1), gives

$$\begin{aligned} &\left(1 - \frac{\varepsilon_2}{\lambda_1}\right)\frac{p}{n} - \frac{4 + \varepsilon_2}{\lambda_1}\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}) \\ &\qquad \leq \left(1 - \frac{\gamma^2}{2} + \varepsilon_4\right)\frac{p}{n} + (7 + 2\varepsilon_5)\sqrt{\frac{p}{n}}(1 + \sqrt{\beta}), \end{aligned}$$

and by further manipulation,

$$\frac{\gamma^2}{2} - \frac{\varepsilon_2}{\lambda_1} - \varepsilon_4 \leq \frac{11 + 2\varepsilon_5 + \varepsilon_2}{\lambda_1}\sqrt{\frac{n}{p}}(1 + \sqrt{\beta}).$$

For sufficiently large n, p, this yields the bound on $\gamma^2 = |\langle \hat{\mathbf{z}}, \mathbf{z} \rangle|^2$ asserted in (1.5), and completes the proof of Theorem 1.2.

7. Proof of Theorem 1.4 (SDP value). We start with the upper bound on $\langle\hat{\Sigma}, X\rangle$. The idea is to drop the constraint $\|X\|_S \le k$ from SDP (1.2), and show that the value of the resulting SDP, which can only be bigger, is actually $\lambda_{\max}(\hat{\Sigma})$, and is thus bounded by Proposition 5.3.

Formally, let X be a solution to SDP (1.2), and let us argue that (with probability 1)

$$\langle\hat{\Sigma}, X\rangle \le \sup\{\langle\hat{\Sigma}, Y\rangle : Y \in \mathcal{S}_+^p, \operatorname{tr}(Y) = 1\} = \lambda_{\max}(\hat{\Sigma}).$$

Indeed, the inequality holds because we have just relaxed SDP (1.2). The equality holds by the following standard argument. Writing $Y = \sum_i \mu_i \mathbf{y}_i \mathbf{y}_i^T$, where $\{\mu_i\}_i$ are the eigenvectors of Y and $\{\mathbf{y}_i\}_i$ is a corresponding orthonormal eigenbasis, we have

$$\begin{aligned}\langle\hat{\Sigma}, Y\rangle &= \sum_i \mu_i \mathbf{y}_i^T \hat{\Sigma} \mathbf{y}_i \le \lambda_{\max}(\hat{\Sigma}) \cdot \sum_i \mu_i = \lambda_{\max}(\hat{\Sigma}) \cdot \operatorname{tr}(Y) \\ &= \lambda_{\max}(\hat{\Sigma}),\end{aligned}$$

and equality is achieved when maximizing over all relevant Y, by taking $Y = \mathbf{y}_1 \mathbf{y}_1^T$ to be a rank-one matrix where $\mathbf{y}_1$ is a leading eigenvector of $\hat{\Sigma}$.

To conclude the upper bound asserted in the theorem, we combine the above with Proposition 5.3, and get that for a suitable $\varepsilon = \varepsilon(n) \to 0$ with probability tending to one as $n \to \infty$,

$$\langle\hat{\Sigma}, X\rangle \le \lambda_{\max}(\hat{\Sigma}) \le (1+\varepsilon)\left(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta}\right)^2.$$

We turn to proving the lower bound on $\langle\hat{\Sigma}, X\rangle$. The idea is to consider a specific X^* which is feasible (but not necessarily optimal) for SDP (1.2), and compute its objective value $\langle\hat{\Sigma}, X^*\rangle$. Our X^* is based on taking the nonsignal part of $\hat{\Sigma}$ [which is a $(p-k) \times (p-k)$ submatrix], padded with zeros elsewhere, and "forcing" it to satisfy the constraints of SDP (1.2) by scaling it to be trace-one.

Formally, let $X^* = \tilde{\Sigma} / \operatorname{tr}(\tilde{\Sigma})$, where the matrix $\tilde{\Sigma} \in \mathbb{R}^{p\times p}$ is given by

$$\tilde{\Sigma}_{ij} = \begin{cases} 0, & \text{if } i \le k \text{ or } j \le k; \\ \hat{\Sigma}_{ij}, & \text{otherwise.} \end{cases} \tag{7.1}$$

We prove below that with probability tending to one, the following inequalities hold for a suitable $\zeta = \zeta(n) \to 0$:

$$\|X^*\|_S \le \frac{2p}{\sqrt{n}} \le k, \tag{7.2}$$

$$\langle\hat{\Sigma}, X^*\rangle \ge (1-\zeta)\left(1 + \frac{p}{n}\right). \tag{7.3}$$

Combining this with $X^* \in \mathcal{S}_+^p$ and $\operatorname{tr}(X^*) = 1$, which hold by construction, will prove that with probability tending to one, X^* is feasible and has a high-objective value.

Before proceeding to prove (7.3) and (7.2), we observe that the nonzeroed part of $\tilde{\Sigma}$ satisfies the conditions of Proposition 5.5, as it is a $(p-k)\times(p-k)$ sample covariance matrix of n multivariate Gaussian observations whose population covariance matrix is the identity, and furthermore $(\log(p-k))/n \to 0$. The concrete bounds that we get hold for a suitable $\varepsilon = \varepsilon(n) \to 0$ and with probability tending to one as $n \to \infty$, and roughly say that $\operatorname{tr}(\tilde{\Sigma}) \approx p-k$ and $\operatorname{tr}(\tilde{\Sigma}^2) \approx (p-k)(1+\frac{p-k}{n})$.

Let us now prove inequality (7.2). First, using Cauchy–Schwarz,

$$\|X^*\|_S = \frac{1}{\operatorname{tr}(\tilde{\Sigma})}\sum_{i,j>k}|\tilde{\Sigma}_{ij}| \le \frac{\sqrt{(p-k)^2\sum_{i,j>k}\tilde{\Sigma}_{ij}^2}}{\operatorname{tr}(\tilde{\Sigma})} = (p-k)\frac{\sqrt{\operatorname{tr}(\tilde{\Sigma}^2)}}{\operatorname{tr}(\tilde{\Sigma})}.$$

By the above bounds from Proposition 5.5, with probability tending to one,

$$\begin{aligned}\frac{\sqrt{\operatorname{tr}(\tilde{\Sigma}^2)}}{\operatorname{tr}(\tilde{\Sigma})} &\le \frac{\sqrt{(1+\varepsilon)(p-k)(1+(p-k)/n)}}{(1-\varepsilon)(p-k)} \le \frac{\sqrt{(1+\varepsilon)(1/(p-k)+1/n)}}{1-\varepsilon}\\ &\le \frac{1+\varepsilon}{1-\varepsilon}\sqrt{\frac{3}{n}},\end{aligned}$$

which together imply that $\|X^*\|_S \le (p-k)\frac{1+\varepsilon}{1-\varepsilon}\sqrt{\frac{3}{n}} \le \frac{2p}{\sqrt{n}} \le k$.

We next prove inequality (7.3). First, we expand

$$\langle\hat{\Sigma}, X^*\rangle = \frac{1}{\operatorname{tr}(\tilde{\Sigma})}\sum_{i,j}\tilde{\Sigma}_{ij}\hat{\Sigma}_{ij} = \frac{1}{\operatorname{tr}(\tilde{\Sigma})}\sum_{i,j}\tilde{\Sigma}_{ij}^2 = \frac{\operatorname{tr}(\tilde{\Sigma}^2)}{\operatorname{tr}(\tilde{\Sigma})}.$$

By the above bounds from Proposition 5.5, with probability tending to one,

$$\frac{\operatorname{tr}(\tilde{\Sigma}^2)}{\operatorname{tr}(\tilde{\Sigma})} \ge \frac{(1-\varepsilon)(p-k)(1+p/n)(1-k/p)}{(1+\varepsilon)(p-k)} \ge (1-\zeta)\left(1+\frac{p}{n}\right),$$

for a suitable $\zeta = \zeta(n) \to 0$, where we used here that $k/p \to 0$ by assumption (c). Altogether, we conclude that $\langle\hat{\Sigma}, X^*\rangle \ge (1-\zeta)(1+\frac{p}{n})$.

Having proved inequalities (7.3) and (7.2), we conclude that with probability tending to one, X^* is feasible and has a high objective value, which establishes a lower bound on the optimal SDP value $\langle\hat{\Sigma}, X\rangle$, and completes the proof of Theorem 1.4.

8. Proof of Theorem 1.5 (SDP value).

Let F be the set of all vectors $\mathbf{y}$ whose corresponding rank-one matrix $Y = \mathbf{y}\mathbf{y}^T$ is feasible for SDP (1.2), formally,

$$F = \{\mathbf{y} \in \mathbb{R}^p : \|\mathbf{y}\|_2 \le 1 \text{ and } \|\mathbf{y}\|_1 \le \sqrt{k}\}.$$

We need to prove that with probability tending to one as $n \to \infty$, every $Y = \mathbf{y}\mathbf{y}^T$ such that $\mathbf{y} \in F$ satisfies $\langle\hat{\Sigma}, Y\rangle \le \frac{8}{9}\frac{p}{n}$. At a high level, $\langle\hat{\Sigma}, Y\rangle = \mathbf{y}^T\hat{\Sigma}\mathbf{y}$ is continuous, and thus a standard approach is to discretize F with an ε-net, analyze every single point in F separately and apply a union bound argument. The size of an ε-net for the unit ℓ_2-ball in p dimensions is proportional to $(1/\varepsilon)^p$. On the other hand, our upper bound on the probability that a fixed $Y = \mathbf{y}\mathbf{y}^T$ violates $\langle \hat{\Sigma}, Y \rangle \le \frac{8}{9}\frac{p}{n}$ is larger than ε^p; see Lemma 8.2. Therefore, a naive discretization of F fails, and we need to reduce the size of the net by using the additional constraint $\|\mathbf{y}\|_1 \le \sqrt{k}$. To this end, we approximate F by a set $\hat{F}$, whose definition uses an ℓ_0-constraint; the idea is that an ℓ_0-bound is technically more convenient than ℓ_1. We apply an ε-net argument to $\hat{F}$, which indirectly yields a bound for all of F. Specifically, we define

$$\hat{F} = \{\mathbf{y} \in \mathbb{R}^p : \|\mathbf{y}\|_2 \le 1 \text{ and } \|\mathbf{y}\|_0 \le 40\sqrt{pk}\}.$$

To formalize the notion of one set approximating another one, we define the r-*neighborhood* of a set $A \subset \mathbb{R}^p$ to be $A_r = \{\mathbf{y} \in \mathbb{R}^p : \exists \mathbf{y}' \in A, \|\mathbf{y} - \mathbf{y}'\| \le r\}$.

LEMMA 8.1. *The sets $\hat{F}$, F defined above satisfy $F \subseteq \hat{F}_{1/40}$.*

PROOF. Fix $\mathbf{y} \in F$, and let $I = \{i \in [p] : |\mathbf{y}_i| \ge 1/(40\sqrt{p})\}$. Since $\|\mathbf{y}\|_1 \le \sqrt{k}$, the size of I is at most $|I| \le 40\sqrt{kp}$. Now define $\mathbf{y}' \in \mathbb{R}^p$ as follows: $\mathbf{y}'_i = \mathbf{y}_i$ if $i \in I$, and $\mathbf{y}'_i = 0$ otherwise. By construction, $\mathbf{y}' \in \hat{F}$ and $\|\mathbf{y}' - \mathbf{y}\|^2 \le p \cdot 1/(40\sqrt{p})^2 = 1/40^2$. $\square$

We proceed to the discretization of $\hat{F}$, which uses the following notation. For $B \subseteq \mathbb{R}^p$ and a subset of the coordinates $I \subseteq [p]$, let $B_I \subseteq B$ denote the vectors in B whose support is contained in I. Recall that an ε-*net* of $B \subseteq \mathbb{R}^p$ is a subset $N \subseteq B$ satisfying $B \subseteq N_\varepsilon$ and that for all $\mathbf{x} \ne \mathbf{y} \in N$, $\|\mathbf{x} - \mathbf{y}\| > \varepsilon$. Setting $\mathcal{I} = \{I \subseteq [p] : |I| = 40\sqrt{pk}\}$, clearly $\hat{F} = \bigcup_{I \in \mathcal{I}} \hat{F}_I$. Let N_I be an ε-net of $\hat{F}_I$ with $\varepsilon = 1/40$, and let $\tilde{N}$ be the union of all these nets, that is,

$$\tilde{N} = \bigcup_{I \in \mathcal{I}} N_I.$$

Then $\hat{F} = \bigcup_{I \in \mathcal{I}} \hat{F}_I \subseteq \bigcup_{I \in \mathcal{I}} (N_I)_{1/40} \subseteq \tilde{N}_{1/40}$. Now using Lemma 8.1 and the triangle inequality, we get that $F \subseteq \hat{F}_{1/40} \subseteq \tilde{N}_{1/20}$. The key to completing the proof is to show that for all sufficiently large n,

$$\Pr[\forall \tilde{\mathbf{y}} \in \tilde{N}, \tilde{\mathbf{y}}^T \hat{\Sigma} \tilde{\mathbf{y}} \le 2(1 + \beta)] \ge 1 - e^{-n/10}. \tag{8.1}$$

Before proving this inequality, let us rely on it to complete the proof of Theorem 1.5. Assume the high-probability event in (8.1) indeed occurs, and similarly for Proposition 5.3, hence $\lambda_{\max}(\hat{\Sigma}) \le (1 + \varepsilon)(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta})^2$. Now because $F \subseteq \tilde{N}_{1/20}$, for every $\mathbf{y} \in F$ there exists $\tilde{\mathbf{y}} \in \tilde{N}$ such that $\mathbf{a} = \mathbf{y} - \tilde{\mathbf{y}}$ is of length $\|\mathbf{a}\| \le 1/20$, and therefore for $Y = \mathbf{y}\mathbf{y}^T$,

$$\langle \hat{\Sigma}, Y \rangle = \mathbf{y}^T \hat{\Sigma} \mathbf{y} = \tilde{\mathbf{y}}^T \hat{\Sigma} \tilde{\mathbf{y}} + 2\mathbf{a}^T \hat{\Sigma} \tilde{\mathbf{y}} + \mathbf{a}^T \hat{\Sigma} \mathbf{a}. \tag{8.2}$$

The assumption we made using (8.1) implies that $\tilde{\mathbf{y}}^T\hat{\Sigma}\tilde{\mathbf{y}} \le 2(1+\beta)$. To bound the two summands, observe that for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^p$ we have $\mathbf{u}^T\hat{\Sigma}\mathbf{v} \le \|\mathbf{u}\|\|\mathbf{v}\|\lambda_{\max}(\hat{\Sigma})$. Plugging all these into (8.2), we get

$$\langle\hat{\Sigma}, Y\rangle \le 2(1+\beta) + (1+\varepsilon)\left(\frac{2}{20} + \frac{1}{20^2}\right)\left(1 + \sqrt{\frac{p}{n}} + \sqrt{\beta}\right)^2.$$

Recall that $\beta \le \sqrt{\frac{p}{n}}$ and that for sufficiently large n, p we have $\frac{p}{n} \ge 20$. Hence by straightforward manipulations, we conclude that as $(n, p, k) \to \infty$, with probability tending to one $\langle\hat{\Sigma}, Y\rangle \le \frac{8}{9}\frac{p}{n}$, which proves Theorem 1.5.

It remains to prove (8.1), which we do via a union bound argument, using the two lemmas below. The first lemma estimates the probability that an arbitrary fixed $\mathbf{y} \in \tilde{N}$ violates the inequality $\mathbf{y}^T\hat{\Sigma}\mathbf{y} \le 2(1+\beta)$, and the second one bounds the size of the ε-net $\tilde{N}$.

LEMMA 8.2. *Under the conditions of Theorem* 1.5, *there exists an integer* $n_0 > 0$, *such that for every* $n \ge n_0$ *and every* $\mathbf{y} \in \mathbb{R}^p$ *of length at most* 1 (*in particular, every* $y \in \tilde{N}$),

$$\Pr\bigl[\mathbf{y}^T\hat{\Sigma}\mathbf{y} \ge 2(1+\beta)\bigr] \le e^{-n/9}.$$

PROOF. Fix $\mathbf{y} \in \mathbb{R}^p$ with $\|y\| \le 1$, and expand

$$\mathbf{y}^T\hat{\Sigma}\mathbf{y} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{y} = \frac{1}{n}\sum_{i=1}^{n}\langle\mathbf{x}_i, \mathbf{y}\rangle^2.$$

Recall from (1.3) that $\mathbf{x}_i = \sqrt{\beta}u_i\mathbf{z} + \boldsymbol{\xi}_i$, where $\boldsymbol{\xi}_i$ is a vector of independent standard Gaussian random variables, and u_i is also a standard Gaussian. Therefore,

$$\langle\mathbf{x}_i, \mathbf{y}\rangle = \langle\boldsymbol{\xi}_i, \mathbf{y}\rangle + u_i\sqrt{\beta}\langle\mathbf{y}, \mathbf{z}\rangle. \tag{8.3}$$

The first term $\langle\boldsymbol{\xi}_i, \mathbf{y}\rangle$ has distribution $N(0, \|\mathbf{y}\|^2)$. Since u_i is independent of $\boldsymbol{\xi}_i$, the distribution of $\langle\mathbf{x}_i, \mathbf{y}\rangle$ is just $N(0, \|\mathbf{y}\|^2 + \beta\langle\mathbf{y}, \mathbf{z}\rangle^2)$. Furthermore, since $\mathbf{y}$ is fixed and the $\boldsymbol{\xi}_i$'s and u_i's are all independent, the random variables $\langle\mathbf{x}_i, \mathbf{y}\rangle$ for $i = 1, \dots, n$ are i.i.d., and thus

$$\sum_{i=1}^{n}\langle\mathbf{x}_i, \mathbf{y}\rangle^2 \sim \bigl(\|\mathbf{y}\|^2 + \beta\langle\mathbf{y}, \mathbf{z}\rangle^2\bigr)\chi_n^2.$$

Lemma 5.1 with $x = n/9$ implies that $\Pr[\chi_n^2 \ge 2n] \le e^{-n/9}$. We conclude that with probability at least $1 - e^{-n/9}$,

$$\bigl\langle\hat{\Sigma}, \mathbf{y}\mathbf{y}^T\bigr\rangle \le \frac{1}{n}\cdot 2n\bigl(\|\mathbf{y}\|^2 + \beta\langle\mathbf{y}, \mathbf{z}\rangle^2\bigr) \le 2(1+\beta),$$

where the second inequality uses the Cauchy–Schwarz inequality. □

LEMMA 8.3. *The ε-net $\tilde{N}$ has size $|\tilde{N}| \le p^{20\sqrt{pk}}$.*

PROOF. By the definition of $\tilde{N}$ and the fact that $|N_I|$ is the same for all I, we can fix arbitrary $I \in \mathcal{I}$ and write

$$|\tilde{N}| \le \binom{p}{40\sqrt{pk}} |N_I|. \tag{8.4}$$

We thus need to bound $|N_I|$. By definition, N_I is contained in an axis-aligned subspace of $\mathbb{R}^p$ of dimension $p' = 40\sqrt{pk}$, and we can use the following standard volume argument. Ignoring henceforth all coordinates outside I, let $\mathcal{B}_r(\mathbf{x})$ be a closed ball (in $\mathbb{R}^{p'}$) of radius $r > 0$ centered at $\mathbf{x}$. Since N_I is an ε-net (of $\hat{F}_I$), for every two distinct points in it, $\mathbf{x} \ne \mathbf{y} \in N_I$, the corresponding balls $\mathcal{B}_{\varepsilon/2}(\mathbf{x})$ and $\mathcal{B}_{\varepsilon/2}(\mathbf{y})$ are disjoint (as otherwise $\|\mathbf{x}-\mathbf{y}\| \le \varepsilon$). In addition, the union of these balls $\mathcal{B}_{\varepsilon/2}(\mathbf{x})$ over all $\mathbf{x} \in N_I$ is contained in $\mathcal{B}_{1+\varepsilon/2}(\mathbf{0})$ (because all $\mathbf{x} \in N_I \subseteq \hat{F}_I$ have length at most 1). Recalling that the Euclidean volume of a ball of radius $r > 0$ in dimension d grows with r proportionally to r^d, and plugging in $\varepsilon = 1/40$, we obtain

$$|N_I| \le \frac{\mathrm{vol}(\mathcal{B}_{1+\varepsilon/2}(\mathbf{0}))}{\mathrm{vol}(\mathcal{B}_{\varepsilon/2}(\mathbf{0}))} \le \left(\frac{1+\varepsilon/2}{\varepsilon/2}\right)^{p'} = 81^{40\sqrt{pk}}.$$

Plugging into (8.4), we get $|\tilde{N}| \le (\frac{ep}{40\sqrt{pk}})^{40\sqrt{pk}} \cdot 81^{40\sqrt{pk}} \le p^{20\sqrt{pk}}$. $\square$

Finally, observe that (8.1) indeed follows from Lemmas 8.2 and 8.3 by a union bound,

$$\Pr\bigl[\exists \tilde{\mathbf{y}} \in \tilde{N}, \tilde{\mathbf{y}}^T \hat{\Sigma} \tilde{\mathbf{y}} \ge 2(1+\beta)\bigr] \le p^{20\sqrt{pk}} \cdot e^{-n/9} \le e^{-n/10},$$

where the last inequality follows from the assumption in Theorem 1.5 that $k/(p/\log^2 p) \to 0$ and that $p/n \to c$. This completes the proof of (8.1) and of Theorem 1.5.

9. Deferred proofs from Section 5 (preliminaries).

PROOF OF PROPOSITION 5.3. Let us rotate $\mathbb{R}^p$ so that the spike $\mathbf{z}$ becomes the first standard basis vector $\mathbf{e}_1$. Obviously, $\lambda_{\max}(\hat{\Sigma})$ would not change at all, and since the normal distribution is rotation invariant, the noise would still be normally distributed. In effect, we may assume henceforth that $\mathbf{z} = \mathbf{e}_1$. Recalling from (1.3) that the samples are given by $\mathbf{x}_i = \sqrt{\beta} u_i \mathbf{z} + \boldsymbol{\xi}_i$, we can write $\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T$ as

$$\hat{\Sigma} = \frac{1}{n}(\sqrt{\beta}U + \Xi)(\sqrt{\beta}U + \Xi)^T,$$

where U is a $p \times n$ matrix whose first row is $(u_1, \dots, u_n)$ and the remaining rows are zero (recall $\mathbf{z} = \mathbf{e}_1$), and Ξ is an $p \times n$ matrix whose ith column is $\boldsymbol{\xi}_i$. Let

$\|A\| = \sqrt{\lambda_{\max}(A^T A)}$ be the spectral norm of a matrix A. Using also $\|A\| = \|A^T\|$ and the triangle inequality,

$$\lambda_{\max}(\hat{\Sigma}) = \frac{1}{n}\|\sqrt{\beta}U + \Xi\|^2 \le \frac{1}{n}(\sqrt{\beta}\|U\| + \|\Xi\|)^2. \tag{9.1}$$

The matrix $\Xi^T \Xi$ follows a Wishart distribution (note that the roles of p and n are reversed). Therefore by [17], Theorem 2, which applies to the regime $p/n \to 0$ and $p/n \to \infty$, and by [20], Theorem 1.1, which applies to $p/n \to c \in (0, \infty)$, we know that with probability tending to one,

$$\|\Xi\|^2 = \lambda_{\max}(\Xi^T \Xi) \le (1 + \varepsilon_1)(\sqrt{p} + \sqrt{n})^2,$$

for some $\varepsilon_1 = \varepsilon_1(n) \to 0$. Since $U^T U$ has rank one, $\|U\|^2 = \lambda_{\max}(U^T U) = \operatorname{tr}(U^T U) = \sum_{i=1}^n u_i^2 \sim \chi_n^2$. Lemma 5.1 with $x = 3\log n$ implies that with probability at least $1 - 1/n^3$,

$$\|U\|^2 \le \left(1 + O\left(\frac{\log n}{n}\right)\right)n,$$

and thus with probability tending to one, $\|U\| \le (1 + \varepsilon_2)\sqrt{n}$ for some $\varepsilon_2 = \varepsilon_2(n) \to 0$.

Plugging these bounds into (9.1), we conclude that with probability tending to one as $n \to \infty$,

$$\begin{aligned}\lambda_{\max}(\hat{\Sigma}) &\le \left[(1 + \varepsilon_2)\sqrt{\beta} + (1 + \varepsilon_1)\left(1 + \sqrt{\frac{p}{n}}\right)\right]^2 \\ &\le \left[(1 + \varepsilon_1 + \varepsilon_2)\left(\sqrt{\beta} + 1 + \sqrt{\frac{p}{n}}\right)\right]^2,\end{aligned}$$

which completes the proof of Proposition 5.3. □

PROOF OF PROPOSITION 5.5. Starting with $\operatorname{tr}(\hat{\Sigma})$, observe that $\hat{\Sigma}_{ii} \sim \frac{1}{n}\chi_n^2$. Lemma 5.1 with $x = 5\ln p$ implies that with probability at least $1 - 1/p^5$, $\chi_n^2 \le (1 + \varepsilon_1)n$ for $\varepsilon_1 = O(\sqrt{(\log p)/n}) \to 0$. Taking a union bound over $i = 1, \dots, p$, we obtain that with probability at least $1 - 1/p^4$, all entries $\hat{\Sigma}_{ii} \in [1 - \varepsilon_1, 1 + \varepsilon_1]$, which implies

$$\operatorname{tr}(\hat{\Sigma}) = \sum_{i=1}^p \hat{\Sigma}_{ii} \in [(1 - \varepsilon_1)p, (1 + \varepsilon_1)p].$$

We now turn to bound

$$\operatorname{tr}(\hat{\Sigma}^2) = \sum_{i,j=1}^p \hat{\Sigma}_{ij}^2 = \sum_{i=1}^p \hat{\Sigma}_{ii}^2 + \sum_{i=1}^p \sum_{j \ne i} \hat{\Sigma}_{ij}^2.$$

By the preceding paragraph, with probability at least $1 - 1/p^4$, $\sum_{i=1}^{p} \hat{\Sigma}_{ii}^2 \in [(1-\varepsilon_1)^2 p, (1+\varepsilon_1)^2 p]$. Using the notation of (1.3), we write off-diagonal entries in $\hat{\Sigma}$ as $\hat{\Sigma}_{ij} = \frac{1}{n}\sum_{s=1}^{n} \boldsymbol{\xi}_{si}\boldsymbol{\xi}_{sj} := \frac{1}{n}\boldsymbol{\rho}_i^T \boldsymbol{\rho}_j$, where $\boldsymbol{\rho}_i = (\boldsymbol{\xi}_{si})_{s=1}^{n}$, and notice that $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_p$ are independent.

Now fix i and condition on $\boldsymbol{\rho}_i$. Then Lemma 5.2 implies that each off-diagonal entry along row i is distributed $\hat{\Sigma}_{ij} \sim \frac{1}{n}\|\boldsymbol{\rho}_i\| \cdot \hat{y}_j$, $\hat{y}_j \sim N(0,1)$. Moreover the $\hat{y}_j$'s (for different $j \neq i$) are independent, hence, $\sum_{j\neq i} \hat{\Sigma}_{ij}^2 \sim \frac{1}{n^2}\|\boldsymbol{\rho}_i\|^2 \chi_{p-1}^2$. Using Lemma 5.1 with $x = 4\log p$, with probability at least $1 - 1/p^4$,

$$\sum_{j\neq i} \hat{\Sigma}_{ij}^2 \in \left[(1-\varepsilon_2)(p-1)\cdot \frac{1}{n^2}\|\boldsymbol{\rho}_i\|^2, (1+\varepsilon_2)(p-1)\cdot \frac{1}{n^2}\|\boldsymbol{\rho}_i\|^2\right],$$

for $\varepsilon_2 = O(\sqrt{(\log p)/p})$.

Next, remove the conditioning on $\boldsymbol{\rho}_i$ (still for a fixed i), observing that $\|\boldsymbol{\rho}_i\|^2 \sim \chi_n^2$. Lemma 5.1 with $x = 4\log p$ then implies that with probability at least $1 - 1/p^4$, we have $\|\boldsymbol{\rho}_i\|^2 \in [(1-\varepsilon_3)n, (1+\varepsilon_3)n]$ for $\varepsilon_3 = O(\sqrt{(\log p)/n})$.

Finally, taking the union bound over rows $i = 1, \dots, p$ and also the sum along the diagonal, with probability at least $1 - 3/p^3$,

$$\operatorname{tr}(\hat{\Sigma}^2) \leq (1+\varepsilon_1)^2 p + (1+\varepsilon_2)(p-1)\cdot \frac{1}{n^2}\cdot (1+\varepsilon_3)n \leq (1+\varepsilon_4)p\left(1 + \frac{p}{n}\right),$$

for a suitably chosen $\varepsilon_4 = \varepsilon_4(n) \to 0$. Similarly, $\operatorname{tr}(\hat{\Sigma}^2) \geq (1-\varepsilon_5)p(1+\frac{p}{n})$ for $\varepsilon_5 = \varepsilon_5(n) \to 0$. To complete the proof of Proposition 5.5, set $\varepsilon = \max\{\varepsilon_1, \varepsilon_4, \varepsilon_5\}$. □

REFERENCES

[1] ALON, N., KRIVELEVICH, M. and SUDAKOV, B. (1998). Finding a large hidden clique in a random graph. *Random Structures Algorithms* **13** 457–466. MR1662795

[2] AMES, B. P. W. and VAVASIS, S. A. (2011). Nuclear norm minimization for the planted clique and biclique problems. *Math. Program.* **129** 69–89. MR2831403

[3] AMINI, A. A. and WAINWRIGHT, M. J. (2009). High-dimensional analysis of semidefinite relaxations for sparse principal components. *Ann. Statist.* **37** 2877–2921. MR2541450

[4] ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York. MR0771294

[5] BAIK, J. and SILVERSTEIN, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680

[6] BERTHET, Q. and RIGOLLET, P. (2013). Optimal detection of sparse principal components in high dimension. *Ann. Statist.* **41** 1780–1815. MR3127849

[7] BERTHET, Q. and RIGOLLET, P. (2013). Complexity theoretic lower bounds for sparse principal component detection. In *COLT* 1046–1066. JMLR.org.

[8] BICKEL, P. J. and LEVINA, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227. MR2387969

[9] BIRNBAUM, A., JOHNSTONE, I. M., NADLER, B. and PAUL, D. (2013). Minimax bounds for sparse PCA with noisy high-dimensional data. *Ann. Statist.* **41** 1055–1084. MR3113803

[10] CAI, T. T., MA, Z. and WU, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist.* **41** 3074–3110. MR3161458

[11] D'ASPREMONT, A., BANERJEE, O. and EL GHAOUI, L. (2008). First-order methods for sparse covariance selection. *SIAM J. Matrix Anal. Appl.* **30** 56–66. MR2399568

[12] D'ASPREMONT, A., EL-GHAOUI, L., JORDAN, M. and LANCKRIET, G. (2004). A direct formulation for sparse PCA using semidefinite programming. *SIAM Rev.* **49** 434–448.

[13] DEBASHIS, P. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865

[14] DEKEL, Y., GUREL-GUREVICH, O. and PERES, Y. (2014). Finding hidden cliques in linear time with high probability. *Combin. Probab. Comput.* **23** 29–49. MR3197965

[15] DESHPANDE, Y. and MONTANARI, A. (2013). Finding hidden cliques of size $\sqrt{n/e}$ in nearly linear time. Available at arXiv:1304.7047.

[16] DESHPANDE, Y. and MONTANARI, A. (2014). Information-theoretically optimal sparse PCA. Available at arXiv:1402.2238.

[17] EL-KAROUI, N. (2003). On the largest eigenvalue of wishart matrices with identity covariance when n, p and $p/n \to \infty$. Available at arXiv:math/0309355.

[18] FEIGE, U. and KRAUTHGAMER, R. (2000). Finding and certifying a large hidden clique in a semirandom graph. *Random Structures Algorithms* **16** 195–208. MR1742351

[19] FEIGE, U. and RON, D. (2010). Finding hidden cliques in linear time. In *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms* (*AofA'10*) 189–203. Assoc. Discrete Math. Theor. Comput. Sci., Nancy. MR2735341

[20] JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961

[21] JOHNSTONE, I. M. and LU, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *J. Amer. Statist. Assoc.* **104** 682–693. MR2751448

[22] JOLLIFFE, I. T. (2002). *Principal Component Analysis*, 2nd ed. Springer, New York. MR2036084

[23] JOLLIFFE, I. T., TRENDAFILOV, N. T. and UDDIN, M. (2003). A modified principal component technique based on the LASSO. *J. Comput. Graph. Statist.* **12** 531–547. MR2002634

[24] LAURENT, B. and MASSART, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** 1302–1338. MR1805785

[25] LEDOIT, O. and WOLF, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. MR1926169

[26] LEI, J. and VU, V. Q. (2015). Sparsistency and agnostic inference in sparse PCA. *Ann. Statist.* **43** 299–322. MR3311861

[27] LU, Z. and ZHANG, Y. (2012). An augmented Lagrangian approach for sparse principal component analysis. *Math. Program.* **135** 149–193. MR2968253

[28] MA, Z. (2013). Sparse principal component analysis and iterative thresholding. *Ann. Statist.* **41** 772–801. MR3099121

[29] MOGHADDAM, B., WEISS, S. and AVIDAN, Y. (2006). Generalized spectral bounds for sparse LDA. In *Proceedings of the 23rd International Conference on Machine Learning* 641–648. ACM, New York.

[30] MOGHADDAM, B., WEISS, Y. and AVIDAN, S. (2006). Spectral bounds for sparse PCA: Exact and greedy algorithms. In *Advances in Neural Information Processing Systems* 915–922. MIT Press, Cambridge.

[31] MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York. MR0652932

[32] NADLER, B. (2008). Finite sample approximation results for principal component analysis: A matrix perturbation approach. *Ann. Statist.* **36** 2791–2817. MR2485013

[33] NATARAJAN, B. K. (1995). Sparse approximate solutions to linear systems. *SIAM J. Comput*. **24** 227–234. MR1320206

[34] SHEN, H. and HUANG, J. Z. (2008). Sparse principal component analysis via regularized low rank matrix approximation. *J. Multivariate Anal*. **99** 1015–1034. MR2419336

[35] STEWART, G. W. and SUN, J. G. (1990). *Matrix Perturbation Theory*. Academic Press, Boston, MA. MR1061154

[36] STURM, J. F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **11/12** 625–653. MR1778433

[37] VU, V. Q. and LEI, J. (2013). Minimax sparse principal subspace estimation in high dimensions. *Ann. Statist*. **41** 2905–2947. MR3161452

[38] WANG, T., BERTHET, Q. and SAMWORTH, R. (2014). Statistical and computational trade-offs in estimation of sparse principal components. Available at arXiv:1408.5369.

[39] WITTEN, D., TIBSHIRANI, R. and HASTIE, R. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* **10** 515–534.

[40] ZHANG, Z., ZHA, H. and SIMON, H. (2002). Low-rank approximations with sparse factors. I. Basic algorithms and error analysis. *SIAM J. Matrix Anal. Appl*. **23** 706–727 (electronic). MR1896815

[41] ZHAORAN WANG, Z., LU, H. and LIU, H. (2014). Nonconvex statistical optimization: Minimax-optimal sparse pca in polynomial time. Available at arXiv:1408.5352.

[42] ZOU, H., HASTIE, T. and TIBSHIRANI, R. (2006). Sparse principal component analysis. *J. Comput. Graph. Statist*. **15** 265–286. MR2252527

R. KRAUTHGAMER
B. NADLER
WEIZMANN INSTITUTE OF SCIENCE
REHOVOT
ISRAEL
E-MAIL: robert.krauthgamer@weizmann.ac.il
boaz.nadler@weizmann.ac.il

D. VILENCHIK
BEN-GURION UNIVERSITY
BEER-SHEVA
ISRAEL
E-MAIL: vilenchi@bgu.ac.il