Sparse PSD approximation of the PSD cone

Grigoriy Blekherman[1] Santanu S. Dey[1] Marco Molinaro [2]
Shengding Sun[1]

[1]Georgia Institute of Technology.

[2]Pontifical Catholic University of Rio de Janeiro.

Oct 2020

1 Introduction

Semi definite programming

$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A^i, X \rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \tag{SDP}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N \rangle := \sum_{i,j} M_{ij} N_{ij}$, and

Semi definite programming

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \tag{SDP}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N\rangle := \sum_{i,j} M_{ij} N_{ij}$, and

$$\mathcal{S}^n_+ = \{X \in \mathbb{R}^{n\times n} \mid X = X^T,\ u^\top X u \geq 0,\ \forall u \in \mathbb{R}^n\}.$$

Semi definite programming

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N\rangle := \sum_{i,j} M_{ij} N_{ij}$, and

$$\mathcal{S}^n_+ = \{X \in \mathbb{R}^{n\times n} \,|\, X = X^T,\ u^\top X u \geq 0,\ \forall u \in \mathbb{R}^n\}.$$

- Polynomial-time algorithm— but often challenging to solve in practice.

A relaxation: Sparse SDP

$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A^i, X \rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

A relaxation: Sparse SDP

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A^i, X \rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{aligned} \qquad \text{(SDP)}$$

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A^i, X \rangle \leq b_i \ \forall i \in \{1, \ldots, m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{aligned} \qquad \text{(Sparse SDP)}$$

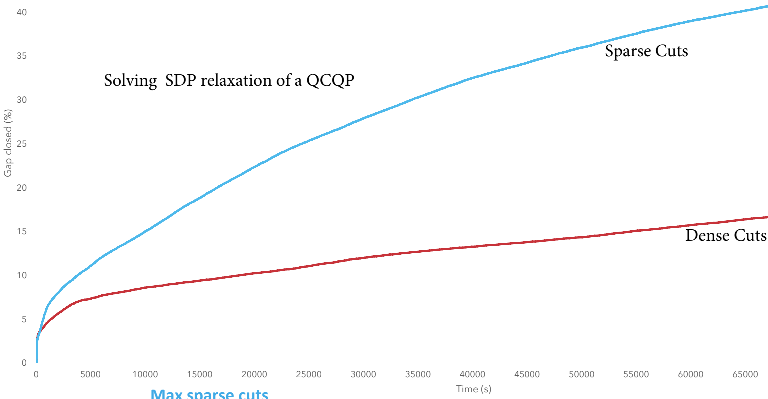
A relaxation: Sparse SDP

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \; \forall i \in \{1, \ldots, m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

- We can enforce PSD constraints by iteratively separating linear constraints. Enforcing PSD-ness on smaller $k \times k$ principal submatrix leads to linear constraints that are sparser, an important property leveraged by linear programming solvers that greatly improve their efficiency.

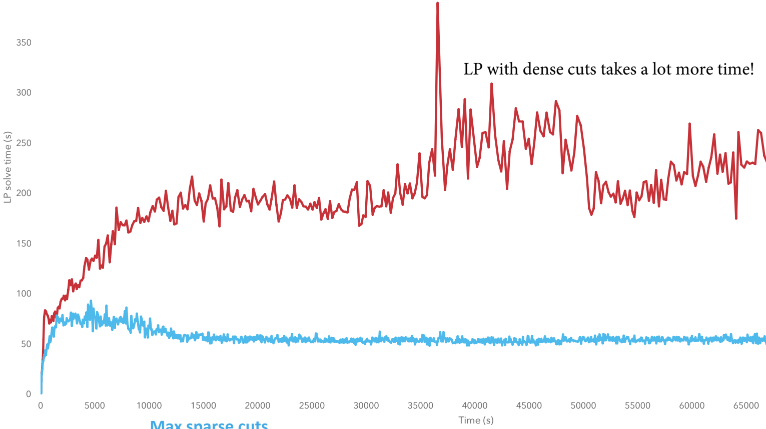
Example from [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]



Sparsity level
$k=0.25(n+1)$

Max sparse cuts per iteration
$K=5n$

Example from [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]



Sparsity level
$k = 0.25(n+1)$

Max sparse cuts per iteration
$K = 5n$

Sparse SDP

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \; \forall i \in \{1, \dots, m\} \qquad \text{(Sparse SDP)} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}_+^k. \end{array}$$

- [A. Qualizza, P. Belotti, and F. Margot (2012)]
- [R. Baltean-Lugojan, P. Bonami, R. Misener, and A. Tramontani (2018)]
- [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]

Sparse SDP

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \ \forall i \in \{1, \dots, m\} \qquad \text{(Sparse SDP)} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array}$$

- [A. Qualizza, P. Belotti, and F. Margot (2012)]
- [R. Baltean-Lugojan, P. Bonami, R. Misener, and A. Tramontani (2018)]
- [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]
- [S. Sojoudi and J. Lavaei (2014)]
- [B. Kocuk, SSD, and X. A. Sun (2016)]

Sparse SDP

$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A^i, X \rangle \leq b_i \ \forall i \in \{1, \ldots, m\} \qquad \text{(Sparse SDP)} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array}$$

- [A. Qualizza, P. Belotti, and F. Margot (2012)]
- [R. Baltean-Lugojan, P. Bonami, R. Misener, and A. Tramontani (2018)]
- [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]
- [S. Sojoudi and J. Lavaei (2014)]
- [B. Kocuk, SSD, and X. A. Sun (2016)]
- [E. G. Boman, D. Chen, O. Parekh, and S. Toledo (2005)]
- [A. A. Ahmadi and A. Majumdar (2019)])

Our question

$$z^{\text{SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \; \forall i \in \{1,\dots,m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

Our question

$$z^{\text{SDP}} := \min \ \langle C, X\rangle \quad \text{s.t.} \ \langle A^i, X\rangle \leq b_i \ \ \forall i \in \{1,\dots,m\}, \ X \in \mathcal{S}^n_+, \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \min \ \langle C, X\rangle \quad \text{s.t.} \ \langle A^i, X\rangle \leq b_i \ \forall i \in \{1,\dots,m\}, \ \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \qquad \text{(Sparse SDP)}$$

Relationship between z^{SDP} and $z^{\text{Sparse-SDP}}$?

Our question

$$z^{\text{SDP}} := \min \ \langle C, X \rangle \quad \text{s.t.} \ \langle A^i, X \rangle \le b_i \ \ \forall i \in \{1, \dots, m\}, \ X \in \mathcal{S}^n_+, \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \min \ \langle C, X \rangle \quad \text{s.t.} \ \langle A^i, X \rangle \le b_i \ \forall i \in \{1, \dots, m\}, \ \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \qquad \text{(Sparse SDP)}$$

Relationship between z^{SDP} and $z^{\text{Sparse-SDP}}$?

- Seems like a difficult question to analyze.

Easier question

$$z^{\text{SDP}} := \min \langle C, X\rangle \quad \text{s.t.} \quad \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\}, \quad X \in \mathcal{S}^n_+, \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \min \langle C, X\rangle \quad \text{s.t.} \quad \langle A^i, X\rangle \leq b_i \ \forall i \in \{1,\dots,m\} \qquad \text{(Sparse SDP)}$$

selected $k \times k$ principal submatrices of $X \in \mathcal{S}^k_+$.

Easier question

$$z^{\text{SDP}} := \min \langle C, X\rangle \quad \text{s.t.} \quad \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\}, \quad X \in \mathcal{S}^n_+, \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \min \langle C, X\rangle \quad \text{s.t.} \quad \langle A^i, X\rangle \leq b_i \ \forall i \in \{1,\dots,m\} \qquad \text{(Sparse SDP)}$$

~~selected~~ $k \times k$ principal submatrices of $X \in \mathcal{S}^k_+$.

How far is cone with all $k \times k$ submatrices PSD from $\mathcal{S}^n_+$?

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- How far is $\mathcal{S}^{n,k}$ from $\mathcal{S}^{n}_{+}$?

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- How far is $\mathcal{S}^{n,k}$ from $\mathcal{S}^{n}_{+}$?
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- How far is $\mathcal{S}^{n,k}$ from $\mathcal{S}^{n}_{+}$?
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:
  1. The norm to measure this distance and
  2. A normalization method

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- How far is $\mathcal{S}^{n,k}$ from $\mathcal{S}^n_+$?
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:
  1. The norm to measure this distance and
  2. A normalization method

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) = \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \text{dist}_F(M, \mathcal{S}^n_+)$$

$$= \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \inf_{N \in \mathcal{S}^n_+} \|M - N\|_F.$$

2

Main results

2.1

Upper bounds on $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$

Upper bound 1

Theorem (Upper Bound 1)

For all $2 \le k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \le \frac{n-k}{n+k-2}. \quad (1)$$

Upper bound 1

Theorem (Upper Bound 1)

For all $2 \leq k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}. \qquad (1)$$

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$.

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Theorem (Upper bound 2)

Assume $n \geq 97$ and $k \geq \frac{3n}{4}$. Then

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq 96 \left(\frac{n-k}{n}\right)^{3/2}. \quad (2)$$

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Theorem (Upper bound 2)

Assume $n \geq 97$ and $k \geq \frac{3n}{4}$. Then

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq 96 \left(\frac{n-k}{n}\right)^{3/2}. \quad (2)$$

- This bound dominates the previous bound when $\frac{k}{n}$ is sufficiently large.

2.2

Lower bounds on $\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$

Lower bound 1

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}. \quad (3)$$

Lower bound 1

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \tag{3}$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{n-k}{n}$$

So first upper bound (Thm 1) is tight (upto constant).

Lower bound 1

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \tag{3}$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{n-k}{n}$$

  So first upper bound (Thm 1) is tight (upto constant).

- When k is very large: $n - k = c$ where c is very small

Lower bound 1

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \quad (3)$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{n-k}{n}$$

So first upper bound (Thm 1) is tight (upto constant).

- When k is very large: $n - k = c$ where c is very small

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{c}{n^{3/2}}$$

So second upper bound (Thm 2) is tight (upto constant).

Lower bound 2: What happens when $k = rn$?

- Upper bound: $\frac{n-k}{n} = 1 - r$, a constant independent of n
- Lower bound 1: $\approx (1/r - 1)\frac{1}{n^{1/2}}$.

So is upper bound weak in this regime?

Lower bound 2: What happens when $k = rn$?

- Upper bound: $\frac{n-k}{n} = 1 - r$, a constant independent of n
- Lower bound 1: $\approx (1/r - 1)\frac{1}{n^{1/2}}$.

So is upper bound weak in this regime?

Theorem (Lower bound 2)

Fix a constant $r < \frac{1}{93}$ and $k = rn$. Then for all $k \geq 2$,

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) > \frac{\sqrt{r - 93r^2}}{\sqrt{162r + 3}},$$

which is independent of n.

2.3
Do we need $\binom{n}{k}$ PSD constraints?

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem

Let $2 \leq k \leq n-1$. *Consider* $\varepsilon, \delta > 0$ *and let*

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \ldots, I_m)$ *be a sequence of random* k*-sets independently uniformly sampled from* $\binom{[n]}{k}$,

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem

Let $2 \leq k \leq n-1$. Consider $\varepsilon, \delta > 0$ and let

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \ldots, I_m)$ be a sequence of random k-sets independently uniformly sampled from $\binom{[n]}{k}$, and define $\mathcal{S}_{\mathcal{I}}$ as the set of matrices satisfying the PSD constraints for the principal submatrices indexed by the I_i's, namely

$$\mathcal{S}_{\mathcal{I}} := \{M \in \mathbb{R}^{n\times n} : M_{I_i} \succeq 0,\ \forall i \in [m]\}.$$

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem

Let $2 \leq k \leq n-1$*. Consider* $\varepsilon, \delta > 0$ *and let*

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \dots, I_m)$ *be a sequence of random* k*-sets independently uniformly sampled from* $\binom{[n]}{k}$*, and define* $\mathcal{S}_{\mathcal{I}}$ *as the set of matrices satisfying the PSD constraints for the principal submatrices indexed by the* I_i*'s, namely*

$$\mathcal{S}_{\mathcal{I}} := \{M \in \mathbb{R}^{n \times n} : M_{I_i} \succeq 0,\ \forall i \in [m]\}.$$

Then with probability at least $1-\delta$ *we have*

$$\overline{\text{dist}}_F(\mathcal{S}_{\mathcal{I}}, \mathcal{S}^n_+) \leq (1+\varepsilon)\frac{n-k}{n+k-2}.$$

3

Proof sketch

3.1

Proof of:

Theorem (Upper Bound 1)

For all $2 \leq k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}.$$

Proof of Upper bound 1

- If

$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$

then red-submatrix is $k \times k$ PSD matrix.

Proof of Upper bound 1

- If

$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$

then red-submatrix is $k \times k$ PSD matrix.

- So

$$\begin{bmatrix} * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathcal{S}^n_+.$$

Proof of Upper bound 1

- If

$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$

then red-submatrix is $k \times k$ PSD matrix.

- So

$$\begin{bmatrix} * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathcal{S}^n_+.$$

- Take average of all the above matrices for different principal $k \times k$ submatrices (and suitably scale with a positive number), then the resulting matrix is in S^n_+.
- The distance between this average PSD matrix and X gives bound.

3.2

Proof of:

Theorem (Upper bound 2)

Assume $n \geq 97$ and $k \geq \frac{3n}{4}$. Then

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq 96 \left(\frac{n-k}{n}\right)^{3/2}.$$

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n-k$ negative eigenvalues.

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n-k$ negative eigenvalues.
- Since the PSD cone consists of symmetric matrices with non-negative eigenvalues, the distance from a unit-norm matrix $M \in \mathcal{S}^{n,k}$ to $\mathcal{S}^n_+$ is upper bounded by

  (absolute value of most negative eigenvalue of M)$\times\sqrt{n-k}$.

- So we need to upper bound absolute value of most negative eigenvalue of M for $M \in \mathcal{S}^{n,k}$ and $\|M\|_F = 1$.

Proof of upper bound 2 -contd.

- Let
$$M = -\lambda v v^\top + \sum_{i=2}^{n} \mu_i v_i v_i^\top$$
where $\lambda > 0$.
- Proof uses probabilitic method: Randomly sparsify (with some scaling) v and let the resulting random vector be V.

Proof of upper bound 2 -contd.

- Let
$$M = -\lambda v v^\top + \sum_{i=2}^{n} \mu_i v_i v_i^\top$$
where $\lambda > 0$.
- Proof uses probabilitic method: Randomly sparsify (with some scaling) v and let the resulting random vector be V. Think of V having the following properties:
  - $V \approx v$, i.e. $V^\top v \approx 1$ and $V^\top v_i \approx 0$. (♣)
  - V has a support of k (♠)

Proof of upper bound 2 -contd.

- Let
$$M = -\lambda v v^\top + \sum_{i=2}^{n} \mu_i v_i v_i^\top$$
where $\lambda > 0$.
- Proof uses probabilitic method: Randomly sparsify (with some scaling) v and let the resulting random vector be V. Think of V having the following properties:
  - $V \approx v$, i.e. $V^\top v \approx 1$ and $V^\top v_i \approx 0$. (♣)
  - V has a support of k (♠)
- So (♣) implies:

$$V^\top M V \approx -\lambda \cdot 1 + \sum_{i=2}^{n} \mu_i 0 \approx -\lambda + \text{small error} \qquad (A)$$

Proof of upper bound 2 -contd.

- Let
$$M = -\lambda v v^\top + \sum_{i=2}^{n} \mu_i v_i v_i^\top$$
where $\lambda > 0$.
- Proof uses probabilitic method: Randomly sparsify (with some scaling) v and let the resulting random vector be V. Think of V having the following properties:
  - $V \approx v$, i.e. $V^\top v \approx 1$ and $V^\top v_i \approx 0$. (♣)
  - V has a support of k (♠)
- So (♣) implies:
$$V^\top M V \approx -\lambda \cdot 1 + \sum_{i=2}^{n} \mu_i 0 \approx -\lambda + \text{small error} \qquad (A)$$
- On the other hand (♠) implies:
$$V^\top M V \geq 0, \qquad (B)$$
since V has a support of k.

Proof of upper bound 2 -contd.

- Let
$$M = -\lambda v v^\top + \sum_{i=2}^{n} \mu_i v_i v_i^\top$$
where $\lambda > 0$.
- Proof uses probabilitic method: Randomly sparsify (with some scaling) v and let the resulting random vector be V. Think of V having the following properties:
  - $V \approx v$, i.e. $V^\top v \approx 1$ and $V^\top v_i \approx 0$. (♣)
  - V has a support of k (♠)
- So (♣) implies:
$$V^\top MV \approx -\lambda \cdot 1 + \sum_{i=2}^{n} \mu_i 0 \approx -\lambda + \text{small error} \qquad (A)$$
- On the other hand (♠) implies:
$$V^\top MV \geq 0, \qquad (B)$$
since V has a support of k.
- So (A) and (B) imply:
$$\boxed{-\lambda + \text{small error} \geq 0 \Rightarrow \lambda \leq \text{small error}}.$$

3.3

Proof of:

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}.$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u =$$

Proof of lower bound 1

- Consider the matrix:

$$G(a, b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2 \geq (a+b) - ak$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^n u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2 \geq (a+b) - ak$$

- So $G(a,b) \in \mathcal{S}^{n,k}$ iff $(1-k)a + b \geq 0$.
- Use these explicit matrices to obtain lower bound from S^n_+

3.4

Proof of:

Theorem (Lower bound 2)

Fix a constant $r < \frac{1}{93}$ and $k = rn$. Then for all $k \geq 2$,

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) > \frac{\sqrt{r - 93r^2}}{\sqrt{162r + 3}},$$

which is independent of n.

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \ldots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \ldots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \ldots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \ldots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

- This normalization makes $\|M\|_F \approx 1$.
- $\text{dist}_F(M, \mathcal{S}^n_+) \gtrsim \sqrt{\left(\frac{1}{\sqrt{n}}\right)^2 \cdot \frac{n}{2}} = cst$ independent of n.

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \ldots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \ldots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

- This normalization makes $\|M\|_F \approx 1$.
- $\text{dist}_F(M, \mathcal{S}^n_+) \gtrsim \sqrt{\left(\frac{1}{\sqrt{n}}\right)^2 \cdot \frac{n}{2}} = cst$ independent of n.
- So we only need to guarantee that M belongs to the k-PSD closure.

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top M x$:

$$x^\top M x = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)
- So *if we could construct the matrix W so that for all k-sparse vectors* $x \in \mathbb{R}^n$ *we had* $\boxed{\|Wx\|_2^2 \approx \|x\|_2^2}$:

$$x^\top M x \gtrsim -\frac{1}{\sqrt{n}} \|x\|_2^2 + \frac{1}{\sqrt{n}} \|x\|_2^2 \gtrsim 0 \qquad \boxed{\text{for all } k\text{-sparse vectors } x},$$

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)
- So if we *could construct the matrix W so that for all k-sparse vectors* $x \in \mathbb{R}^n$ *we had* $\boxed{\|Wx\|_2^2 \approx \|x\|_2^2}$:

$$x^\top Mx \gtrsim -\frac{1}{\sqrt{n}} \|x\|_2^2 + \frac{1}{\sqrt{n}} \|x\|_2^2 \gtrsim 0 \qquad \boxed{\text{for all } k\text{-sparse vectors } x},$$

- This approximate preservation of norms of sparse vectors is precisely the notion of the *Restricted Isometry Property*.

Thank You.