MATHEMATICS OF OPERATIONS RESEARCH
Vol. 33, No. 4, November 2008, pp. 965–975
ISSN 0364-765X | EISSN 1526-5471 | 08 | 3304 | 0965
DOI 10.1287/moor.1080.0331

On the Low Rank Solutions for Linear Matrix Inequalities

Wenbao Ai
School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China,
wenbaoai@gmail.com

Yongwei Huang, Shuzhong Zhang
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong
{ywhuang@se.cuhk.edu.hk, zhang@se.cuhk.edu.hk, http://www.se.cuhk.edu.hk/~zhang/}

In this paper we present a polynomial-time procedure to find a low-rank solution for a system of linear matrix inequalities (LMI). The existence of such a low-rank solution was shown in the work of Au-Yeung and Poon and the work of Barvinok. In the approach of Au-Yeung and Poon an earlier unpublished manuscript of Bohnenblust played an essential role. Both proofs in the work of Au-Yeung and Poon and that of Barvinok are nonconstructive in nature. The aim of this paper is to provide a polynomial-time constructive procedure to find such a low-rank solution approximatively. Extensions of our new results and their relations to some of the known results in the literature are discussed.

Key words: rank reduction; linear matrix inequality; joint numerical range
MSC2000 subject classification: Primary: 90C22, 15A03; secondary: 15A60, 15A48
OR/MS subject classification: Primary: mathematics, matrices; secondary: programming, nonlinear
History: Received October 1, 2006; revised October 11, 2007, and March 28, 2008. Published online in *Articles in Advance* November 3, 2008.

1. Introduction. Finding a low-rank matrix solution for a system of linear matrix inequalities (LMI), or a semidefinite program (SDP), is of great importance in theory as well as in practice. As one example, SDP is often used as a relaxation for quadratic optimization, where a rank-1 SDP solution is automatically optimal for the quadratic model. Thus, finding a rank-1 matrix solution for an SDP problem is equivalent to solving a quadratic program. Ye and Zhang [19] showed a number of examples where such rank-1 solutions can be found. In sensor network localization problems (see Biswas and Ye [4]), one is required to find rank 2 (planar) or rank 3 (spatial) solutions for an SDP relaxation problem. Similarly, in graph realization (see Biswas et al. [5]), the problem is to find a solution to an LMI with the rank no more than a given value. Unexpectedly, the famous kissing number problem of identical spheres (originally due to Newton and Gregory) is related as well. More information on the ball kissing number problem can be found, e.g., in Pfender and Ziegler [14].

Barvinok [2] presented an upper bound on the lowest rank among all matrix solutions of a feasible LMI system. The same bound was independently rediscovered by Pataki in [13]. Later it was shown by Barvinok in [3] that the bound is essentially tight. Along a different line, Sturm and Zhang [18] proposed a matrix rank-1 decomposition scheme, leading to an algorithmic approach to finding rank-1 solutions for SDP, provided that the number of constraints in the SDP problem is small. The matrix decomposition scheme of Sturm and Zhang was extended to the complex matrix case by Huang and Zhang [9]. Moreover, Huang and Zhang [9] also proposed a polynomial-time procedure, in the complex case, for finding a low-rank solution with a bound on the rank similar to that in Barvinok [2] and Pataki [13]. In the case where the number of constraints in the SDP is small, the rank-1 decomposition method of Sturm and Zhang [18] (and Huang and Zhang [9] for the complex SDP) can be considered as an alternative polynomial-time procedure to find the low-rank matrix solutions (in this case rank-1), other than the procedure suggested in Barvinok [2] and Pataki [13].

Although the bound in Barvinok [2] cannot be improved in general, it does permit an unexpected strengthening under an additional mild condition. A first reference in which such an enhanced bound had appeared was Au-Yeung and Poon [1]. Essentially, Au-Yeung and Poon [1] showed that there is a chance to further reduce the rank by one. For instance, for a bounded standard real SDP with three constraints, Barvinok [2] and Pataki [13]'s bound can only conclude that there exists a rank-2 feasible solution. However, Au-Yeung and Poon [1]'s result concludes that there is in fact a rank-1 feasible solution! Although this further rank reduction (by 1) may appear insignificant when there is a large number of constraints, it can be interesting when the number of constraints is small (see more examples in §3). The cornerstone in Au-Yeung and Poon's proof is a result established by Bohnenblust [6] in an unpublished note. Barvinok [3] gave an alternative proof. Unfortunately, both proofs are not constructive. As a matter of fact, Barvinok posed as an open problem in Barvinok [3] to find such a low-rank solution constructively and efficiently. The main goal of the current paper is indeed to solve this open problem by presenting a polynomial-time method to actually get hold of a low-rank solution as promised in Au-Yeung and Poon [1] and Barvinok [3], albeit in an approximative sense. That we speak only of approximation in this context is inevitable, because finding *any* feasible solution (even without rank constraints) for an LMI system itself is basically a semidefinite program, which can only be solved approximately. The aim of this paper is to present a polynomial-time algorithm to find the low-rank solutions in terms of the problem dimension and $\log(1/\epsilon)$, where ϵ is the error in satisfying the equality constraints. As a consequence, the existence of the exact low-rank solution follows by taking a limit. The organization of the paper is as follows. In §2 we present an algorithmic approach and the new results, and in §3 we discuss extensions of our results and their connections with other known results in the literature.

2. A polynomial-time rank reduction procedure. Let $\mathcal{F}$ be either $\Re$ (real) or $\mathbf{C}$ (complex), and $\mathcal{SF}^n$ be either $\mathcal{S}^n$ (space of real n by n symmetric matrices) or $\mathcal{H}^n$ (space of complex n by n Hermitian conjugate symmetric matrices). Furthermore, let

$$d_{\mathcal{F}}(n) = \begin{cases} n(n+1)/2, & \text{if } \mathcal{F} = \Re; \\ n^2, & \text{if } \mathcal{F} = \mathbf{C}. \end{cases}$$

In other words, $d_{\mathcal{F}}(n) = \dim(\mathcal{SF}^n)$. We also use X^{H} to denote the conjugate transpose of X in the Hermitian complex case, and as a special case the simple transpose of X in the real case. We denote $\|X\|_F$ to be the Frobenius norm of matrix X.

Let $\mathcal{A} \subseteq \mathcal{SF}^n$ be an affine subspace. The following result was shown by Barvinok [2] and Pataki [13] for the real case, and Huang and Zhang [9] for the complex case.

Theorem 2.1. *Suppose that $\mathcal{A}$ is an affine subspace of $\mathcal{SF}^n$, which does not contain the origin, and that $\mathcal{A} \cap \mathcal{SF}^n_+ \neq \varnothing$. If $\operatorname{codim}(\mathcal{A}) \le d_{\mathcal{F}}(p+1) - 1$ with $p \ge 1$, then one can find in polynomial time $X \in \mathcal{A} \cap \mathcal{SF}^n_+$ with $\operatorname{rank}(X) \le p$.*

In a similar vein, in case $\mathcal{A}$ is a subspace, then the above theorem has a parallel statement as follows.

Theorem 2.2. *Let $\mathcal{L}$ be a subspace of $\mathcal{SF}^n$. Suppose that $\dim(\mathcal{L} \cap \mathcal{SF}^n_+) \ge 1$, i.e., $\mathcal{L}$ has a nontrivial intersection with $\mathcal{SF}^n_+$. If $\operatorname{codim}(\mathcal{L}) \le d_{\mathcal{F}}(p+1) - 2$ with $p \ge 1$, then one can find in polynomial time $0 \neq X \in \mathcal{L} \cap \mathcal{SF}^n_+$, such that $\operatorname{rank}(X) \le p$.*

Proof. Let us denote $m = \operatorname{codim}(\mathcal{L})$, and specifically write

$$\mathcal{L} = \{X \in \mathcal{SF}^n \mid A_i \cdot X = 0,\ i = 1, \ldots, m\},$$

where $A_i \in \mathcal{SF}^n$, $b_i \in \Re$, $i = 1, \ldots, m$, and A_is are linearly independent.

Find X such that $0 \neq X \in \mathcal{L} \cap \mathcal{SF}^n_+$, say by semidefinite programming in polynomial time. Let $r = \operatorname{rank}(X)$, and decompose $X = UU^{\mathrm{H}}$, where $U \in \mathcal{F}^{n \times r}$ has orthogonal column vectors and $\operatorname{rank}(U) = r$. Consider the system of linear equations

$$(U^{\mathrm{H}} A_i U) \cdot \Delta = 0, \quad i = 1, \ldots, m, \tag{1}$$

where $\Delta \in \mathcal{SF}^r$. The above system of linear equations (1) has two linearly independent solutions when $d_{\mathcal{F}}(r) > m + 1$, which can be found in polynomial time by Gaussian elimination. One of these two solutions, simply denoted by Δ ($\neq 0$), can be chosen to be linearly independent of I_r. Therefore, if $d_{\mathcal{F}}(r) > m + 1$, then one is able to further reduce the rank of X by letting

$$X := U(I_r + \hat{t}\Delta)U^{\mathrm{H}}$$

with $\hat{t} = \arg\max\{t > 0 \mid I_r + t\Delta \succeq 0\}$. (If $\Delta \succeq 0$, then switch to $\Delta := -\Delta$.) This rank-reduction procedure terminates only when (1) does not have any solution that is linearly independent of I_r, implying that $d_{\mathcal{F}}(r) \le m + 1$. Therefore, if $m \le d_{\mathcal{F}}(p+1) - 2$ (under the condition of Theorem 2.2), then $d_{\mathcal{F}}(r) \le m + 1 \le d_{\mathcal{F}}(p+1) - 1$, and consequently $\operatorname{rank}(X) = r < p + 1$ due to the strict monotonicity of $d_{\mathcal{F}}(\cdot)$. Thus, $\operatorname{rank}(X) \le p$. This proves Theorem 2.2. □

In general, the above bound on $\operatorname{codim}(\mathcal{L})$ is tight; see example below.

Example 2.1. Consider $p = n - 1$, and $\mathcal{L} = \{\lambda I_n \mid \lambda \in \Re\}$. Clearly, $\operatorname{codim}(\mathcal{L}) = d_{\mathcal{F}}(p+1) - 1$ in this case. There is no $0 \neq X \in \mathcal{L} \cap \mathcal{SF}^n_+$ with $\operatorname{rank}(X) \le n - 1$. This shows that the condition "$\operatorname{codim}(\mathcal{L}) \le d_{\mathcal{F}}(p+1) - 2$" in Theorem 2.2 cannot be relaxed in general.

It turns out, however, that if $p \le n - 2$, then the bound on $\operatorname{codim}(\mathcal{L})$ can be further improved, as the following theorem shows. This theorem was originally due to Bohnenblust [6] in an unpublished note written in 1948. The contribution of the current paper is to present a constructive procedure that runs in polynomial time to actually find such a low-rank matrix solution.

THEOREM 2.3. *Let $\mathcal{L}$ be a subspace of $\mathcal{SF}^n$. Suppose that* $\dim(\mathcal{L}\cap\mathcal{SF}^n_+)\geq 1$. *If* $\mathrm{codim}(\mathcal{L})\leq d_{\mathcal{F}}(p+1)-1$ *and* $1\leq p\leq n-2$, *then in polynomial time (in terms of n and* $\log(1/\epsilon)$) *one can find* $X\succeq 0$, *with* $\|X\|_F=1$, $\mathrm{rank}(X)\leq p$, *and* $\mathrm{dist}(X,\mathcal{L})\leq\epsilon$. *By taking any cluster point as* $\epsilon\to 0$, *we conclude that there is a nontrivial* $X\in\mathcal{L}\cap\mathcal{SF}^n_+$, *such that* $\mathrm{rank}(X)\leq p$.

Before we present our construction, let us first comment on the result and discuss the main ideas of the new proof here. As Barvinok remarked in Barvinok [3], Theorem 2.1 and Theorem 2.3 (or as an exact counterpart, Theorem 2.4) are quite different in nature. Theorem 2.1 gives a bound on the rank that is satisfied by *any extremal solution* of the LMI system, whereas Theorem 2.3 (or Theorem 2.4) only assures the existence of *one* extremal solution of the LMI system with an even lower rank. This lower-rank solution is in some way an analog of a *degenerate vertex* in the context of a polyhedron. As is well known, degeneracy occurs in a polyhedron like an accident, and it will almost surely not happen for a generally positioned polyhedron. Interestingly, Theorems 2.3 and 2.4 suggest that this rank degeneracy is inevitable for a general LMI system. To look for this particular "degenerate matrix solution," we introduce a single parameter in the range space of the matrix solution. It is possible to confine this parameter to a finite interval, say $[0,1]$. Moreover, it is easy to find the corresponding matrix solutions with respect to the parameters at the ends of the interval, initially being 0 and 1. One of the matrix solutions is known to be positive definite with rank $p+1$, and the other has a rank no more than $p+1$ and is indefinite. Now, we can apply bisection to reduce the parameter search interval, retaining the property that one end yields a matrix in $\mathcal{L}\cap\mathcal{SF}^n_+$ with rank $p+1$, and the other end yields a matrix that is indefinite. It is also possible to combine the positive definite matrix with the indefinite matrix (after modifying its range space to fit that of the positive definite one), hence further reducing the rank of the positive definite matrix. However, this will lead to a matrix solution that does not precisely reside in $\mathcal{L}$, because of the nonlinear modification made on the indefinite matrix. The main effort is then to show that the distance to $\mathcal{L}$, nonetheless, converges to zero as the bisection reduces the size of the parameter search interval to zero.

PROOF. We need only to continue when the construction in the proof for Theorem 2.2 has finished. Under the conditions of Theorem 2.3 ($m=\mathrm{codim}(\mathcal{L})\leq d_{\mathcal{F}}(p+1)-1$ and $1\leq p\leq n-2$), to complete the proof we shall continue to discuss what to do when the above rank-reduction procedure terminates, and yet $\mathrm{rank}(X)=p+1$. Specifically, in this remaining case we have exactly $m=d_{\mathcal{F}}(p+1)-1$, and $\mathrm{rank}(X)=p+1\leq n-1$. In that case, let us first do a spectral decomposition on X to yield

$$X=U\Lambda U^{\mathrm{H}},\tag{2}$$

where $U=[u_1,u_2,\ldots,u_p,u_{p+1}]\in\mathcal{F}^{n\times(p+1)}$ and u_js are normalized orthogonal vectors with $j=1,2,\ldots,p+1$, and $\Lambda\in\mathcal{SF}^{p+1}_{++}$ is a diagonal positive definite matrix. Because the rank-reduction procedure has terminated at the solution X, we conclude that

$$(U^{\mathrm{H}}A_iU)\cdot\Delta=0,\quad i=1,\ldots,m;\ \ \Delta\in\mathcal{SF}^{p+1}\tag{3}$$

has the solutions that form a one-dimensional subspace, that is, all the solutions of (3) are multiples of Λ.

Because $p+1\leq n-1$, there must exist $u_{p+2}\in\mathcal{F}^n$ with $\|u_{p+2}\|=1$ and $u_j\cdot u_{p+2}=0$, $j=1,2,\ldots,p,p+1$. Consider now the system of linear equations

$$(U^{\mathrm{H}}A_iU)\cdot\Delta=r_i,\quad i=1,\ldots,m;\ \ \Delta\in\mathcal{SF}^{p+1},\tag{4}$$

where $r\in\Re^m$ is an arbitrary given real vector. There are m ($=d_{\mathcal{F}}(p+1)-1$) equations and $m+1$ ($=d_{\mathcal{F}}(p+1)$) variables in (4). Because (3) defines a one-dimensional null space, it follows that the rank of linear system (4) is precisely m. In other words, the system of linear equations (4) always has a solution for any real vector $r\in\Re^m$. In particular, there exists $\Delta\in\mathcal{SF}^{p+1}$ satisfying

$$(U^{\mathrm{H}}A_iU)\cdot\Delta=-u_{p+2}^{\mathrm{H}}A_iu_{p+2},\quad i=1,\ldots,m.\tag{5}$$

In the remainder, we assume that $(u_{p+2}^{\mathrm{H}}A_1u_{p+2},u_{p+2}^{\mathrm{H}}A_2u_{p+2},\ldots,u_{p+2}^{\mathrm{H}}A_mu_{p+2})\neq(0,0,\ldots,0)$, because otherwise $u_{p+2}u_{p+2}^{\mathrm{H}}$ would be exactly a rank-1 solution of (3), and the proof is completed in that case. Combining (5) and (3), it follows immediately that

$$A_i\cdot U(\Delta+\tau\Lambda)U^{\mathrm{H}}=-u_{p+2}^{\mathrm{H}}A_iu_{p+2},\quad i=1,\ldots,m,\tag{6}$$

for any given $\tau \in \Re$. We specifically choose $\hat{\tau} := -\lambda_{\max}(\Lambda^{-1/2}\Delta\Lambda^{-1/2})$. Thus, $\Delta + \hat{\tau}\Lambda \preceq 0$, and $\operatorname{rank}(\Delta + \hat{\tau}\Lambda) \le p$. Let the spectral decomposition of $\Delta + \hat{\tau}\Lambda$ be

$$\Delta + \hat{\tau}\Lambda = Q\begin{bmatrix} \beta_1 & & & \\ & \ddots & & \\ & & \beta_p & \\ & & & 0 \end{bmatrix} Q^{\mathrm{H}}$$

where $\beta_j \le 0$, $j = 1, \ldots, p$, and Q is an order $p+1$ unitary matrix. Note that β_js are not all zero, because $(u_{p+2}^{\mathrm{H}} A_1 u_{p+2}, u_{p+2}^{\mathrm{H}} A_2 u_{p+2}, \ldots, u_{p+2}^{\mathrm{H}} A_m u_{p+2}) \ne (0, 0, \ldots, 0)$. To simplify, denote

$$W := UQ, \qquad \text{and} \qquad [w_1, \ldots, w_{p+1}] := W \qquad \text{and} \qquad w_{p+2} := u_{p+2}. \tag{7}$$

Clearly, the vectors w_js, $j = 1, \ldots, p+1, p+2$, remain unit vectors and orthogonal to each other. Therefore, Equation (3) can be written as

$$([w_1, \ldots, w_p, w_{p+1}]^{\mathrm{H}} A_i [w_1, \ldots, w_p, w_{p+1}]) \cdot (Q^{\mathrm{H}} \Lambda Q) = 0, \quad i = 1, \ldots, m.$$

Similarly, by ignoring the zero term in front of $w_{p+1} w_{p+1}^{\mathrm{H}}$, Equation (6) can be rearranged and its solution can be written as

$$([w_1, \ldots, w_p, w_{p+2}]^{\mathrm{H}} A_i [w_1, \ldots, w_p, w_{p+2}]) \cdot \begin{bmatrix} \beta_1 & & & \\ & \ddots & & \\ & & \beta_p & \\ & & & 1 \end{bmatrix} = 0, \quad i = 1, \ldots, m.$$

Let

$$W_t := [w_1, w_2, \ldots, w_p, (1-t)w_{p+1} + t w_{p+2}] \in \mathscr{F}^{n \times (p+1)}, \tag{8}$$

which is always a tall matrix; i.e., $\operatorname{rank}(W_t) = p+1$ for any $t \in [0, 1]$. For a given $t \in [0, 1]$, consider nontrivial unit solutions to the following system of linear equations

$$(W_t^{\mathrm{H}} A_i W_t) \cdot \Delta = 0, \quad i = 1, \ldots, m. \tag{9}$$

Because $m = d_{\mathscr{F}}(p+1) - 1$, for any fixed $t \in [0, 1]$ the above linear equations must always have nontrivial solutions, which can be found in polynomial time by Gaussian elimination. Take a nontrivial matrix solution for (9), normalize the solution, and denote it to be Δ_t ($\|\Delta_t\|_F = 1$). Before discussing the properties of the solutions of (9) for any given $t \in [0, 1]$, let us check what we have got so far at the two special points, $t = 0$ and $t = 1$. In fact, we have identified two solutions corresponding to $t = 0$ and $t = 1$, respectively, namely

$$\Delta_0 = Q^{\mathrm{H}} \Lambda Q / \|\Lambda\|_F \qquad \text{and} \qquad \Delta_1 = \frac{1}{\sqrt{\sum_{j=1}^{p} \beta_j^2 + 1}} \begin{bmatrix} \beta_1 & & & \\ & \ddots & & \\ & & \beta_p & \\ & & & 1 \end{bmatrix}. \tag{10}$$

Observe that Δ_0 is positive definite, whereas Δ_1 is *indefinite*. We now apply bisection on the interval $[0, 1]$, according to the positive (semi)definiteness of the normalized solution Δ_t for (9). By resetting $\Delta_t = -\Delta_t$ if Δ_t is negative semidefinite, we ensure that Δ_t is either positive (semi)definite or indefinite. We shall stop the procedure immediately once Δ_t is positive semidefinite with rank less than $p+1$, because we have found a solution with rank no more than p. In the remainder we need only to be concerned with the situation when this procedure does not stop finitely. At any rate, this enables us to get hold of two sequences of matrices, $\{\Delta_{t_k} \mid k \in K\}$ and $\{\Delta_{t_l} \mid l \in L\}$, where Δ_{t_k} is positive definite for $k \in K$, and Δ_{t_l} is indefinite for $l \in L$; moreover, $t_l - t_k \to 0$ as $k, l \to \infty$, $k \in K$, $l \in L$. However, it is always possible to combine a positive definite matrix with an indefinite matrix to come up with a *positive semidefinite* matrix (i.e., reducing the rank of the positive definite matrix). This rank-reduction operation, however, may violate Equation (3). Fortunately, because t_k and t_l are sufficiently close, this "equation violation" can be controlled. Below we shall formally present the above-described procedure, followed by the error analysis.

Finding a Lower-Rank Solution

Initialization: Let $p_0 = 0$ and $i_0 = 1$, $k = 0$, and the solutions for (9), at $t = p_0$ (=0) and $t = i_0$ (=1), respectively, be as given by (10).

Checking termination: Let $t = (p_k + i_k)/2$, and solve (9) (with $\|\Delta_t\|_F = 1$). If Δ_t is positive semidefinite with $\operatorname{rank}(\Delta_t) \le p$, stop, and we have found a nontrivial solution with rank less than $p+1$ in the set $\mathcal{L} \cap \mathcal{SF}_+^n$; else, continue with the next step.

Update: If Δ_t is positive definite, then let $p_{k+1} := t$ and $i_{k+1} := i_k$; else, if Δ_t is indefinite, then let $p_{k+1} := p_k$ and $i_{k+1} := t$.

Let

$$D_k := (\Delta_{p_k} + \tau_k \Delta_{i_k})/\|\Delta_{p_k} + \tau_k \Delta_{i_k}\|_F,$$

where

$$\tau_k = \begin{cases} \arg\max\{\tau > 0 \mid \Delta_{p_k} + \tau\Delta_{i_k} \succeq 0\}, & \text{if } \Delta_{p_k} \cdot \Delta_{i_k} \ge 0; \\ \arg\max\{\tau > 0 \mid \Delta_{p_k} - \tau\Delta_{i_k} \succeq 0\}, & \text{if } \Delta_{p_k} \cdot \Delta_{i_k} < 0, \end{cases}$$

and

$$X_k := W_{p_k} D_k W_{p_k}^{\mathrm{H}}. \tag{11}$$

Increase the iteration counter to $k := k+1$, and return to *Checking termination*.

A few comments are in order here.

(i) In this notation, the "p" in the index "p_k" stands for the "positive definiteness" nature of Δ_{p_k} (corresponding to the index set K as we discussed earlier), and "i" in the index "i_k" is for the "indefiniteness" nature of Δ_{i_k} (corresponding to the index set L), where k is the iteration counter.

(ii) Clearly, $i_k - p_k = 2^{-k}$, $k = 0, 1, 2, \ldots$, due to the bisection.

(iii) Because Δ_{p_k} is positive definite and Δ_{i_k} is indefinite in our algorithmic procedure, τ_k is always finite, and $\Delta_{p_k} + \tau_k \Delta_{i_k}$ can never be zero either. Moreover, by this construction, we have $\operatorname{rank}(X_k) \le p$. By renaming $-\Delta_{i_k}$ to Δ_{i_k} if necessary, we assume from now on that $\Delta_{p_k} \cdot \Delta_{i_k} \ge 0$ for all k.

Because

$$\|(1-t)w_{p+1} + t w_{p+2}\|^2 = (1-t)^2 + t^2 \in [0.5, 1] \quad \text{for all } t \in [0, 1],$$

and

$$W_{p_k}^{\mathrm{H}} W_{p_k} = \begin{bmatrix} I_p & 0_p \\ 0_p^{\mathrm{T}} & \|(1-p_k)w_{p+1} + p_k w_{p+2}\|^2 \end{bmatrix}, \tag{12}$$

we have $0.5 I_{p+1} \preceq W_{p_k}^{\mathrm{H}} W_{p_k} \preceq I_{p+1}$, and $p \le \operatorname{tr}(W_{p_k}^{\mathrm{H}} W_{p_k}) \le p+1$. Thus,

$$\|X_k\|_F^2 = \operatorname{tr}\big(W_{p_k} D_k W_{p_k}^{\mathrm{H}} W_{p_k} D_k W_{p_k}^{\mathrm{H}}\big) \ge \frac{1}{2}\operatorname{tr}\big(D_k^2 W_{p_k}^{\mathrm{H}} W_{p_k}\big) \ge \frac{1}{4}\|D_k\|_F^2 = 0.25,$$

and also $\|X_k\|_F \le 1$, for all iteration counter k.

Therefore, the matrix solution X_k satisfies $0.25 \le \|X_k\|_F \le 1$, and $\operatorname{rank}(X_k) \le p$ for all k. However, X_k may not exactly be on the subspace $\mathcal{L}$ anymore. In the remaining part of the analysis, we shall bound this error. Before proceeding, we first note the following estimations:

$$\|\Delta_{p_k} + \tau_k \Delta_{i_k}\|_F^2 = \|\Delta_{p_k}\|_F^2 + \tau_k^2 \|\Delta_{i_k}\|_F^2 + 2\tau_k \Delta_{p_k} \cdot \Delta_{i_k} \ge 1 + \tau_k^2. \tag{13}$$

Let $\widehat{W} := [\underbrace{0_n, \ldots, 0_n}_{p}, w_{p+1} - w_{p+2}] \in \mathcal{F}^{n\times(p+1)}$. We have $\|\widehat{W}\|_F = \sqrt{2}$, and

$$W_{p_k} = W_{i_k} + (i_k - p_k)\widehat{W} = W_{i_k} + \widehat{W}/2^k. \tag{14}$$

For any $1 \le j \le m$ and $k \ge 1$, we have

$$\begin{aligned} |A_j \cdot X_k| &= \big|A_j \cdot (W_{p_k} D_k W_{p_k}^{\mathrm{H}})\big| \\ &= \frac{1}{\|\Delta_{p_k} + \tau_k \Delta_{i_k}\|_F}\big|A_j \cdot \big(W_{p_k}(\Delta_{p_k} + \tau_k \Delta_{i_k})W_{p_k}^{\mathrm{H}}\big)\big| \\ &= \frac{\tau_k}{\|\Delta_{p_k} + \tau_k \Delta_{i_k}\|_F}\big|A_j \cdot \big(W_{p_k} \Delta_{i_k} W_{p_k}^{\mathrm{H}}\big)\big| \end{aligned}$$

$$\begin{aligned}
&\le \frac{\tau_k}{\sqrt{1+\tau_k^2}}\left|A_j\cdot\left(W_{p_k}\Delta_{i_k}W_{p_k}^{\mathrm{H}}\right)\right| \quad \text{[here, we use (13)]}\\
&\le \left|A_j\cdot\left(W_{p_k}\Delta_{i_k}W_{p_k}^{\mathrm{H}}\right)\right|\\
&= |A_j\cdot((W_{i_k}+\widehat{W}/2^k)\Delta_{i_k}(W_{i_k}+\widehat{W}/2^k)^{\mathrm{H}})| \quad \text{[here, we use (14)]}\\
&= |A_j\cdot((W_{i_k}\Delta_{i_k}\widehat{W}^{\mathrm{H}}+\widehat{W}\Delta_{i_k}W_{i_k}^{\mathrm{H}})/2^k+\widehat{W}\Delta_{i_k}\widehat{W}^*/4^k)|\\
&\le \|A_j\|_F(2\|W_{i_k}\Delta_{i_k}\|_F\cdot\|\widehat{W}\|_F/2^k+\|\widehat{W}\|_F^2\cdot\|\Delta_{i_k}\|_F/4^k)\\
&\le \frac{5\|A_j\|_F}{2^k}. \quad \left[\because \text{ by (12)}, \|W_{i_k}\Delta_{i_k}\|_F^2=\mathrm{tr}(\Delta_{i_k}W_{i_k}^TW_{i_k}\Delta_{i_k})\le\mathrm{tr}(\Delta_{i_k}^2)=1\right].
\end{aligned}$$

Therefore, for any given precision $\epsilon>0$, if $k\ge\log(4\max_{1\le j\le m}\|A_j\|_F/\epsilon)$, then $|A_j\cdot X_k|\le\epsilon$ for all $1\le j\le m$. Using the famous error bound result of Hoffman for the polyhedral systems (thus includes the linear subspace as a special case; see Hoffman [8] or Pang [12]), it follows that

$$\mathrm{dist}(X_k,\mathcal{L})\le O(\epsilon).$$

By taking a cluster point of $\{X_k\mid k=1,2,\ldots\}$, say $\widehat{X}$, then $\widehat{X}\in\mathcal{L}$, $\mathrm{rank}(\widehat{X})\le p$ and $\widehat{X}\succeq 0$. The theorem is proven. □

The following example is due to Bohnenblust [6], which shows that the bound in Theorem 2.3 is tight if no further information is available.

Example 2.2. Let $1\le p\le n-2$ and consider

$$\mathcal{L}=\left\{X=\begin{bmatrix}\lambda I_{p+1} & X_{12}\\ X_{21} & X_{22}\end{bmatrix}\in\mathcal{SF}^n\;\middle|\;\lambda\in\Re,\ \mathrm{tr}X=n\lambda,\ X_{12}\in\mathcal{F}^{(p+1)\times(n-p-1)},\ X_{22}\in\mathcal{SF}^{n-p-1}\right\}.$$

One can compute that $\mathrm{codim}(\mathcal{L})=d_{\mathcal{F}}(p+1)$ in this case. It is clear that $\mathcal{L}$ does not contain any nontrivial positive semidefinite matrices with rank less than $p+1$, whereas it indeed does contain a nontrivial positive semidefinite matrix with rank equal to $p+1$:

$$\begin{bmatrix}
I_p & 0 & 0 & 0_{p\times(n-p-2)}\\
0 & 1 & \sqrt{n-p-1} & 0\\
0 & \sqrt{n-p-1} & n-p-1 & 0\\
0_{(n-p-2)\times p} & 0 & 0 & 0_{(n-p-2)\times(n-p-2)}
\end{bmatrix}.$$

However, if the linear subspace $\mathcal{L}$ in Theorem 2.3 is replaced by an affine space with a bounded intersection with the cone of positive semidefinite matrices, then the bound on the codimension can be further relaxed; see the theorem below. This result is an extension of Au-Yeung and Poon [1] and Barvinok [2].

Theorem 2.4. *Suppose that $\mathcal{A}$ is an affine subspace of $\mathcal{SF}^n$ and that $\mathcal{A}\cap\mathcal{SF}_+^n$ is nontrivial and bounded. If $\mathrm{codim}(\mathcal{A})\le d_{\mathcal{F}}(p+1)$ and $1\le p\le n-2$, then in polynomial time (in terms of n and $\log(1/\epsilon)$) one can find $X\succeq 0$, with $\mathrm{rank}(X)\le p$, and $\mathrm{dist}(X,\mathcal{A})\le\epsilon$. In particular, by taking limits, we conclude that there is $X\in\mathcal{A}\cap\mathcal{SF}_+^n$, such that $\mathrm{rank}(X)\le p$.*

Proof. Let

$$\mathcal{A}=\{X\in\mathcal{SF}^n\mid A_j\cdot X=b_j,\ j=1,\ldots,m\},$$

where A_js are linearly independent, b_js are not all zero, and $m=\mathrm{codim}(\mathcal{A})$. Without loss of generality, we assume $b_m>0$. Consider the subspace as follows

$$\mathcal{L}=\left\{X\in\mathcal{SF}^n\;\middle|\;\left(A_i-\frac{b_i}{b_m}A_m\right)\cdot X=0,\ i=1,2,\ldots,m-1\right\}.$$

Of course, $\mathrm{codim}(\mathcal{L})=m-1\le d_{\mathcal{F}}(p+1)-1$. Using Theorem 2.3, we can get $X(\epsilon)\succeq 0$ in polynomial time, such that:

(a) $\mathrm{rank}(X(\epsilon))\le p$;
(b) $0.25\le\|X(\epsilon)\|_F\le 1$;
(c) $\mathrm{dist}(X(\epsilon),\mathcal{L})\le\epsilon$.

We shall see that in this case, $\liminf_{\epsilon\to 0} A_m \cdot X(\epsilon) > 0$. We prove this by contradiction. If there is a subsequence ϵ_k, such that $\lim_{\epsilon_k\to 0} A_m \cdot X(\epsilon_k) = 0$ and $\lim_{\epsilon_k\to 0} X(\epsilon_k) = \widehat{X}$, then $A_i \cdot \widehat{X} = 0$, $i = 1, \ldots, m$, and $\widehat{X} \succeq 0$, contradicting the fact that $\mathcal{A} \cap \mathcal{SF}^n_+$ is bounded. Similarly, if there is a subsequence ϵ_k with $\lim_{\epsilon_k\to 0} A_m \cdot X(\epsilon_k) < 0$ and $\lim_{\epsilon_k\to 0} X(\epsilon_k) = \widehat{X}$, then $\widehat{X} := -(b_m/A_m \cdot \widehat{X})\widehat{X}$ satisfies $A_i \cdot \widehat{X} = -b_i$, $i = 1, \ldots, m$. Taking any $X \in \mathcal{A} \cap \mathcal{SF}^n_+$, and letting $\Delta := X + \widehat{X} \succeq 0$, we have $A_i \cdot \Delta = 0$, $i = 1, \ldots, m$. This is again in contradiction with the fact that $\mathcal{A} \cap \mathcal{SF}^n_+$ is bounded. Now that $\liminf_{\epsilon\to 0} A_m \cdot X(\epsilon) > 0$, let us define $X(\epsilon) := (b_m/A_m \cdot X(\epsilon))X(\epsilon)$, which is the desired matrix solution, satisfying:

(i) $X(\epsilon) \in \mathcal{SF}^n_+$;
(ii) $\limsup_{\epsilon\to 0} \|X(\epsilon)\|_F < \infty$;
(iii) $\operatorname{rank}(X(\epsilon)) \le p$;
(iv) $\operatorname{dist}(X(\epsilon), \mathcal{A}) = O(\epsilon)$.

Using the error bound result for the LMI systems (Theorem 3.3 in Sturm [17], or Theorem 7.4.2 in Luo and Sturm [11]), we have $\lim_{\epsilon\to 0} \operatorname{dist}(X(\epsilon), \mathcal{A} \cap \mathcal{SF}^n_+) = 0$. If, additionally, $\mathcal{A} \cap \mathcal{SF}^n_{++} \neq \varnothing$, then by the error bound result in Zhang [20] (Theorem 2.5), we have $\operatorname{dist}(X(\epsilon), \mathcal{A} \cap \mathcal{SF}^n_+) = O(\epsilon)$. By taking a limit, the desired result follows. □

Observe that Theorem 2.4 implies Theorem 2.3 by constructing an affine subspace from a linear subspace as: $\mathcal{A} = \{X \in \mathcal{L} \cap \mathcal{SF}^n_+ \mid I \cdot X = 1\}$. In fact, one observes that Theorem 2.2 follows from Theorem 2.1 by the same construction. Similar to Example 2.2, the bound in Theorem 2.4 cannot be improved in general.

Example 2.3. Let $1 \le p \le n-2$ and consider

$$\mathcal{A} = \left\{ X = \begin{bmatrix} I_{p+1} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \in \mathcal{SF}^n \mid \operatorname{tr} X_{22} = 1 \right\}.$$

One can compute that $\operatorname{codim}(\mathcal{A}) = d_{\mathcal{F}}(p+1) + 1$ in this case. It is clear that $\mathcal{A}$ does not contain any nontrivial positive semidefinite matrices with rank less than $p+1$, whereas it does indeed contain a nontrivial positive semidefinite matrix with rank equal to $p+1$:

$$\begin{bmatrix} I_{p+1} & \dfrac{1}{\sqrt{p+1}} 1_{p+1} & 0_{(p+1)\times(n-p-2)} \\ \dfrac{1}{\sqrt{p+1}} 1^{\mathrm{T}}_{p+1} & 1 & 0 \\ 0_{(n-p-2)\times(p+1)} & 0 & 0_{(n-p-2)\times(n-p-2)} \end{bmatrix}.$$

Similarly, consider

$$\mathcal{A} = \left\{ X = \begin{bmatrix} I_{p+1} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \in \mathcal{SF}^n \right\}.$$

Then, $\operatorname{codim}(\mathcal{A}) = d_{\mathcal{F}}(p+1)$. However, it does not contain any positive semideinite matrix with rank less than $p+1$. This shows that the boundedness assumption on $\mathcal{A} \cap \mathcal{SF}^n_+$ in Theorem 2.4 cannot be removed in general.

Using the same technique, the following analogous results in the context of linear matrix inequalities (replacing linear subspace by a polyhedral cone) can be similarly shown.

Corollary 2.1. *Let $\mathcal{A} = \{X \in \mathcal{SF}^n \mid A_i \cdot X \trianglelefteq_i 0,\ i = 1, \ldots, m\}$, where $\trianglelefteq_i \in \{\le, \ge, =\}$, $i = 1, \ldots, m$. Suppose that $\dim(\mathcal{A} \cap \mathcal{SF}^n_+) \ge 1$, $m \le d_{\mathcal{F}}(p+1) - 1$, and $1 \le p \le n-2$. Then, one can find in polynomial time (of n and $\log(1/\epsilon)$) a matrix X such that $\|X\|_F = 1$, $X \in \mathcal{SF}^n_+$, $\operatorname{rank}(X) \le p$, and $\operatorname{dist}(X, \mathcal{A}) \le \epsilon$. In particular, this implies that there is a nontrivial $X \in \mathcal{A} \cap \mathcal{SF}^n_+$ with $\operatorname{rank}(X) \le p$.*

Proof. Follow exactly the same steps as in the proof for Theorem 2.3, except that the direction-finding equations are now expanded sequentially,

$$(U^{\mathrm{H}} A_i U) \cdot \Delta = 0, \quad \text{for } i \text{ with } A_i \cdot X = 0, \tag{15}$$

where $X = UU^{\mathrm{H}}$. By letting $X(t) := U(I + t\Delta)U^{\mathrm{H}}$ and increasing t, we either arrive at a reduction of the rank of X, or adding a new index to the direction-finding equations (15). In the worst case, all the equations are added, and then the situation is the same as in Theorem 2.3. □

Corollary 2.2. *Let* $\mathcal{A} = \{X \in \mathcal{SF}^n \mid A_i \cdot X \trianglelefteq b_i,\ i = 1, \ldots, m\}$, *where* $\trianglelefteq_i \in \{\leq, \geq, =\}$, $i = 1, \ldots, m$. *Suppose that* $\mathcal{A} \cap \mathcal{SF}^n_+$ *is nonempty and bounded. Moreover, suppose that* $m \leq d_{\mathcal{F}}(p+1)$ *and* $1 \leq p \leq n-2$. *Then, one can find in polynomial time (of* n *and* $\log(1/\epsilon)$*) a matrix* X *such that* $X \in \mathcal{SF}^n_+$, $\operatorname{rank}(X) \leq p$, *and* $\operatorname{dist}(X, \mathcal{A}) \leq \epsilon$. *In particular, this implies that there is a nontrivial* $X \in \mathcal{A} \cap \mathcal{SF}^n_+$ *with* $\operatorname{rank}(X) \leq p$.

If $\mathcal{A}$ contains inequality constraints (e.g., $\mathcal{A} \cap \mathcal{SF}^n_+ = \{X \in \mathcal{SF}^n_+ \mid A_i \cdot X \leq b_i,\ i = 1, \ldots, m\}$), then Pataki (Theorem 2.2 in Pataki [13]) asserted that $d_{\mathcal{F}}(\operatorname{rank}(X)) + s = m$, where s is the number of *inactive* constraints at X. Therefore, if there is at least one inactive constraint, say $A_i \cdot X < b_i$ for some i, then $d_{\mathcal{F}}(\operatorname{rank}(X)) \leq m - 1 \leq d_{\mathcal{F}}(p+1) - 1$, and so consequently $\operatorname{rank}(X) \leq p$. In other words, if there is a solution $X \in \mathcal{A} \cap \mathcal{SF}^n_+$ with at least one inactive constraint, then Pataki's result would also lead to the conclusion of Corollary 2.2.

3. Connections and extensions. Consider a given Euclidean space $\mathcal{V}$. For a convex cone $\mathcal{K} \subseteq \mathcal{V}$, its dual cone is denoted as $\mathcal{K}^* := \{y \in \mathcal{V} \mid x^{\mathrm{T}} y \geq 0,\ \text{for all } x \in \mathcal{K}\}$. We first note the following useful fact regarding the convex cones and their dual objects.

Lemma 3.1. *Let* $\mathcal{U}, \mathcal{K} \subseteq \mathcal{V}$ *be two nonempty closed convex cones. Suppose that* $\mathcal{K}$ *is pointed, i.e.,* $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$. *Then,* $\dim(\mathcal{U} \cap \mathcal{K}) = 0$ *if and only if* $(-\mathcal{U}^*) \cap \operatorname{int}(\mathcal{K}^*) \neq \varnothing$.

Proof. "If" part. Let $0 \neq y \in (-\mathcal{U}^*) \cap \operatorname{int}(\mathcal{K}^*)$. Then, for any $0 \neq x \in \mathcal{K}$ it follows that $x^{\mathrm{T}} y > 0$. Thus, if there is any $0 \neq x \in \mathcal{U} \cap \mathcal{K}$, then because $y \in -\mathcal{U}^*$ it follows that $x^{\mathrm{T}} y \leq 0$, yielding a contradiction. Thus, we must have $\mathcal{U} \cap \mathcal{K} = \{0\}$.

"Only if" part. If $\mathcal{U} \cap \mathcal{K} = \{0\}$, then by the duality relation we have

$$\mathcal{V} = \{0\}^* = (\mathcal{U} \cap \mathcal{K})^* = \operatorname{cl}(\mathcal{U}^* + \mathcal{K}^*). \tag{16}$$

Suppose by contradiction that $(-\mathcal{U}^*) \cap \operatorname{int}(\mathcal{K}^*) = \varnothing$. Because $\mathcal{K}$ is pointed, $\operatorname{int}(\mathcal{K}^*)$ is nonempty ($-\mathcal{U}^*$ is obviously nonempty), and so we can apply the separation theorem to conclude that there exists $0 \neq d \in \mathcal{V}$ such that $d^{\mathrm{T}}(-y) \leq 0$ for all $-y \in -\mathcal{U}^*$ (i.e., $d^{\mathrm{T}} y \geq 0$ for all $y \in \mathcal{U}^*$), and $d^{\mathrm{T}} z \geq 0$ for all $z \in \operatorname{int}(\mathcal{K}^*)$. Now, by (16), because $-d \in \mathcal{V}$, there should exist a sequence $y_i \in \mathcal{U}^*$ and $z_i \in \mathcal{K}^*$, $i = 1, 2, \ldots$, such that

$$-d = \lim_{i \to \infty} (y_i + z_i).$$

This leads to the following contradiction:

$$0 > -\|d\|^2 = \lim_{i \to \infty} d^{\mathrm{T}}(y_i + z_i) \geq 0.$$

The lemma is thus proven. □

Now let us take $\mathcal{V} = \mathcal{SF}^n$, $\mathcal{K} = \mathcal{SF}^n_+$, $\mathcal{U} = \mathcal{L} = \{X \in \mathcal{SF}^n \mid A_i \cdot X = 0,\ i = 1, \ldots, m\}$. It follows that $\mathcal{K}^* = \mathcal{K} = \mathcal{SF}^n_+$, and $-\mathcal{U}^* = \mathcal{U}^* = \mathcal{L}^\perp = \{Z \in \mathcal{SF}^n \mid Z = \sum_{i=1}^k y_i A_i,\ y \in \Re^m\}$. Moreover, let us define $\mathcal{L}_p := \{X \in \mathcal{SF}^n \mid A_i \cdot X = 0,\ i = 1, \ldots, m,\ \operatorname{rank}(X) \leq p\}$. Then, the following holds.

Theorem 3.1. *Let* $A_1, A_2, \ldots, A_m \in \mathcal{SF}^n$. *The following four statements are equivalent*:

(i) *There exists* $y \in \Re^m$ *such that* $\sum_{i=1}^m y_i A_i \succ 0$.

(ii) $\dim(\mathcal{L} \cap \mathcal{SF}^n_+) = 0$.

(iii) $\dim(\mathcal{L}_p \cap \mathcal{SF}^n_+) = 0$, *provided that* $1 \leq m \leq d_{\mathcal{F}}(p+1) - 1$ *and* $1 \leq p \leq n-2$.

(iv) $\dim(\mathcal{L}_p \cap \mathcal{SF}^n_+) = 0$ *and the set* $K := \{(A_1 \cdot X, A_2 \cdot X, \ldots, A_m \cdot X) \mid X \in \mathcal{SF}^n_+, \operatorname{rank}(X) \leq p\}$ *is a pointed closed convex cone, provided that* $1 \leq m \leq d_{\mathcal{F}}(p+1)$ *and* $1 \leq p \leq n-2$.

Proof.

(i) $\iff$ (ii). By Lemma 3.1, (i) $\iff \mathcal{L}^\perp \cap \mathcal{SF}^n_+ \neq \varnothing \iff \mathcal{L} \cap \mathcal{SF}^n_+ = \{0_{n\times n}\}$, i.e., (ii).

(ii) $\iff$ (iii). The "(ii) $\Longrightarrow$ (iii)" part is due to $\mathcal{L}_p \subseteq \mathcal{L}$, and the "(ii) $\Longleftarrow$ (iii)" part follows from Theorem 2.3.

(iv) $\Longrightarrow$ (i). Because K is a pointed closed convex cone, by the separation theorem there is $y \in \Re^m$, such that $y^T v > 0$ for all $0 \neq v \in K$. Due to $\mathcal{L}_p \cap \mathcal{SF}^n_+ = \{0_{n\times n}\}$, it follows that

$$(y_1 A_1 + y_2 A_2 + \cdots + y_m A_m) \cdot X > 0 \quad \text{for all } 0_{n\times n} \neq X \in \mathcal{SF}^n_+ \text{ and } \operatorname{rank}(X) \leq p.$$

Hence, $y_1 A_1 + y_2 A_2 + \cdots + y_m A_m \succ 0$.

(ii) $\Longrightarrow$ (iv). $\mathcal{L} \cap \mathcal{SF}^n_+ = \{0_{n\times n}\}$ implies that, for any $0_m \neq v \in \Re^m$, the set $\{X \mid A_i \cdot X = v_i,\ i = 1, 2, \ldots, m\} \cap \mathcal{SF}^n_+$ must always be bounded. Using Theorem 2.4, we conclude that for any $v \in \operatorname{conv}(K) =$

$\{(A_1 \cdot X, A_2 \cdot X, \ldots, A_m \cdot X) \mid X \in \mathcal{SF}_+^n\}$ it actually follows that $v \in K = \{(A_1 \cdot X, A_2 \cdot X, \ldots, A_m \cdot X) \mid X \in \mathcal{SF}_+^n, \operatorname{rank}(X) \le p\}$. That is, K is a convex cone. Now, $\mathcal{L} \cap \mathcal{SF}_+^n = \{0_{n\times n}\}$ implies that K is pointed. To see this, for the sake of deriving a contradiction, let us suppose that there is $0_m \neq v = (A_1 \cdot X_1, A_2 \cdot X_1, \ldots, A_m \cdot X_1)$ and $-v = (A_1 \cdot X_2, A_2 \cdot X_2, \ldots, A_m \cdot X_2)$ for some $X_1, X_2 \succeq 0$. Then, $X_1 + X_2 \in \mathcal{L} \cap \mathcal{SF}_+^n$ and so $X_1 = X_2 = 0_{n\times n}$, and this is in contradiction with the fact that $v \neq 0_m$. □

In particular, the equivalence between (i) and (iii) was first shown by Bohnenblust in an unpublished manuscript (Bohnenblust [6]); see also §2.4 of Hiriart-Urruty and Torki [10]. As we shall see below, the equivalence between (i) and (iv) leads to a result of Polyak [16], by setting $\mathcal{F} = \Re$ and $p = 1$.

Proposition 3.1 (Theorem 2.1 in Polyak [16]). *For $n \ge 3$ and real symmetric matrices A_1, A_2, A_3, the following statements are equivalent*:

(i) *There is $\mu \in \Re^3$ such that*

$$\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 \succ 0.$$

(ii) *The set $K := \{(x^{\mathrm T} A_1 x, x^{\mathrm T} A_2 x, x^{\mathrm T} A_3 x) \mid x \in \Re^n\}$ is a pointed closed convex cone, and*

$$(x^{\mathrm T} A_1 x, x^{\mathrm T} A_2 x, x^{\mathrm T} A_3 x) \neq (0,0,0),\ \forall\, x \in \Re^n \backslash \{0_n\}.$$

As an example, if we let $\mathcal{F} = \mathbf{C}$ and $p = 1$, then Theorem 3.1 leads to the following extension of Polyak's result.

Theorem 3.2. *For $n \ge 3$ and Hermitian matrices A_1, A_2, A_3, A_4, the following statements are equivalent*:

(i) *There is $\mu \in \Re^4$ such that*

$$\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \mu_4 A_4 \succ 0.$$

(ii) *The set $K := \{(x^{\mathrm H} A_1 x, x^{\mathrm H} A_2 x, x^{\mathrm H} A_3 x, x^{\mathrm H} A_4 x) \mid x \in \mathbf{C}^n\}$ is a pointed closed convex cone, and*

$$(x^{\mathrm H} A_1 x, x^{\mathrm H} A_2 x, x^{\mathrm H} A_3 x, x^{\mathrm H} A_4 x) \neq (0,0,0,0), \quad \forall\, x \in \mathbf{C}^n \backslash \{0_n\}.$$

Another important aspect of Theorem 3.1 is its connection to the so-called S-lemma, which is stated as follows. For any two n by n symmetric matrices B, A_1, if there is $x_0 \in \Re^n$ satisfying $x_0^{\mathrm T} A_1 x_0 < 0$, then the implication $x^{\mathrm T} A_1 x \le 0 \Longrightarrow x^{\mathrm T} B x \ge 0$ is equivalent to the existence of $\mu_1 \ge 0$ such that $B + \mu_1 A_1 \succeq 0$. The S-lemma has played an instrumental role in optimization and control theory (see Pólik and Terlaky [15] for a recent survey on the S-lemma). In fact, Theorem 2.4 and Theorem 3.1 lead to the following extension of the S-lemma.

Theorem 3.3. *Let n, m, p be integers satisfying $1 \le m \le d_{\mathcal{F}}(p+1) - 1$ and $1 \le p \le n-2$. Suppose that $B \in \mathcal{SF}^n$ and $A_k \in \mathcal{SF}^n$, $k = 1, \ldots, m$ satisfy the following two conditions*: (1) *The equations $B \cdot X = 0$, $A_1 \cdot X = 0, \ldots, A_m \cdot X = 0$, and $X \in \mathcal{SF}_+^n$ have only trivial solution*; (2) *there is $X_0 \in \mathcal{SF}_+^n$ such that $A_k \cdot X_0 < 0$, $k = 1, \ldots, m$. Then, the implication that $B \cdot X \ge 0$ holds for all $A_k \cdot X \le 0$, $k = 1, \ldots, m$, $X \in \mathcal{SF}_+^n$, and $\operatorname{rank}(X) \le p$, is equivalent to the existence of $\mu_k \ge 0$, $k = 1, \ldots, m$, such that $B + \mu_1 A_1 + \cdots + \mu_m A_m \succeq 0$.*

Proof. The part "⟸" is obvious.

We need only to show the part "⟹." Due to Condition (1), it follows from Theorem 2.4 (also see the proof for Theorem 3.1) that the cone

$$K = \{(B \cdot X, A_1 \cdot X, A_2 \cdot X, \ldots, A_m \cdot X) \mid X \in \mathcal{SF}_+^n,\ \operatorname{rank}(X) \le p\}$$

is a closed convex cone. Moreover, $K \cap (\Re^1_{--} \times \Re^m_-) = \varnothing$. Applying the separation theorem on these two convex sets, there exist $\mu_0 \ge 0, \mu_1 \ge 0, \ldots, \mu_m \ge 0$, which are not zero simultaneously, such that $\mu_0 B \cdot X + \sum_{k=1}^m \mu_k A_k \cdot X \ge 0$ for all $X \in \mathcal{SF}_+^n$ with $\operatorname{rank}(X) \le p$. Therefore, $\mu_0 B + \sum_{k=1}^m \mu_k A_k \succeq 0$. We now invoke Condition (2) to ensure that $\mu_0 > 0$, and the theorem follows by letting $\mu_k := \mu_k / \mu_0$, $k = 1, \ldots, m$. □

Clearly, if the conditions in Theorem 3.3 hold, then it follows immediately from Theorem 3.1 directly that there is $y \in \Re^{m+1}$ such that $y_0 B + \sum_{k=1}^m y_k A_k \succ 0$. However, Theorem 3.3 gives more information on the sign of the y vector. In case $p = 1$ (corresponding to the original S-lemma), Theorem 3.3 requires two more conditions, as compared to the original S-lemma, to be effective: Condition (1) and $n \ge 3$. However, it asserts that the S-procedure is lossless for two quadratic forms ($m = d_{\Re}(2) - 1 = 2$) in the real case (see also Theorem 4.1 of Polyak [16]), and three quadratic forms ($m = d_{\mathrm C}(2) - 1 = 3$) in the complex case.

We shall mention that one can extend the result in Theorem 3.1 to accommodate polyhedral cones instead of subspaces. Let $\trianglelefteq \in \{\leq, \geq, =\}$, and the corresponding dual operators are

$$\trianglelefteq^* \text{ is } \begin{cases} \leq, & \text{if } \trianglelefteq \text{ is } \leq; \\ \geq, & \text{if } \trianglelefteq \text{ is } \geq; \\ \text{unrestricted}, & \text{if } \trianglelefteq \text{ is } =. \end{cases}$$

By combining Corollary 2.1 with Lemma 3.1, we have the following result.

THEOREM 3.4. *Let $1 \leq m \leq d_{\mathcal{F}}(p+1) - 1$, and $1 \leq p \leq n-2$. Let $\trianglelefteq_i \in \{\leq, \geq, =\}$, $i = 1, \ldots, m$. Then, the following implication*

$$A_i \cdot X \trianglelefteq_i 0, \quad i = 1, \ldots, m, \quad X \succeq 0, \qquad \operatorname{rank}(X) \leq p \implies X \text{ is a zero matrix}$$

is equivalent to the fact that there exists $y \in \Re^m$, with $-y_i \trianglelefteq_i^* 0$, $i = 1, \ldots, m$, such that $\sum_{i=1}^m y_i A_i \succ 0$.

There is a natural connection between the rank of a positive semidefinite matrix and the multiplicity of its eigenvalues. Let X belong to $\mathcal{SF}^n$, and arrange the eigenvalues of X in ascending order

$$\lambda_1(X) \leq \lambda_2(X) \leq \cdots \lambda_n(X).$$

We say that $\lambda_1(X)$ is of multiplicity k if

$$\lambda_1(X) = \cdots = \lambda_k(X) < \lambda_{k+1}(X).$$

Friedland and Loewy [7] showed the following result (see Theorem 1 of Friedland and Loewy [7]):

THEOREM 3.5. *Let $\mathcal{L}$ be a subspace of $\mathcal{SF}^n$, and $m = \operatorname{codim}(\mathcal{L})$. Let $1 \leq p \leq n-2$ be a given integer. If $m \leq d_{\mathcal{F}}(p+1)$, then there is a nontrivial $X \in \mathcal{L}$ such that the minimum eigenvalue of X is at least of multiplicity $n-p$.*

In fact, the above theorem is shown to be equivalent to Bohnenblust's result [6] (equivalence between (i) and (iii) in Theorem 3.1); see §3 of Friedland and Loewy [7].

Now we shall prove that there is a polynomial algorithm to find a nontrivial matrix X such that $\operatorname{dist}(X, \mathcal{L}) \leq \epsilon$ and the minimum eigenvalue of X is at least of multiplicity $n-p$.

THEOREM 3.6. *Let $\mathcal{L}$ be a subspace of $\mathcal{SF}^n$, and $m = \operatorname{codim}(\mathcal{L})$. Let $1 \leq p \leq n-2$ be a given integer. If $m \leq d_{\mathcal{F}}(p+1)$, then there is a polynomial-time algorithm to find a nontrivial matrix X such that $\operatorname{dist}(X, \mathcal{L}) = O(\epsilon)$ and the minimum eigenvalue of X has a multiplicity at least $n-p$.*

PROOF. Let

$$\mathcal{L} = \{X \mid A_i \cdot X = 0,\ i = 1, \ldots, m\}.$$

It will be trivial if the identity matrix $I_n \in \mathcal{L}$. Suppose $I_n \notin \mathcal{L}$, and let $A_i \cdot I_n = b_i$, $i = 1, \ldots, m$. Without loss of generality, we assume that $b_m > 0$. Let $\mathcal{U}$ be the linear span of $\mathcal{L} \cup \{I_n\}$, which can be written as

$$\mathcal{U} = \left\{ X \,\middle|\, \left(A_i - \frac{b_i}{b_m} A_m\right) \cdot X = 0,\ i = 1, \ldots, m-1 \right\}.$$

Clearly, $\operatorname{codim}(\mathcal{U}) = \operatorname{codim}(\mathcal{L}) - 1 \leq d_{\mathcal{F}}(p+1) - 1$, due to $\dim(\mathcal{U}) = \dim(\mathcal{L}) + 1$. Now, $I_n \in \mathcal{U} \cap \mathcal{SF}_+^n$ and $1 \leq p \leq n-2$, and it follows by Theorem 2.3 that there is a polynomial procedure to find an $X(\epsilon) \succeq 0$ such that $\operatorname{dist}(X(\epsilon), \mathcal{U}) = O(\epsilon)$ and $\operatorname{rank} X(\epsilon) \leq p$. Define $\mathcal{A} = \mathcal{L} + I_n$. Evidently, $\mathcal{A} = \{X \mid A_i \cdot X = b_i,\ i = 1, \ldots, m\}$.

If $A_m \cdot X(\epsilon) \to 0$, then $\operatorname{dist}(X(\epsilon), \mathcal{L}) = O(A_m \cdot X(\epsilon))$ and the minimum eigenvalue of $X(\epsilon)$, which is zero, is at least of multiplicity $n-p$. Otherwise, if $\liminf_{\epsilon \to o} A_m \cdot X(\epsilon) > 0$, then we have

$$\frac{b_m}{A_m \cdot X(\epsilon)} X(\epsilon) \succeq 0, \qquad \operatorname{rank}\left(\frac{b_m}{A_m \cdot X(\epsilon)} X(\epsilon)\right) \leq p, \qquad \text{and} \qquad \operatorname{dist}\left(\frac{b_m}{A_m \cdot X(\epsilon)} X(\epsilon),\ \mathcal{A}\right) = O(\epsilon).$$

Let

$$\frac{b_m}{A_m \cdot X(\epsilon)} X(\epsilon) = Q \begin{bmatrix} \operatorname{diag}(0_{n-p}) & & & \\ & \lambda_{n-p+1} & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} Q^{\mathrm{H}},$$

where $\lambda_{n-p+1} \geq 0, \ldots, \lambda_n \geq 0$, and let $Y(\epsilon) = (b_m/(A_m \cdot X(\epsilon)))X(\epsilon) - I_n$, and we have

$$Y(\epsilon) = Q \begin{bmatrix} -I_{n-p} & & & \\ & -1+\lambda_{n-p+1} & & \\ & & \ddots & \\ & & & -1+\lambda_n \end{bmatrix} Q^{\mathrm{H}}.$$

At the same time, $\operatorname{dist}(Y(\epsilon), \mathscr{L}) = O(\epsilon)$. Summarizing these two cases, the theorem is proven. □

To observe the other direction, one can also use Theorem 3.6 to prove Theorem 2.4 straightforwardly as follows. Take any two linearly independent matrices $X_1, X_2 \in \mathscr{A} \cap \mathscr{SF}^n_+$, where $\mathscr{A} = \{X \in \mathscr{SF}^n \mid A_i \cdot X = b_i,\ i = 1, \ldots, m\}$ with $m \leq d_{\mathscr{F}}(p+1)$. In case any one of X_1 or X_2 already has a rank of no more than p, then we are done. Otherwise, let $X = (X_1 + X_2)/2 \in \mathscr{A} \cap \mathscr{SF}^n_+$, and so $\operatorname{rank}(X) = r \geq p+2$. Let $X = UU^{\mathrm{H}}$ with $U \in \mathscr{F}^{n\times r}$ and consider the linear subspace $\mathscr{L} = \{\Delta \in \mathscr{SF}^r \mid (U^{\mathrm{H}} A_i U) \cdot \Delta = 0,\ i = 1, \ldots, m\}$. Applying Theorem 3.6 gives us $\Delta(\epsilon)$ such that $(U^{\mathrm{H}} A_i U) \cdot \Delta(\epsilon) = O(\epsilon)$, $i = 1, \ldots, m$, and the minimum eigenvalue of $\Delta(\epsilon)$, say $\lambda_{\min}$, has a multiplicity of at least $r - p$. Also, $\lambda_{\min} < 0$ due to the fact that $\mathscr{A} \cap \mathscr{SF}^n_+$ is bounded. Let $X(\epsilon) := U(I_r - (1/\lambda_{\min})\Delta(\epsilon))U^{\mathrm{H}} \succeq 0$. Clearly, $\operatorname{rank}(X(\epsilon)) \leq r - (r-p) = p$, and $A_i \cdot X(\epsilon) = b_i - (1/\lambda_{\min})(U^{\mathrm{H}} A_i U) \cdot \Delta(\epsilon)$, $i = 1, \ldots, m$. Therefore, $\operatorname{dist}(X(\epsilon), \mathscr{A}) = O(\epsilon)$, and Theorem 2.4 is proven.

Acknowledgments. The authors would like to thank the two referees for their comments, which led to improvements of the presentation. Research was partially supported by Chinese NSFC Earmarked Grant Project 10701016 and Hong Kong RGC Earmarked Grant CUHK418406.

References

[1] Au-Yeung, Y.-H., Y. T. Poon. 1979. A remark on the convexity and positive definiteness concerning Hermitian matrices. *Southeast Asian Bull. Math.* **3** 85–92.
[2] Barvinok, A. 1995. Problems of distance geometry and convex properties of quadratic maps. *Discrete Computat. Geometry* **12** 189–202.
[3] Barvinok, A. 2001. A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete Computat. Geometry* **25** 23–31.
[4] Biswas, P., Y. Ye. 2004. Semidefinite programming for ad hoc wireless sensor network. *Proc. 3rd Internat. Sympos. Inform. Processing Sensor Networks*, Berkeley, CA, 46–54.
[5] Biswas, P., T. C. Liang, K. C. Toh, Y. Ye. 2005. An SDP based approach for anchor-free 3D graph realization. Working paper, Stanford University, Palo Alto, CA.
[6] Bohnenblust, F. 1948. Joint positiveness of matrices. Unpublished manuscript. http://orion.math.uwaterloo.ca/~hwolkowi/henry/book/fronthandbk.d/Bohnenblust.pdf.
[7] Friedland, S., R. Loewy. 1976. Subspaces of symmetric matrices containing matrices with a multiple first egenvalue. *Pacific J. Math.* **62** 389–399.
[8] Hoffman, A. J. 1952. On approximate solutions of systems of linear inequalities. *J. Res. National Bureau Standards* **49** 263–265.
[9] Huang, Y., S. Zhang. 2007. Complex matrix decomposition and quadratic programming. *Math. Oper. Res.* **32** 758–768.
[10] Jiriart-Urruty, J.-B., M. Torki. 2002. Permanently going back and forth between the "quadratic world" and the "convexity world" in optimization. *Appl. Math. Optim.* **45** 169–184.
[11] Luo, Z. Q., J. F. Sturm. 2000. Error analysis. H. Wolkowicz, R. Saigal, L. Vandenberghe, eds. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 163–189.
[12] Pang, J. S. 1997. Error bounds in mathematical programming. *Math. Programming* **79** 299–332.
[13] Pataki, G. 1998. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.* **23** 339–358.
[14] Pfender, F., G. M. Ziegler. 2004. Kissing numbers, sphere packings, and some unexpected proofs. *Notices Amer. Math. Soc.* **51**(8) 873–883.
[15] Pólik, I., T. Terlaky. 2007. A survey of the S-lemma. *SIAM Rev.* **49** 371–418.
[16] Polyak, B. T. 1998. Convexity of quadratic transformations and its use in control and optimization. *J. Optim. Theory Appl.* **99** 553–583.
[17] Sturm, J. F. 2000. Error bounds for linear matrix inequalities. *SIAM J. Optim.* **10** 1228–1248.
[18] Sturm, J. F., S. Zhang. 2003. On cones of nonnegative quadratic functions. *Math. Oper. Res.* **28** 246–267.
[19] Ye, Y., S. Zhang. 2003. New results on quadratic minimization. *SIAM J. Optim.* **14** 245–267.
[20] Zhang, S. 2000. Global error bounds for convex conic problems. *SIAM J. Optim.* **10** 836–851.