Disk matrices and the proximal mapping for the numerical radius

X.Y. Han [*]   A.S. Lewis [†]

April 29, 2020

Abstract

Optimal matrices for problems involving the matrix numerical radius often have fields of values that are disks, a phenomenon associated with partial smoothness. Such matrices are highly structured: we experiment in particular with the proximal mapping for the radius, which often maps n-by-n random matrix inputs into a particular manifold of disk matrices that has real codimension $2n$. The outputs, computed via semidefinite programming, also satisfy an unusual rank property at optimality.

Key words: field of values, numerical radius, proximal mapping, partial smoothness, semidefinite program

AMS 2020 Subject Classification: 15A60, 49J52, 90C22

1 Introduction

For any matrix A in the vector space $\mathbf{M}^n$ of n-by-n complex matrices (for $n > 1$), the field of values (or numerical range) is the set

$$W(A) \;=\; \{u^*Au : u \in \mathbf{C}^n, \|u\|_2 = 1\}.$$

It is a compact convex set, by the Toeplitz-Hausdorff Theorem. The numerical radius of A is

$$r(A) \;=\; \max\{|z| : z \in W(A)\}.$$

_____
[*]ORIE, Cornell University, Ithaca, NY 14853, USA.
[†]ORIE, Cornell University, Ithaca, NY 14853, USA. `people.orie.cornell.edu/aslewis`

The numerical radius is a norm on $\mathbf{M}^n$, so in particular vanishes only for the zero matrix. As a consequence of the "power inequality"

$$r(A^k) \leq \big(r(A)\big)^k \qquad (k = 0, 1, 2, \ldots),$$

it provides a simple measure of the transient stability of the discrete-time dynamical system $x_{k+1} = Ax_k$. For a broad discussion, see [5, §25].

Given its role as a measure of system stability, we may seek to optimize the numerical radius over some given set of matrices. Experiments reported in [7] suggest empirically that the optimal solution of a random feedback-control-type problem is often a *disk matrix*: its field of values is a disk centered at zero. At first sight, this observation is surprising: disk matrices have been studied extensively [9], and although random matrices in $\mathbf{M}^n$ often have field of values that approximate disks, especially for large n (see [2]), the true disk-matrix property is quite restrictive. In particular, any disk matrix must be singular, and in fact [3] its zero eigenvalue must have algebraic multiplicity at least two.

As proved in [7, Theorem 7.4], however, around disk matrices belonging to a special class (which we call *strongly certified*), the numerical radius is *partly smooth*. Specifically, the numerical radius is smooth on the manifold $\mathcal{M}$ consisting of all nearby disk matrices, but along paths crossing $\mathcal{M}$ transversally, its rate of change increases in a jump discontinuity at $\mathcal{M}$. For any minimization problem, partial smoothness of the objective relative to some manifold ensures that a random perturbation produces a distribution of random outputs with local support the manifold. As pointed out in [7], this explains why disk matrices tend to appear as solutions of appropriately randomized optimization problems.

In the current work, we explore the prevalence of disk matrices in numerical radius optimization first noted in [7]. We make three numerical observations and explore the supporting theory. We first consider a canonical optimization problem, computing the proximal map for the numerical radius at random matrix inputs and observing that the outputs are often disk matrices. Secondly, although strong certification appears a stringent requirement, we observe that our output disk matrices are indeed always strongly certified. We therefore conjecture a converse of [7, Theorem 7.4]: that the numerical radius is partly smooth around a manifold of disk matrices if and only if those matrices are strongly certified.

Both the proximal map for the numerical radius and the numerical radius itself can be calculated via semidefinite programming. We observe thirdly that the disk matrix property corresponds to a rank condition in the semidefinite program for the numerical radius. We relate this property to strong certification.

2 A numerical experiment

We consider the *proximal map* for the numerical radius, $\text{prox}_r \colon \mathbf{M}^n \to \mathbf{M}^n$ defined by

$$\text{prox}_r(Y) \;=\; \text{argmin}\big\{r(A) + \lambda\|A - Y\|^2 : A \in \mathbf{M}^n\big\} \qquad (Y \in \mathbf{M}^n),$$

where $\|\cdot\|$ denotes the Frobenius norm, and $\lambda > 0$ is a fixed constant. For each dimension n, we generate a thousand independent random matrices Y of unit norm by rescaling normally distributed matrices. By solving a semidefinite program (described below), we then compute the matrix $A = \text{prox}_r(Y)$ and its numerical radius $r(A)$, along with several associated measures, each of which we plot in its own row in Figure 1.

The first row measures how far the field of values of a matrix A diverges from a disk. By definition, we have

$$\max_{z \in W(A)} \text{Re}(w^* z) \;\leq\; r(A) \quad \text{for all } w \in \mathbf{T}$$

where $\mathbf{T}$ denotes the unit circle in the complex plane $\mathbf{C}$, and disk matrices are characterized by equality for all w. The left-hand side is the maximum eigenvalue of the Hermitian matrix $\frac{1}{2}(w^* A + w A^*)$, so calculating a natural measure of divergence,

$$(2.1) \qquad\qquad 1 \;-\; \frac{1}{r(A)} \min_{w \in \mathbf{T}} \max_{z \in W(A)} \text{Re}(w^* z),$$

is straightforward using the `chebfun` package [4]. This measure vanishes for a nonzero matrix A if and only A is either a disk matrix, or a multiple of the identity (an exceptional case that we do not observe experimentally).

We sequence our thousand random matrices A in ranked increasing order of the divergence measure and preserve this order in each of the subsequent rows of graphs, which plot this same ranking against the various measures for the scaled matrix $\hat{A} = \frac{1}{r(A)}A$. The step functions in the first row of Figure 1, involving several orders of magnitude, clearly suggest an underlying partition of the output set into disk and non-disk matrices. The second row of Figure 1 plots another precise measure of divergence from the disk-matrix property, which we introduce shortly. Numerical limitations inevitably blur the clarity of the step functions: for comparison, for Figure 2 in the Appendix, we repeat the calculations in dimensions $n = 2, 3$ to higher precision, refining the step-function behavior.

The experimental results also suggest that the random output A is often singular, and indeed has a multiple zero eigenvalue. We measure these latter properties, through singular values of A and A^2, in the third and fourth rows of Figure 1 (which incidentally also verify that A is never a multiple of the identity matrix). In fact, in our experiments, these various properties, along with two others related to the associated semidefinite program and plotted in the fifth and sixth rows, appear almost surely equivalent.
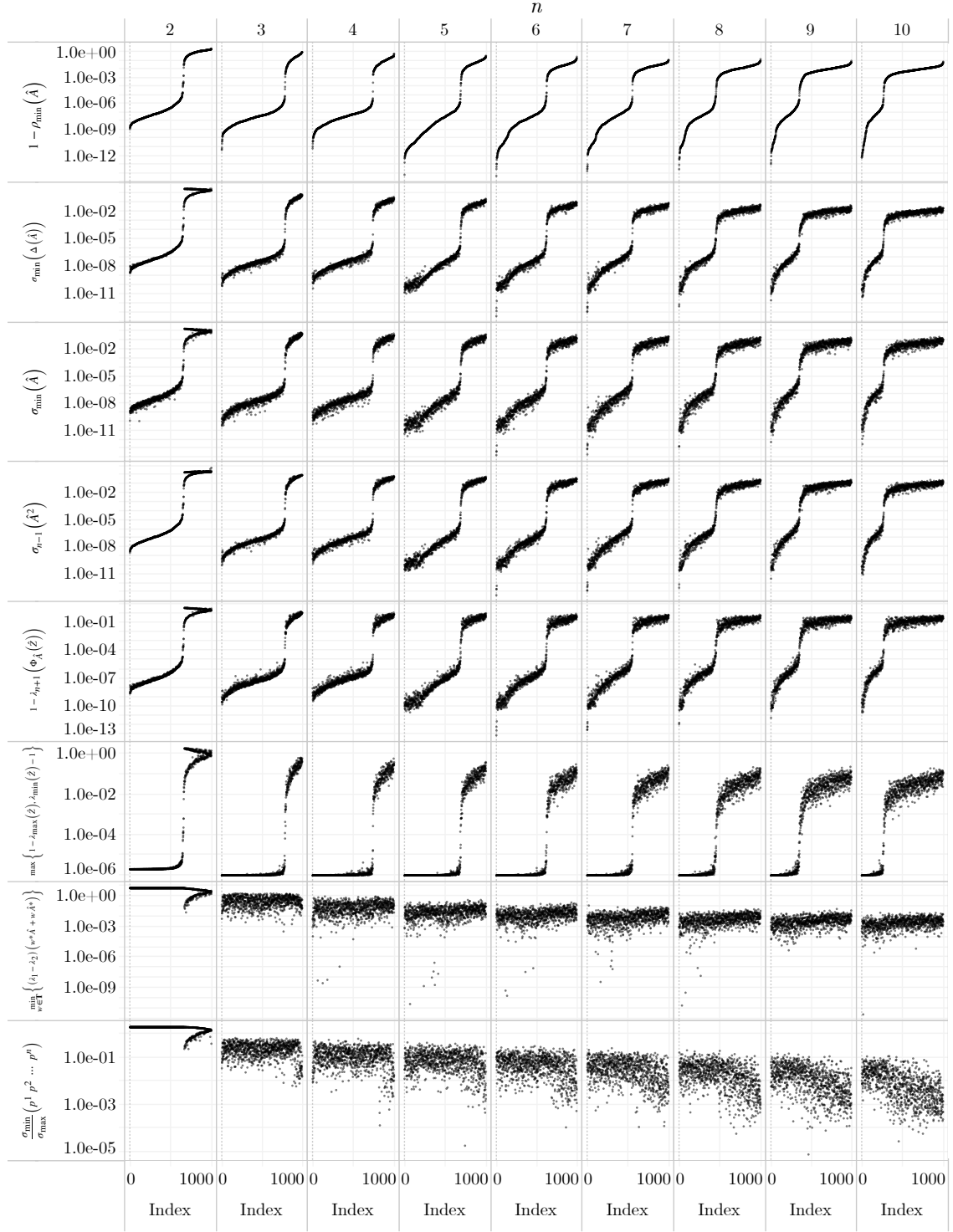
Figure 1: Disk and non-disk matrices

We note that normalizing the random input $Y \in \mathbf{M}^n$ encourages the output $\mathrm{prox}_r(Y)$ to be a disk matrix. All sufficiently small Y satisfy $\mathrm{prox}_r(Y) = 0$: specifically, this holds exactly when λY lies in the unit ball for the norm dual to the radius. On the other hand, since the set of disk matrices is a cone, the distance to this set for a scaled matrix αY is proportional to the scaling $\alpha > 0$: if α is large, $\mathrm{prox}_r(\alpha Y)$ will not be a disk matrix. The proportion of disk matrices seen depends on the balance between the normalization and the choice of the constant λ, which we fixed to be $\frac{3}{4}$ in the experiment.

3   The semidefinite program

Our calculation of the proximal map relies on the semidefinite representability of the numerical radius. We denote the space of n-by-n Hermitian matrices by $\mathbf{H}^n$, the cone of positive-semidefinite matrices by $\mathbf{H}^n_+$, and the corresponding semidefinite ordering by $\succeq$. We consider the eigenvalues of a matrix $X \in \mathbf{H}^n$, listed by multiplicity, in nonincreasing order,

$$\lambda_{\max}(X) = \lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X) = \lambda_{\min}(X),$$

and we use analogous notation for the singular values of an arbitrary matrix Y:

$$\sigma_j(Y) = \sqrt{\lambda_j(Y^*Y)}.$$

For a fixed matrix $A \in \mathbf{M}^n$, we define an affine map $\Phi_A \colon \mathbf{H}^n \to \mathbf{H}^{2n}$ by

$$\Phi_A(Z) \;=\; \begin{bmatrix} Z & A \\ A^* & -Z \end{bmatrix} \qquad \text{for } Z \in \mathbf{H}^n.$$

The following standard characterization of the numerical radius is due to Mathias [8], based in part on classical work of Ando [1].

Theorem 3.1 (SDP representation of the numerical radius)
The numerical radius of any matrix $A \in \mathbf{M}^n$ is the minimum value of the function $\lambda_{\max} \circ \Phi_A$, and the minimum is attained.

Any minimizer of the function $\lambda_{\max} \circ \Phi_A$ is called an *SDP-representor* for A.
    Using this result, we deduce that the minimum

$$(3.2) \qquad \min_{A \in \mathbf{M}^n, \; Z \in \mathbf{H}^n} \left\{ \lambda_{\max} \begin{bmatrix} Z & A \\ A^* & -Z \end{bmatrix} + \frac{1}{2}\|A - Y\|^2 \right\},$$

is attained, and at optimality, the matrix A must be $\mathrm{prox}_r(Y)$. Both the functions $\lambda_{\max}$ and $\|\cdot\|^2$ are semidefinite representable, so we can easily convert this minimization problem into an equivalent semidefinite program: in our numerical results, we solve it using the `cvx` package.

We can check the disk matrix property for any matrix, and in particular for the output of our experiment, using straightforward numerical linear algebra, rather than relying on the divergence measure (2.1) and the `chebfun` package. We apply the following result, which we derive later from work of Ando [1].

Theorem 3.3 *A matrix $A \in \mathbf{M}^n$ has field of values the unit disk if and only its numerical radius is one and there exists a nonzero polynomial n-vector $p(\cdot)$ of degree less than n satisfying*

$$(3.4) \qquad (w^*A + wA^*)p(w) \;=\; 2p(w) \quad \text{for all } w \in \mathbf{T}.$$

We call such a polynomial vector $p(\cdot)$ a *disk certificate*. A disk certificate for a general nonzero matrix $A \in \mathbf{M}^n$ is just a disk certificate for $\frac{1}{r(A)}A$. If the image $p(\mathbf{T})$ spans the space $\mathbf{C}^n$, then we call the certificate *spanning*.

To convert this result into a numerical tool, we identify any polynomial n-vector $p(\cdot)$ of degree less than n with its coefficient vectors:

$$(3.5) \qquad p(w) = \sum_{j=1}^{n} p^j w^{j-1} \qquad \text{for } p^j \in \mathbf{C}^n \quad (j = 1, 2, \ldots, n).$$

Certificates $p(\cdot)$ satisfying equation (3.4) then correspond to nontrivial solutions of the recursion

$$
\begin{aligned}
Ap^{k+1} + A^*p^{k-1} &= 2p^k \qquad (k = 0, 1, 2, \ldots, n+1) \\
p^{-1} = p^0 &= 0 = p^{n+1} = p^{n+2},
\end{aligned}
$$

which we can rewrite in matrix form,

$$
\Delta(A)
\begin{bmatrix}
p^1 \\ p^2 \\ p^3 \\ \vdots \\ p^n
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ \vdots \\ 0
\end{bmatrix},
$$

for the $n(n+2)$-by-n^2 matrix

$$(3.6) \qquad \Delta(A) \;=\;
\begin{bmatrix}
A & 0 & 0 & 0 & \cdots & 0 \\
-2I & A & 0 & 0 & \cdots & 0 \\
A^* & -2I & A & 0 & \cdots & 0 \\
0 & A^* & -2I & A & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & A^* & -2I & A \\
0 & 0 & \cdots & 0 & A^* & -2I \\
0 & 0 & \cdots & 0 & 0 & A^*
\end{bmatrix}.
$$

Thus, the disk property amounts simply to linear dependence of the columns of the matrix $\Delta(\frac{1}{r(A)}A)$ — the second row of Figure 1 plots the smallest singular value of this matrix. It is easy to see that a certificate $p(\cdot)$ is spanning exactly when its coefficient vectors p^j comprise a basis for $\mathbf{C}^n$, a property that we can write in terms of a condition measure:

$$(3.7) \qquad \frac{\sigma_{\min}}{\sigma_{\max}}(p^1 \ p^2 \cdots p^n) \ > \ 0.$$

The final row of Figure 1 plots this measure.

We call a matrix $A \in \mathbf{M}^n$ *separated* if the largest eigenvalue of the Hermitian matrix $w^*A + wA^*$ is simple for all complex units w, or in other words

$$(3.8) \qquad \min_{w \in \mathbf{T}}(\lambda_1 - \lambda_2)(w^*A + wA^*) \ > \ 0.$$

In that case, any disk certificate must be unique, up to nonzero scalar multiplication, or equivalently, the matrix on the left-hand side has a one-dimensional null space. The penultimate row of Figure 1 plots this measure. We call separated matrices with spanning certificates *strongly certified*.

As we already observed, the output of our experiment, $A = \text{prox}_r(Y)$ is often a disk matrix. In those instances, we can determine whether A is strongly certified by checking the conditions (3.7) and (3.8). Indeed, we find, numerically, that the disk matrices generated are always strongly certified.

4 Disk certificates

We begin our more formal development by justifying our definition of a disk certificate. For any matrix $A \in \mathbf{M}^n$ and complex unit $w \in \mathbf{T}$, we have

$$2\max\{\text{Re}(w^*z) : z \in W(A)\} \ = \ \lambda_{\max}(w^*A + wA^*).$$

Hence A is a disk matrix if and only if, for all $w \in \mathbf{T}$, there exists a nonzero vector $p(w) \in \mathbf{C}^n$ satisfying

$$(4.1) \qquad (w^*A + wA^*)p(w) \ = \ 2r(A)p(w).$$

However, using two key results, due to Ando and Crouzeix, which we summarize next, we can restrict attention to vectors $p(w)$ satisfying equation (4.1) whose components are polynomials in w of degree less than n. To see this, we rely on the following restatement of [1, Theorem 1].

Theorem 4.2 *A matrix $A \in \mathbf{M}^n$ has numerical radius $r(A) \leq 1$ if and only if it has an* **Ando representation** *$A = 2SUC$, for matrices $S, C \in \mathbf{H}_+^n$ satisfying $S^2 + C^2 = I$, and a unitary matrix U.*

We deduce the following result, subsuming Theorem 3.3.

Theorem 4.3 *Suppose that a matrix $A \in \mathbf{M}^n$ with unit numerical radius has Ando representation $A = 2SUC$, as in Theorem 4.2. Then disk certificates for A are exactly nonzero polynomial n-vectors p of degree less than n satisfying*

(4.4) $$(UC - wS)p(w) = 0 \ \ for \ all \ w \in \mathbf{C}.$$

Furthermore, the following properties are equivalent:

(i) *The matrix $UC - wS$ is singular for all $w \in \mathbf{C}$;*

(ii) *A has a disk certificate;*

(iii) *A is a disk matrix.*

Proof   The first claim follows directly from [7, Proposition 8.2]. The equivalence of properties (i) and (iii) follows [3, Lemma 10.1]. Clearly the existence of a disk certificate implies the singularity property (i). Conversely, if the singularity property (i) holds then the matrix pencil $UC - wS$ is singular. The Kronecker form for singular matrix pencils [10, pp. 106–108] now guarantees the existence of a nonzero polynomial n-vector p of degree less than n satisfying equation (4.4), and as we have already seen, this p is a disk certificate. □

By rescaling, we deduce the justification for our disk certificate definition.

Corollary 4.5 *A matrix is a disk matrix if and only if it has a disk certificate.*

To consider more specifically the set of strongly certified matrices, we begin with the following easy result.

Proposition 4.6 *In the space $\mathbf{M}^n$, the set of separated matrices is open.*

Proof   Consider a sequence of non-separated matrices $A_k \in \mathbf{M}^n$, for $k = 1, 2, 3, \ldots$, with limit $A \in \mathbf{M}^n$. By assumption, there exist corresponding complex units $w_k \in \mathbf{C}$ such that the largest two eigenvalues of the matrix $w_k^* A_k + w_k A_k^*$ are equal. Taking a subsequence, we can suppose that w_k converges to some complex unit $w \in \mathbf{C}$. By continuity, the largest two eigenvalues of the matrix $w^* A + w A^*$ are also equal, so the result follows. □

Theorem 4.7 *Within the set of all disk matrices in $\mathbf{M}^n$, the set of strongly certified disk matrices is relatively open.*

Proof   Consider a strongly certified disk matrix $A \in \mathbf{M}^n$. We claim that all nearby disk matrices are strongly certified. If not, then there exists a sequence of disk matrices $A_k \to A$, each of which is not not strongly certified. By the previous result, each A_k is separated, so its disk certificate p_k (unique up to scalar multiplication) is not spanning. By scaling, we can suppose that each certificate has unit norm. (Here, we define the norm of any polynomial vector p, as in equation (3.5), as the Frobenius norm of the matrix with columns its coefficients p^j.) Taking a subsequence, we can suppose that p_k converges to some nonzero limit polynomial p, which by continuity must be the unique certificate for A, and furthermore is not spanning. This contradiction completes the proof.   □

Example 4.8 (A strongly certified disk matrix)   The matrix

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

has field of values the unit disk, since it has the Ando representation

$$2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

and the polynomial vector

$$\begin{bmatrix} 1 \\ w \end{bmatrix}$$

is a spanning disk certificate. Furthermore, the matrix A is separated, since the matrix $w^*A + wA^*$ has the distinct eigenvalues ± 2, so A is strongly certified.

Example 4.9 (A disk matrix with no spanning certificate)   The matrix

$$A = \begin{bmatrix} 0 & 0 & 2 \\ 0 & \frac{4}{5} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

has field of values the unit disk, since it has the Ando representation

$$2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ \frac{3}{5} & \frac{4}{5} & 0 \\ \frac{4}{5} & -\frac{3}{5} & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and the polynomial vector

$$\begin{bmatrix} 1 \\ 0 \\ w \end{bmatrix}$$

is a disk certificate. However, a direct application of Theorem 4.3 shows that any disk certificate for A must have second component zero, and hence A has no spanning certificate.

9

Singular matrices need not be disk matrices, and as Example 4.9 highlights, disk matrices need not be strongly certified. However, no such examples emerge from our computational experiment: whenever the output matrix A is singular, it is a strongly certified disk matrix. The numerical radius must be partly smooth around such matrices, as a consequence of the following result, which, in the light of Theorem 4.7, is just a restatement of [7, Theorem 7.4].

Theorem 4.10 *The set of all strongly certified disk matrices in $\mathbf{M}^n$ is a real-analytic manifold of real codimension $2n$, with respect to which the numerical radius is partly smooth.*

5 Random disk matrices and SDP-representors

For any matrix $A \in \mathbf{M}^n$, Theorem 3.1 ensures the existence of an SDP-representor, or in other words a minimizer of the function $\lambda_{\max} \circ \Phi_A$. Our computational experiment (3.2) suggests a variety of interesting properties of SDP-representors. Each random instance solves a single semidefinite program which, by decomposing into two subproblems, we can view as an experiment on random disk matrices:

- Compute $A = \mathrm{prox}_r(Y)$ for a random input matrix $Y \in \mathbf{M}^n$;

- Minimize $\lambda_{\max} \circ \Phi_A$ to find an SDP-representor Z for A.

We note that the SDP-representor Z may not be unique.

In the experiment, the following properties often hold. Furthermore, in each random instance, with occasional numerical anomalies, either all the properties hold, or none do.

  (i) A is singular.

 (ii) A^2 has rank strictly less than $n - 1$.

(iii) $\lambda_{\max}(Z) = r(A)$ and $\lambda_{\min}(Z) = -r(A)$.

 (iv) A is a disk matrix.

  (v) The largest eigenvalue of $\Phi_A(Z)$ has multiplicity strictly larger than n.

 (vi) A is a strongly certified disk matrix.

In light of property (v), we say that a matrix A has the *multiplicity property* if there exists an SDP-representor Z (not necessarily unique) such that the largest eigenvalue of $\Phi_A(Z)$ has multiplicity strictly larger than n.

We can quantify each of the first five properties with a corresponding nonnegative measure that vanishes exactly when the property holds:

- $\sigma_{\min}(A)$

- $\sigma_{n-1}(A^2)$

- $\max\left\{r(A) - \lambda_{\max}(Z), \lambda_{\min}(Z) - r(A)\right\}$

- $\sigma_{\min}\left(\Delta\left(\frac{1}{r(A)}A\right)\right)$

- $r(A) - \lambda_{n+1}\left(\Phi_A(Z)\right).$

We verify the strong certification (property (vi)) by checking the conditions (3.7) and (3.8) in conjunction with property (iv). To illustrate the coincidence of the six properties, for each matrix size $n = 1, 2, 3, \ldots$, we order the random instances in increasing order of the disk-property divergence measure (2.1), and then simultaneously plot the measures, scaled appropriately, in Figure 1.

Of course, for *arbitrary* matrices A, the six properties are far from equivalent. We shall see (in Corollary 5.5) the general implications

(5.1) $$\text{(iv)} \;\Rightarrow\; \text{(iii)} \;\Rightarrow\; \text{(ii)} \;\Rightarrow\; \text{(i)}.$$

Furthermore, when the SDP-representor Z is unique, we shall see (in Theorems 6.1 and 6.8) the implications

$$\text{(vi)} \;\Rightarrow\; \text{(v)} \;\Rightarrow\; \text{(iv)}.$$

On the other hand, examples show the relationships

$$\text{(i)} \;\not\Rightarrow\; \text{(ii)} \;\not\Rightarrow\; \text{(iii)} \;\not\Rightarrow\; \text{(iv)} \;\not\Rightarrow\; \text{(v)}.$$

The diagonal matrix $A = \mathrm{Diag}(0, 1)$ illustrates the first relationship, and a direct calculation show that the diagonal matrices $A = \mathrm{Diag}(0, 0, 1)$ and $Z = \mathrm{Diag}(1, 1, 0)$ illustrate the second. For the third relationship, we present an example after Corollary 5.5, and Example 6.4 illustrates the fourth. The precise relationship between the multiplicity property and strong certification remains unclear.

To explore these properties, we begin by characterizing SDP-representors. We first consider matrices with unit numerical radius: the general case follows easily by scaling.

Theorem 5.2 *A matrix $Z \in \mathbf{H}^n$ is an SDP-representor for a matrix $A \in \mathbf{M}^n$ with unit numerical radius if and only if*

(5.3) $I_n \succeq Z \succeq -I_n \quad and \quad A = (I - Z)^{\frac{1}{2}} U (I + Z)^{\frac{1}{2}}$ *for a contraction $U \in \mathbf{M}^n$.*

Furthermore, there exist such matrices Z and U, with U in fact unitary.

Proof By Theorem 3.1, a matrix Z is an SDP-representor for A if and only if the matrix $I_{2n} - \Phi_A(Z)$ is positive semidefinite, and this is equivalent to the property (5.3) by [6, Theorem 7.7.9]. The final result is proved in [1]. $\qquad\square$

The proof of the following result essentially follows [3, Lemma 10.1].

Theorem 5.4 *If a matrix $A \in \mathbf{M}^n$ with field of values the unit disk satisfies the property (5.3), then the matrix*

$$U(I + Z)^{\frac{1}{2}} - w(I - Z)^{\frac{1}{2}}$$

is singular for all $w \in \mathbf{C}$.

Proof Consider the two positive semidefinite Hermitian matrices $S = (I - Z)^{\frac{1}{2}}$ and $C = (I + Z)^{\frac{1}{2}}$. By assumption, for any $w \in \mathbf{C}$ satisfying $|w| = 1$ there exists a unit vector $v \in \mathbf{C}^n$ satisfying $v^*Av = w$. We now observe the inequalities

$$1 = |v^*Av| = 2|v^*SUCv| \le 2\|Sv\|\|UCv\| \le 2\|Sv\|\|Cv\| \le \|Sv\|^2 + \|Cv\|^2 = 1,$$

Equality in the first inequality ensures $UCv = w'Sv$ for some $w' \in \mathbf{C}$. Equality in the last inequality, along with the final equality, ensures $\|Sv\| = \frac{1}{\sqrt{2}} = \|Cv\|$. We deduce

$$w = v^*Av = 2v^*SUCv = 2v^*Sw'Sv = w',$$

so $UCv = wSv$, and hence the matrix $UC - wS$ is singular. The result follows. $\quad\square$

We deduce the implications (5.1), which we summarize in the following result.

Corollary 5.5 *If a matrix $A \in \mathbf{M}^n$ has an SDP-representor with both the eigenvalues $\pm r(A)$, as holds in particular if A is a disk matrix, then $\mathrm{rank}(A^2) < n - 1$, and hence A is singular.*

Proof By scaling, we can assume $r(A) = 1$. By Theorem 5.2, for any SDP-representor Z, there exists a contraction $U \in \mathbf{M}^n$ satisfying the property (5.3). We deduce

$$A^2 = (I - Z)^{\frac{1}{2}}U(I - Z^2)^{\frac{1}{2}}U(I + Z)^{\frac{1}{2}},$$

and the rank condition now follows from the fact that $\mathrm{rank}(I - Z^2) < n - 1$. In the case when A is a disk matrix, we apply Theorem 5.4, and consider the cases $w = 0$ and $w \to \infty$. $\qquad\square$

The converse of Corollary 5.5 is false. For example, it is easy to verify that if we define

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

then A has unit numerical radius, Z is an SDP-representor with eigenvalues ± 1, and yet A is not a disk matrix.

Proposition 5.6 *If matrix $A \in \mathbf{M}^n$ with unit numerical radius has Ando representation $A = 2SUC$, as in Theorem 4.2, then the matrix*

$$Z = I - 2S^2 = 2C^2 - I$$

is an SDP-representor, and the largest eigenvalue of the matrix $\Phi_A(Z)$ has multiplicity at least n.

Proof The matrix

$$(5.7) \qquad B = I - \Phi_A(Z) = 2 \begin{bmatrix} S \\ -CU^* \end{bmatrix} [S \quad -UC]$$

is positive semidefinite. We deduce the inequality $\lambda_{\max}\big(\Phi_A(Z)\big) \leq 1$, and in fact equality must hold, by Theorem 3.1, and Z must be an SDP-representor. Furthermore, the multiplicity of the largest eigenvalue of $\Phi_A(Z)$ is $2n - \operatorname{rank} B$. Since the rank of any matrix Y coincides with the rank of the matrix Y^*Y, we see $\operatorname{rank} B \leq n$, and the result now follows. $\qquad\square$

Turning to the multiplicity property, we first note that the following corollary of the previous result.

Corollary 5.8 *Any matrix $A \in \mathbf{M}^n$ has an SDP-representor $Z \in \mathbf{H}^n$ for which the largest eigenvalue of $\Phi_A(Z)$ has multiplicity at least n.*

Consider, for example, Example 4.8. Any SDP-representor $Z \in \mathbf{H}^2$ has both the eigenvalues ± 1, by Corollary 5.5, and hence $Z^2 = I$. By Theorem 5.2, there exists a contraction $U \in \mathbf{M}^2$ such that

$$A = (I - Z)^{\frac{1}{2}} U (I + Z)^{\frac{1}{2}},$$

so

$$A(I - Z) = (I - Z)^{\frac{1}{2}} U (I - Z^2)^{\frac{1}{2}} (I - Z)^{\frac{1}{2}} = 0.$$

Hence the matrix $I - Z$ must be diagonal, whence so is Z. We deduce that the unique SDP-representor is the matrix

$$Z = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In that case, the matrix

$$\Phi_A(Z) = \begin{bmatrix} -1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & -1 \end{bmatrix}$$

has eigenvalues (by multiplicity) 1, 1, 1, and -3, illustrating property (v).

Question 5.9 Must SDP-representors of strongly certified disk matrices be unique?

6 Unique SDP-representors

As we noted in the previous section, the matrices A generated by our random experiment are disk matrices exactly when they satisfy the multiplicity property. In this section we explain this phenomenon, at least in the case when the SDP-representor is unique. We begin with the following result.

Theorem 6.1 (Sufficient condition for disk matrices)   *Consider a matrix $A \in \mathbf{M}^n$ with a unique SDP-representor $Z \in \mathbf{H}^n$. If A has the multiplicity property, then it must be a disk matrix.*

Proof   After scaling, we may assume $r(A) = 1$. Given an Ando representation $A = 2SUC$, as in Theorem 4.2, Proposition 5.6 implies $Z = I - 2S^2$. Using the multiplicity property, and following the argument of Proposition 5.6, we deduce, for all numbers $w \in \mathbf{C}$,

$$
\begin{aligned}
n \; > \; \mathrm{rank}\big(I - \Phi_A(Z)\big) \; &= \; \mathrm{rank}[S \;\; -UC] \\
\geq \; \mathrm{rank}\left([S \;\; -UC] \begin{bmatrix} wI \\ I \end{bmatrix} \right) \; &= \; \mathrm{rank}(UC - wS),
\end{aligned}
$$

so the matrix $UC - wS$ is singular. The result now follows by Theorem 4.2.   □

Question 6.2 Is the uniqueness assumption necessary? In other words, for a nondisk matrix A with an SDP-representor Z, can the largest eigenvalue of the matrix $\Phi_A(Z)$ have multiplicity exceeding n?

Question 6.3 Is the converse true? That is, if a disk matrix A has a unique SDP-representor Z, can the largest eigenvalue of the matrix $\Phi_A(Z)$ have multiplicity n?

Routine algebra, and the classification in [7] verifies the converse of Theorem 6.1 in the 2-by-2 and 3-by-3 cases. On the other hand, it can fail without the uniqueness assumption, as shown by the following example.

Example 6.4 (A disk matrix with many SDP-representors)
Consider the disk matrix A of Example 4.9, which has no strong (or indeed spanning) certificate. For any real number s, consider the matrix

$$
Z \; = \; \begin{bmatrix} -1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

The matrix

$$\Phi_A(Z) = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 2 \\ 0 & s & 0 & 0 & \frac{4}{5} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{4}{5} & 0 & 0 & -s & 0 \\ 2 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

is unitarily similar (by row and column permutations) to

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \oplus \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \oplus \begin{bmatrix} s & \frac{4}{5} \\ \frac{4}{5} & -s \end{bmatrix},$$

and hence has eigenvalues

$$1, \ 1, \ 1, -3, \ \pm \sqrt{s^2 + \left(\frac{4}{5}\right)^2}.$$

Thus Z is an SDP-representor providing $|s| \leq \frac{3}{5}$, and if $|s| < \frac{3}{5}$, then the multiplicity of the largest eigenvalue of $\Phi_A(Z)$ is three.

On the other hand, in this example, setting $s = \pm\frac{3}{5}$ does produce an SDP-representor Z for which the multiplicity (four) of the largest eigenvalue of the matrix $\Phi_A(Z)$ exceeds $n = 3$. This suggests the following question, more general than Question 6.3.

Question 6.5 Do all disk matrices A satisfy the multiplicity property?

Again, routine algebra and [7] answers this question affimatively in the 2-by-2 and 3-by-3 cases.

Unlike the disk matrix A of Examples 4.9 and 6.4, the disk matrices produced by our random experiment are always strongly certified. In that case we have a more positive result, as we argue next. We begin with an elementary tool.

Lemma 6.6 *The components of polynomial k-vector $q(w)$ (in the variable $w \in \mathbf{C}$) are linearly independent over $\mathbf{C}$ if and only its range $q(\mathbf{C})$ spans the complex vector space $\mathbf{C}^k$.*

Proof The conclusion fails if and only if there exists a nonzero vector $a \in \mathbf{C}^k$ satisfying $\sum_j a_j q_j(w) = 0$ for all $w \in \mathbf{C}$, so the result follows. $\square$

This result extends, as follows.

Lemma 6.7 *The dimension in the complex vector space $\mathbf{C}^k$ of the range of a polynomial k-vector $q(w)$ (in the variable $w \in \mathbf{C}$) equals the dimension of the span over $\mathbf{C}$ of its component polynomials.*

Proof Denote by d the second dimension. Denote by q^d and q^{k-d} the polynomial vectors obtained from q by deleting the last $k - d$ and first d components respectively. Without loss of generality, we can suppose that q^d has linearly independent components and that $q^{k-d} = Eq^d$ for some complex matrix E. By Lemma 6.6, the range $q^d(\mathbf{C})$ has dimension d in the space $\mathbf{C}^d$, and the range $q(\mathbf{C})$ is its image under the injective linear map $\binom{I}{E}$, so has the same dimension. □

Theorem 6.8 (Necessary condition for disk matrices) *If a nonzero matrix $A \in \mathbf{M}^n$ has a spanning disk certificate (or in particular is strongly certified), then it satisfies the multiplicity property.*

Proof After scaling, we can assume that the matrix A has unit numerical radius and an Ando representation $A = 2SUC$, as in Theorem 4.2, with a spanning disk certificate p. By Proposition 5.6, the matrix $Z = I - 2S^2$ is an SDP-representor, and by equation (5.7) the vector

$$q(w) = \left[\begin{array}{c} wp(w) \\ p(w) \end{array} \right]$$

is an eigenvector corresponding to the largest eigenvalue (namely one) of the matrix $\Phi_A(Z)$ for all $w \in \mathbf{C}$. By assumption, the n components of the vector polynomial p are linearly independent. The vector polynomial q must therefore have a linearly independent set of more than n components, since at last one component has degree strictly larger that any component of p. By Lemma 6.7, the matrix $\Phi_A(Z)$ therefore has a linearly independent set of more than n eigenvectors corresponding to the eigenvalue 1, so this eigenvalue has multiplicity larger than n. □

To conclude, in the case of a unique SDP-representor Z, Theorems 6.1 and 6.8 together indicate a strong correlation between the disk property for matrices A and the multiplicity property, exactly as we noted empirically.

References

[1] T. Ando. Structure of operators with numerical radius one. *Acta Sci. Math. (Szeged)*, 34:11–15, 1973.

[2] B. Collins, Pi. Gawron, A.E. Litvak, and K. Życzkowski. Numerical range for random matrices. *J. Math. Anal. Appl.*, 418:516–533, 2014.

[3] M. Crouzeix. Some constants related to numerical ranges. *SIAM J. Matrix Anal. Appl.*, 37(1):420–442, 2016.

[4] T.A. Driscoll, N. Hale, and L.N. Trefethen. *Chebfun Guide*. Pafnuty Publications, Oxford, 2014.

[5] Leslie Hogben, editor. *Handbook of Linear Algebra.* Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, second edition, 2014.

[6] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, second edition, 2013.

[7] A.S. Lewis and M.L. Overton. Partial smoothness of the numerical radius at matrices whose fields of values are disks. `arXiv:1901.00050`, 2019.

[8] R. Mathias. Matrix completions, norms and Hadamard products. *Proc. Amer. Math. Soc.*, 117(4):905–918, 1993.

[9] Bit-Shun Tam and Shangjun Yang. On matrices whose numerical ranges have circular or weak circular symmetry. *Linear Algebra Appl.*, 302/303:193–221, 1999. Special issue dedicated to Hans Schneider (Madison, WI, 1998).

[10] P. Van Dooren. The computation of Kronecker's canonical form of a singular pencil. *Linear Algebra Appl.*, 27:103–140, 1979.
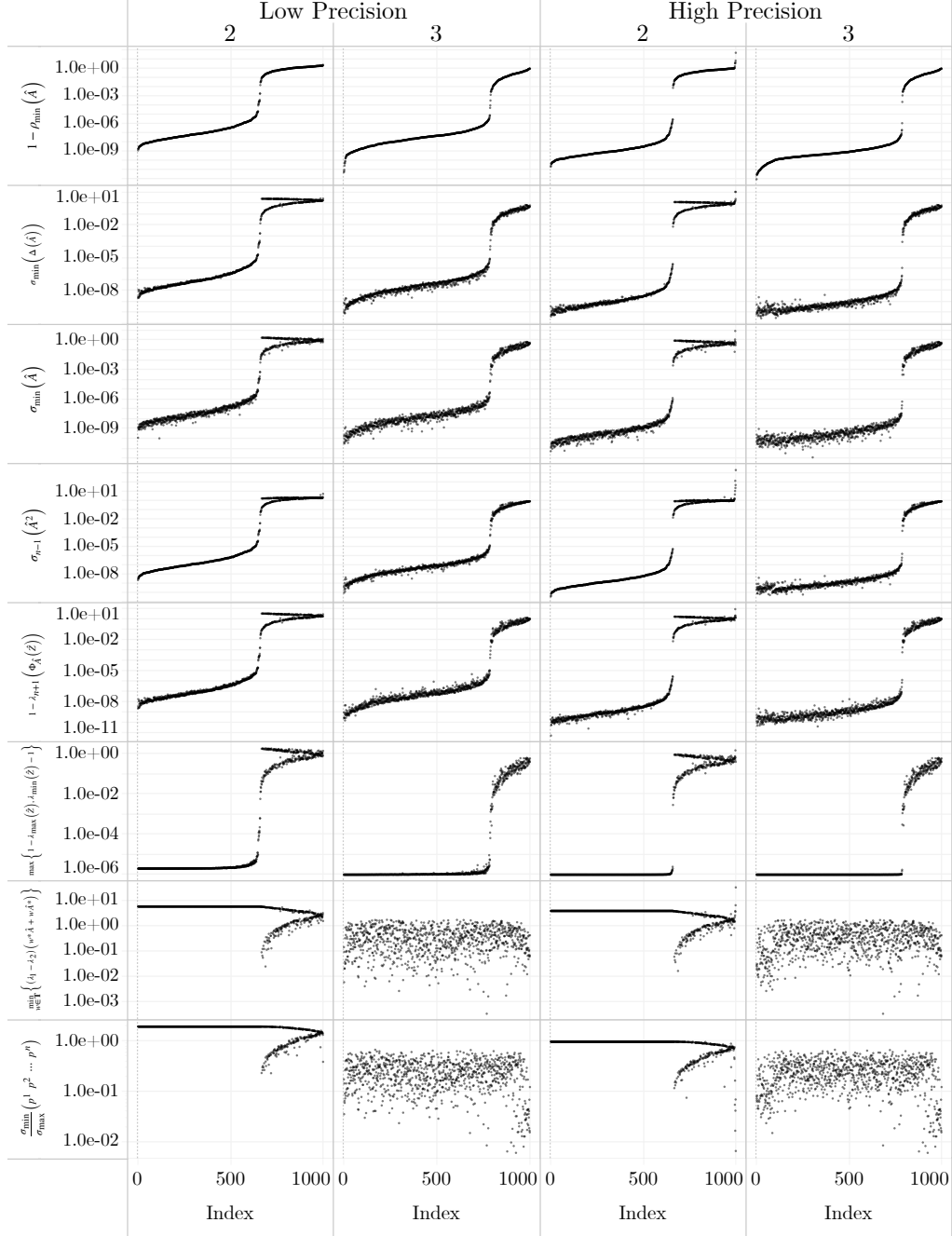
7 Appendix

Figure 2: Disk and non-disk matrices in higher precision