FAST SINGULAR VALUE DECAY FOR LYAPUNOV SOLUTIONS WITH NONNORMAL COEFFICIENTS[*]

JONATHAN BAKER[†], MARK EMBREE[‡], AND JOHN SABINO[§]

Abstract. Lyapunov equations with low-rank right-hand sides often have solutions whose singular values decay rapidly, enabling iterative methods that produce low-rank approximate solutions. All previously known bounds on this decay involve quantities that depend quadratically on the departure of the coefficient matrix from normality: these bounds suggest that the larger the departure from normality, the slower the singular values will decay. We show this is only true up to a threshold, beyond which a larger departure from normality can actually correspond to *faster* decay of singular values: if the singular values decay slowly, the numerical range cannot extend far into the right half-plane.

Key words. Lyapunov equation, singular values, numerical range, nonnormality

AMS subject classifications. 15A18, 15A60, 65F30, 93B05, 93B40

DOI. 10.1137/140993867

1. Introduction. Lyapunov equations of the form

$$(1.1) \qquad \mathbf{AX} + \mathbf{XA}^* = -\mathbf{BB}^*$$

arise from the study of the controllability and observability of linear time-invariant dynamical systems and subsequently in balanced truncation model order reduction [1, 27]. In this setting, the right-hand side $-\mathbf{BB}^*$ often has low rank (equal to the number of inputs or outputs in the system). If the eigenvalues of $\mathbf{A} \in \mathbb{C}^{n \times n}$ are in the left half of the complex plane and $(\mathbf{A}, \mathbf{B})$ is controllable, the solution $\mathbf{X} \in \mathbb{C}^{n \times n}$ is Hermitian positive definite, i.e., $\mathrm{rank}(\mathbf{X}) = n$ [27, sect. 3.8]. Even when the coefficient matrix $\mathbf{A}$ is sparse, $\mathbf{X}$ is typically dense: so for large-scale problems one cannot afford to store all n^2 entries of the solution.

Penzl observed that when the right-hand side of (1.1) has low rank, the singular values $s_1 \geq s_2 \geq \cdots \geq s_n > 0$ of $\mathbf{X}$ often decay exponentially [17]; e.g., $s_k/s_1 \leq C\gamma^k$ for some constants $C > 0$ and $\gamma \in (0, 1)$. This fact now enables numerous iterative methods that seek accurate low-rank approximations to $\mathbf{X}$; see [4, 20] for recent surveys. Since the singular values of $\mathbf{X}$ bound the best possible performance of iterative methods for solving Lyapunov equations, understanding how these singular values depend on the coefficient matrix $\mathbf{A}$ is important. Of course, since $\mathbf{X}$ is Hermitian positive definite, its singular values equal its eigenvalues; it is customary to refer to singular values because (a) we seek low-rank approximations to $\mathbf{X}$, and (b) much of the related analysis generalizes to Sylvester equations, where $\mathbf{X}$ need not even be square. We shall thus always speak of the *singular values of* $\mathbf{X}$, $s_1 \geq s_2 \geq \cdots \geq s_n > 0$, and the *eigenvalues of* $\mathbf{A}$, $\lambda_1, \ldots, \lambda_n$, which comprise the spectrum $\sigma(\mathbf{A})$. Let $\mathbb{C}^-$ and $\mathbb{C}^+$ denote the open left and right halves of the complex plane, and let $\|\cdot\|$ de-

---

[†]Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005-1892 (jpb7@rice.edu).
[‡]Department of Mathematics, Virginia Tech, Blacksburg, VA 24061 (embree@vt.edu).
[§]The Boeing Company, Tukwila, WA 98108-4040 (john-paul.n.sabino@boeing.com).

note the vector 2-norm and the matrix norm it induces. We assume $\mathbf{A}$ is stable, i.e., $\sigma(\mathbf{A}) \subset \mathbb{C}^-$.

The singular values of $\mathbf{X}$ depend on spectral properties of $\mathbf{A}$; grossly speaking, they decay more rapidly the farther $\sigma(\mathbf{A})$ falls in the left half of the complex plane and more slowly as eigenvalues of $\mathbf{A}$ grow in imaginary part. But eigenvalues alone cannot explain the singular values of $\mathbf{X}$. Penzl showed that for any desired singular values of $\mathbf{X}$, one can construct a corresponding $\mathbf{A}$ with *any spectrum* in the left half-plane (for some special choice of $\mathbf{B}$) [18]. Now suppose $\sigma(\mathbf{A})$ is fixed. Recall that $\mathbf{A}$ is *normal* if it commutes with its adjoint ($\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$) or, equivalently, if eigenvectors give an orthonormal basis for $\mathbb{C}^n$. We shall use the term *departure from normality* generically; many different scalar measures of nonnormality have been shown to be essentially equivalent [10]. All previously known bounds suggest that the singular values of $\mathbf{X}$ will decay more slowly as the departure of $\mathbf{A}$ from normality increases, and it is this particular point that concerns us here. In section 2 we describe the variety of bounds that have been proposed in the literature, highlighting how they treat the nonnormality of $\mathbf{A}$. Section 3 gives a simple 2×2 example that clearly illustrates that, in contrast to previously known bounds, beyond a certain threshold a larger departure from normality can actually give singular values that *decay more quickly.* We offer an intuitive explanation for this behavior in section 4 and then prove a decay bound that incorporates this effect in section 5: *the trailing singular values must be small if eigenvalues of the Hermitian part of $\mathbf{A}$ fall far in the right half-plane.*

2. Decay bounds and their inadequacy for nonnormal coefficients. One approach to proving the decay of the singular values of $\mathbf{X}$ uses the low-rank approximations constructed by the ADI algorithm; see, e.g., [1, 8, 11]. Suppose $\text{rank}(\mathbf{B}) = r$. The kth ADI iteration gives an approximate solution $\mathbf{X}_k$ with $\text{rank}(\mathbf{X}_k) \leq kr$ that satisfies

$$\mathbf{X} - \mathbf{X}_k = \phi_k(\mathbf{A})\mathbf{X}\phi_k(\mathbf{A})^*,$$

where

$$\phi_k(z) = \prod_{j=1}^{k} \frac{z + \mu_j}{z - \overline{\mu_j}}$$

is a rational function whose parameters, the *shifts* $\{\mu_j\} \subset \mathbb{C}^+$, are picked from the right half-plane to minimize $\|\mathbf{X} - \mathbf{X}_k\|$. By the optimality of the singular values (the Schmidt–Eckart–Young–Mirsky theorem [1, Thm. 3.6]),

$$(2.1) \qquad \frac{s_{kr+1}}{s_1} \leq \frac{\|\mathbf{X} - \mathbf{X}_k\|}{\|\mathbf{X}\|} \leq \|\phi_k(\mathbf{A})\|\|\phi_k(\mathbf{A})^*\| = \|\phi_k(\mathbf{A})\|^2.$$

Bounds on the singular values of $\mathbf{X}$ then follow by approximating norms of functions of $\mathbf{A}$. Any specific choice of rational function ϕ_k gives an upper bound, and much theoretical and practical work has addressed the selection of optimal $\{\mu_j\}$ parameters. Since our main point does not depend on the choice of ϕ_k, we shall not dwell on that issue here. Our goal is to illustrate that all known bounds on the singular values of $\mathbf{X}$ fail to capture the diverse behavior possible for nonnormal $\mathbf{A}$, so we shall briefly describe the different approaches taken in the literature. If $\mathbf{A}$ is normal, then

$$(2.2) \qquad \|\phi_k(\mathbf{A})\| = \max_{\lambda \in \sigma(\mathbf{A})} |\phi_k(\lambda)|,$$

but for nonnormal $\mathbf{A}$, the left-hand side of (2.2) can be considerably larger than the right-hand side. There are three common ways to bound $\|\phi_k(\mathbf{A})\|$ (cf. [13, sect. 4.11]), each of which then leads to an upper bound on (2.1).

- Eigenvalues: If $\mathbf{A}$ is diagonalizable, $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$, then

$$(2.3) \qquad \|\phi_k(\mathbf{A})\| \leq \|\mathbf{V}\|\|\mathbf{V}^{-1}\| \max_{\lambda \in \sigma(\mathbf{A})} |\phi_k(\lambda)|.$$

Combining (2.3) with (2.1) gives

$$(2.4) \qquad \frac{s_{kr+1}}{s_1} \leq \|\mathbf{V}\|^2\|\mathbf{V}^{-1}\|^2 \max_{\lambda \in \sigma(\mathbf{A})} \prod_{j=1}^{k} \frac{|\lambda + \mu_j|^2}{|\lambda - \overline{\mu_j}|^2}.$$

This bound was first written down for general diagonalizable $\mathbf{A}$ by Sorensen and Zhou [21, Thm. 2.1], based on earlier work on the Hermitian case by Penzl [18]. In that Hermitian case, several concrete bounds were obtained by selecting particular real shifts, $\{\mu_j\}$: using suboptimal shifts, Penzl gave an elegant bound [18, Thm. 1], which was improved using optimal shifts for a real interval in [19, Thm. 2.1.1].

When $\mathbf{A}$ is non-Hermitian and the eigenvector matrix is ill-conditioned, $\|\mathbf{V}\|\|\mathbf{V}^{-1}\| \gg 1$, one might improve upon (2.4) by posing the maximization problem on larger subsets of $\mathbb{C}$ that permit constants smaller than $\|\mathbf{V}\|\|\mathbf{V}^{-1}\|$. We consider two such methods next.

- Numerical range: If ϕ_k is analytic on the numerical range [14, Chap. 1]

$$W(\mathbf{A}) := \{\mathbf{x}^*\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = 1\},$$

then

$$(2.5) \qquad \|\phi_k(\mathbf{A})\| \leq C \max_{z \in W(\mathbf{A})} |\phi_k(z)|,$$

where C denotes Crouzeix's constant, $C \in [2, 11.08]$ [6]. Combining this bound with (2.1) gives

$$(2.6) \qquad \frac{s_{kr+1}}{s_1} \leq C^2 \max_{z \in W(\mathbf{A})} \prod_{j=1}^{k} \frac{|z + \mu_j|^2}{|z - \overline{\mu_j}|^2}.$$

This bound only holds when ϕ_k is analytic on $W(\mathbf{A})$, so, in particular, $\overline{\mu_j} \notin W(\mathbf{A})$. Since $\mu_j \in \mathbb{C}^+$, a sufficient condition to ensure analyticity is that $W(\mathbf{A}) \subseteq \mathbb{C}^-$. The rightmost extent of $W(\mathbf{A})$ in the complex plane plays an important role in analysis of dynamical systems. This value is called the *numerical abscissa,*

$$(2.7) \qquad \omega(\mathbf{A}) = \max_{z \in W(\mathbf{A})} \text{Re}\, z,$$

and it equals the rightmost eigenvalue of the Hermitian part of $\mathbf{A}$:

$$\omega(\mathbf{A}) = \max_{z \in W(\mathbf{A})} \frac{z + \overline{z}}{2} = \max_{\substack{\mathbf{x} \in \mathbb{C}^n \\ \|\mathbf{x}\|=1}} \mathbf{x}^*\left(\frac{\mathbf{A} + \mathbf{A}^*}{2}\right)\mathbf{x} = \max\left\{\lambda : \lambda \in \sigma\left(\frac{\mathbf{A} + \mathbf{A}^*}{2}\right)\right\}.$$

Notice that $\omega(\mathbf{A})$ can be positive even when the spectrum of $\mathbf{A}$ is in the left half-plane and that $|\omega(\mathbf{A})| \leq \|\mathbf{A}\|$. The numerical abscissa describes the small t behavior of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ with $\mathbf{x}(0) = \mathbf{x}_0$:

$$\max_{\substack{\mathbf{x}_0 \in \mathbb{C}^n \\ \|\mathbf{x}_0\|=1}} \frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{x}(t)\| \Big|_{t=0} = \omega(\mathbf{A});$$

see, e.g., [23, Thm. 17.4]. Thus $\omega(\mathbf{A}) > 0$ is a necessary condition for solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ to exhibit transient growth.
- Pseudospectra: The requirement in (2.6) that ϕ_k be analytic throughout $W(\mathbf{A})$, which is typically enforced through the condition $W(\mathbf{A}) \subset \mathbb{C}^-$, excludes many stable $\mathbf{A}$. Thus we consider a more flexible alternative. Given $\varepsilon > 0$, if ϕ_k is analytic on the ε-pseudosopectrum [23]

$$\begin{aligned} \sigma_\varepsilon(\mathbf{A}) &= \{z \in \mathbb{C} : z \in \sigma(\mathbf{A}) \text{ or } \|(z - \mathbf{A})^{-1}\| > 1/\varepsilon\} \\ &= \{z \in \mathbb{C} : z \in \sigma(\mathbf{A} + \mathbf{E}) \text{ for some } \mathbf{E} \in \mathbb{C}^{n \times n} \text{ with } \|\mathbf{E}\| < \varepsilon\}, \end{aligned}$$

then

$$(2.8) \qquad \|\phi_k(\mathbf{A})\| \leq \frac{L_\varepsilon}{2\pi\varepsilon} \sup_{z \in \sigma_\varepsilon(\mathbf{A})} |\phi_k(z)|,$$

where L_ε denotes the contour length of the boundary of $\sigma_\varepsilon(\mathbf{A})$; see, e.g., [23, p. 139]. Substituting (2.8) into (2.1) yields [19, eq. (3.4)]

$$(2.9) \qquad \frac{s_{kr+1}}{s_1} \leq \frac{L_\varepsilon^2}{4\pi^2\varepsilon^2} \max_{z \in \sigma_\varepsilon(\mathbf{A})} \prod_{j=1}^{k} \frac{|z + \mu_j|^2}{|z - \overline{\mu_j}|^2}.$$

The choice of $\varepsilon > 0$ balances the leading constant against the set over which the maximization occurs: increasing ε typically decreases $L_\varepsilon^2/(4\pi^2\varepsilon^2)$ but enlarges $\sigma_\varepsilon(\mathbf{A})$. For any $\{\mu_j\} \subset \mathbb{C}^+$ there exists $\varepsilon > 0$ sufficiently small such that ϕ_k is analytic on $\sigma_\varepsilon(\mathbf{A})$ since $\sigma(\mathbf{A}) \subset \mathbb{C}^-$ and $\sigma_\varepsilon(\mathbf{A})$ converges to $\sigma(\mathbf{A})$ in the Hausdorff metric as $\varepsilon \to 0$; see, e.g., [23, Chap. 4] for related details.

All these bounds derived from (2.1) predict that the decay of singular values will slow as the departure of $\mathbf{A}$ from normality increases, as reflected in increased ill-conditioning of the eigenvector matrix (i.e., the eigenvectors associated with distinct eigenvalues become increasingly aligned), enlargement of the numerical range, or an increase in the sensitivity of the eigenvalues to perturbations.

Several alternative bounds on the singular values of $\mathbf{X}$ have been derived using entirely different approaches, but they share this same property. Convergence theorems for the low-rank approximate solutions to the Lyapunov equation constructed by projection methods also provide upper bounds on the decay of the singular values of $\mathbf{X}$. In this literature, results based on the numerical range have proved most popular. Like (2.6), these bounds predict slower singular value decay as the distance of the numerical range from the origin decreases, and they fail to hold when $0 \in W(\mathbf{A})$ (see, e.g., Theorem 4.2 of [7], which resembles (2.6), and Corollary 2.5 of [3]). Thus they do not apply to the highly nonnormal examples that interest us here.

Antoulas, Sorensen, and Zhou [2, Thm. 3.1] propose a different strategy. For diagonalizable $\mathbf{A}$ they write the solution $\mathbf{X}$ as a finite series to show, for $r = 1$, that

$$(2.10) \qquad s_{k+1} \leq (n - k)^2 \|\mathbf{V}\|^2 \|\mathbf{V}^{-1}\|^2 \|\mathbf{B}\|^2 \delta_{k+1},$$

where

$$\delta_k = -\frac{1}{2\operatorname{Re}\lambda_k}\prod_{j=1}^{k-1}\frac{|\lambda_k-\lambda_j|^2}{|\lambda_k+\overline{\lambda_j}|^2},$$

with the eigenvalues $\lambda_1,\ldots,\lambda_n$ of $\mathbf{A}$ ordered to make $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n$. By (1.1), we have

$$(2.11) \qquad \|\mathbf{B}\|^2 = \|\mathbf{B}\mathbf{B}^*\| = \|\mathbf{A}\mathbf{X}+\mathbf{X}\mathbf{A}^*\| \leq 2\|\mathbf{A}\|\|\mathbf{X}\| = 2\|\mathbf{A}\|s_1,$$

so (2.10) implies the relative bound

$$\frac{s_{k+1}}{s_1} \leq 2(n-k)^2\|\mathbf{A}\|\,\|\mathbf{V}\|^2\|\mathbf{V}^{-1}\|^2\,\delta_{k+1}.$$

(The $r > 1$ case is slightly more complicated [2, Thm. 3.2].) By analyzing the trace of $\mathbf{X}$, Truhar and Veselić [26] derive an alternative to (2.10) that characterizes the departure of $\mathbf{A}$ from normality by terms like $\|\mathbf{V}\|^2\|\widehat{\mathbf{b}}_j\|^2$, where $\widehat{\mathbf{b}}_j^*$ denotes the jth row of $\mathbf{V}^{-1}\mathbf{B}$ (and thus depends on the conditioning of the eigenvectors of $\mathbf{A}$). This bound can be generalized to nondiagonalizable $\mathbf{A}$, with an explicit formula given for 2×2 Jordan blocks [26, Thm. 2.2], and can be further generalized to Sylvester equations [25]. Bounds for coefficients $\mathbf{A}$ that are non-self-adjoint operators on Hilbert space exhibit similar dependence on the square of the condition number of the transformation that orthogonalizes a Riesz basis of eigenvectors [12, Thm. 4.1].

When $W(\mathbf{A}) \subset \mathbb{C}^-$, these bounds can be qualitatively descriptive, even when $\mathbf{A}$ departs significantly from normality. For a simple example, suppose $\mathbf{A}$ is a discretization of the differential operator

$$\frac{\mathrm{d}}{\mathrm{d}x} - 1$$

defined on absolutely continuous functions in $L^2(0,1)$ satisfying $u(1) = 0$. Approximating the operator with forward finite differences on the uniform grid with spacing $1/n$ gives

$$\mathbf{A} = \begin{bmatrix} -1-n & n & & \\ & -1-n & \ddots & \\ & & \ddots & n \\ & & & -1-n \end{bmatrix} \in \mathbb{C}^{n\times n}$$

with spectrum $\sigma(\mathbf{A}) = \{-1-n\}$ in the left half-plane. Since $\mathbf{A}$ is a Jordan block, its numerical range is known in closed form [16]:

$$W(\mathbf{A}) = \left\{z\in\mathbb{C}: |z+1+n|\leq n\cos\left(\frac{\pi}{n+1}\right)\right\},$$

which is a disk centered at $-1-n$ of radius $n\cos(\pi/(n+1))$. Notice that as n increases, $W(\mathbf{A})$ enlarges monotonically: the numerical range includes larger portions of the half-plane $\{z\in\mathbb{C}:\operatorname{Re}z < -1\}$, reflecting the resolvent behavior of the underlying differential operator [23, sect. 5]. As n increases, the singular value decay slows. This behavior is shown in Figure 1, where $\mathbf{B}$ is a constant vector. In this case, as predicted by the bounds we have surveyed, an increasing departure from normality
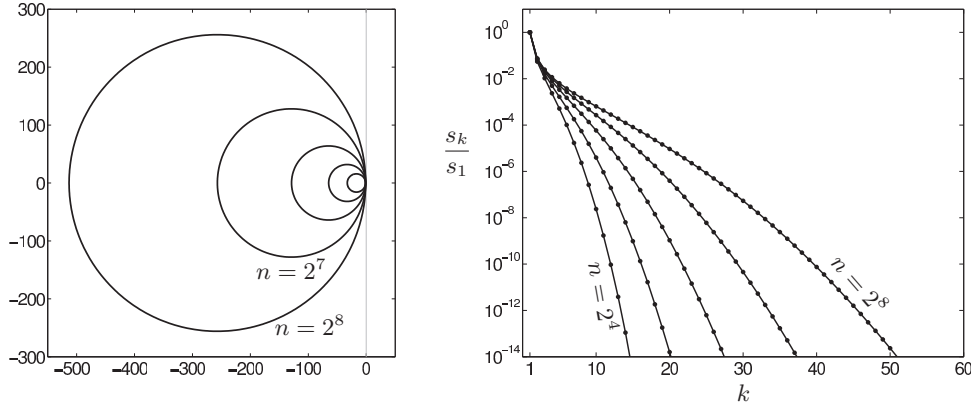
FIG. 1. *Boundaries of $W(\mathbf{A})$ (left) and decay of singular values of $\mathbf{X}$ (right) for discretizations of $d/dx - 1$ of dimension $n = 2^4, 2^5, \ldots, 2^8$. As n increases, the numerical range enlarges, while the singular values decay more slowly. This correlation is consistent with previously known bounds.*

slows convergence. We shall see that $\omega(\mathbf{A}) = -1 - n(1 - \cos(\pi/(n+1))) < 0$ is a crucial property.

Not all nonnormal coefficients give this same behavior. To see how the known bounds fail to capture the rich behavior exhibited by the singular values of Lyapunov solutions with highly nonnormal coefficients, consider those $\mathbf{X}$ that exhibit no decay at all, i.e., $\mathbf{X} = \xi \mathbf{I}$ for some $\xi > 0$, for rank$(\mathbf{B}) < n$. Then (1.1) reduces to

$$\mathbf{A} + \mathbf{A}^* = \frac{-1}{\xi}\mathbf{B}\mathbf{B}^*,$$

which implies that the Hermitian part of $\mathbf{A}$ is a negative semidefinite matrix having rightmost eigenvalue (and hence numerical abscissa, $\omega(\mathbf{A})$) equal to zero. If the numerical abscissa is positive, reflecting a larger departure from normality, the singular values *must* decay faster. Important applications give rise to matrices with $\omega(\mathbf{A}) > 0$; for example, positive $\omega(\mathbf{A})$ can grow with Reynolds number in fluid flows, a fact that complicates studies of transition to turbulence [24]. Lyapunov equations with low-rank right-hand sides have recently been applied to study this problem [9]. To cleanly illustrate the inadequacy of existing bounds, we next study a family of 2×2 matrices.

3. A completely solvable example. Consider the following 2×2 example from [19], where we interpret "singular value decay" to mean the ratio of the first two singular values, s_2/s_1. Consider the coefficient and right-hand side[1]

$$(3.1) \qquad \mathbf{A}(\alpha) = \begin{bmatrix} -1 & \alpha \\ 0 & -1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} t \\ 1 \end{bmatrix}.$$

Note that $W(\mathbf{A}(\alpha))$ is the disk in $\mathbb{C}$ centered at $\lambda = -1$ with radius $|\alpha|/2$. The solution to the Lyapunov equation can be written out explicitly:

$$\mathbf{X} = \frac{1}{4}\begin{bmatrix} 2t^2 + 2\alpha t + \alpha^2 & \alpha + 2t \\ \alpha + 2t & 2 \end{bmatrix}.$$

---

[1]Note the normalization of $\mathbf{B}$; if the second component of $\mathbf{B}$ is zero, then $\mathbf{B}$ is an eigenvector of $\mathbf{A}$, and the corresponding linear system is not controllable [1].
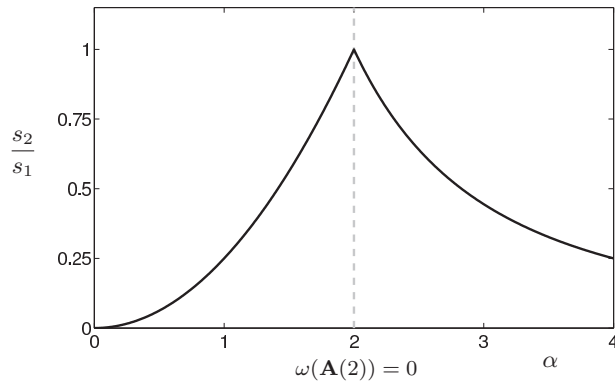
FIG. 2. *Worst-case singular value decay for the matrix* $\mathbf{A}(\alpha)$ *in* (3.1) *and controllable* $\mathbf{B} \in \mathbb{R}^{2 \times 1}$ *as a function of* α. *As* α *increases from zero, so does the departure of* $\mathbf{A}(\alpha)$ *from normality. The singular value ratio* s_2/s_1 *increases up until* $\omega(\mathbf{A}(\alpha)) = 0$; *beyond this point,* s_2/s_1 *decreases.*

We seek the right-hand side $\mathbf{B}$ that gives the *slowest* decay, i.e., that maximizes

$$\frac{s_2}{s_1} = \frac{\text{tr}(\mathbf{X}) - \sqrt{\text{tr}(\mathbf{X})^2 - 4 \det(\mathbf{X})}}{\text{tr}(\mathbf{X}) + \sqrt{\text{tr}(\mathbf{X})^2 - 4 \det(\mathbf{X})}} \leq 1$$

over all controllable $\mathbf{B} \in \mathbb{R}^2$, i.e., over all $t \in \mathbb{R}$. This worst-case decay is attained when $t = -\alpha/2$, giving

$$\frac{s_2}{s_1} = \begin{cases} \alpha^2/4, & 0 < \alpha \leq 2; \\ 4/\alpha^2, & 2 \leq \alpha. \end{cases}$$

As α increases from zero, so too does the departure of $\mathbf{A}(\alpha)$ from normality. The ratio s_2/s_1 also increases but only up to $\alpha = 2$ (when $\omega(\mathbf{A}(\alpha)) = 0$). As α increases beyond $\alpha = 2$, the ratio of singular values decreases significantly: contrary to our expectation from bounds described in section 2, *the decay actually improves.* Figure 2 illustrates this behavior.

4. Krylov conditioning and decay. We can gain some general insight into this decay behavior by writing a Lyapunov solution $\mathbf{X}$ in terms of the solution of a related canonical Lyapunov equation that depends only on the spectrum of $\mathbf{A}$. This formulation is not intended for practical calculations, but it provides some intuition for the results that follow in section 5.

Let $\mathbf{B} \in \mathbb{C}^{n \times 1}$, and suppose $(\mathbf{A}, \mathbf{B})$ is controllable. Thus $\mathbf{A}$ is nonderogatory, so its minimum polynomial equals its characteristic polynomial,

$$\chi(z) = (z - \lambda_1) \cdots (z - \lambda_n) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1} + z^n.$$

Let $\mathbf{A}_c$ be the associated companion matrix,

$$\mathbf{A}_c = \begin{bmatrix} & & & -c_0 \\ 1 & & & -c_1 \\ & \ddots & & \vdots \\ & & 1 & -c_{n-1} \end{bmatrix},$$

whose eigenvalues are the same as those of $\mathbf{A}$. Antoulas, Sorensen, and Zhou [2, Lem. 3.1] describe the following method for constructing the solution $\mathbf{X}$ to $\mathbf{AX} +$

$\mathbf{X}\mathbf{A}^* = -\mathbf{B}\mathbf{B}^*$. Let $\mathbf{K}$ denote the Krylov matrix

$$\mathbf{K} = [\mathbf{B} \ \mathbf{A}\mathbf{B} \ \cdots \ \mathbf{A}^{n-1}\mathbf{B}] \in \mathbb{C}^{n \times n},$$

and let $\mathbf{e}_1$ be the first column of the $n \times n$ identity matrix. Then $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{B}\mathbf{B}^*$ if and only if $\mathbf{X} = \mathbf{K}\mathbf{G}\mathbf{K}^*$, where $\mathbf{G}$ solves the companion Lyapunov equation

$$\mathbf{A}_c\mathbf{G} + \mathbf{G}\mathbf{A}_c^* = -\mathbf{e}_1\mathbf{e}_1^*.$$

Notice that $\mathbf{G}$ depends only on $\mathbf{A}_c$, and hence only on the spectrum of $\mathbf{A}$, not the departure of $\mathbf{A}$ from normality or the right-hand side $\mathbf{B}$: the influence of these latter factors on $\mathbf{X}$ occurs only through the matrix $\mathbf{K}$.

Let $\varsigma_k(\cdot)$ denote the kth singular value of a matrix. Since $\mathbf{G}$ is positive definite, it has a square root, and so

$$s_k := \varsigma_k(\mathbf{X}) = \varsigma_k(\mathbf{K}\mathbf{G}\mathbf{K}^*) = \varsigma_k(\mathbf{K}\mathbf{G}^{1/2})^2 \le \varsigma_k(\mathbf{K})^2\varsigma_1(\mathbf{G}^{1/2})^2 = \varsigma_k(\mathbf{K})^2\|\mathbf{G}\|,$$

using the singular value inequality [14, Thm. 3.3.16(d)]. Use (2.11), $\|\mathbf{B}\mathbf{B}^*\| \le 2\|\mathbf{A}\|s_1$, to obtain the bound

$$(4.1) \qquad \frac{s_k}{s_1} \le \varsigma_k(\mathbf{K})^2\|\mathbf{A}\|\left(\frac{2\|\mathbf{G}\|}{\|\mathbf{B}\mathbf{B}^*\|}\right).$$

The singular values of $\mathbf{X}$ will thus decay (at least) at a rate controlled by the singular values of the Krylov matrix $\mathbf{K}$; note that the term in parentheses in (4.1) is independent of the departure of $\mathbf{A}$ from normality. The columns of $\mathbf{K}$ are iterates of the power method, and hence one can gain insight into the decay of singular values of $\mathbf{X}$ by studying the convergence of the power method for nonnormal $\mathbf{A}$. (See [23, sect. 28], especially the illustration in Figure 28.1 showing how nonnormality can *accelerate* the convergence of the power method.) We shall not pursue this direction here; instead we will imagine fixing $\mathbf{B}$ and the spectrum of $\mathbf{A}$ and then varying the departure of $\mathbf{A}$ from normality, e.g.,

$$\mathbf{A} = \mathbf{\Lambda} + \alpha\mathbf{S},$$

where $\mathbf{\Lambda}$ is diagonal, $\mathbf{S}$ is strictly upper triangular, and α controls the departure of $\mathbf{A}$ from normality. For a concrete example, take $\mathbf{\Lambda} = -\mathbf{I}$ and $\mathbf{S}$ to be the shift matrix, yielding a Jordan block that generalizes the example in section 3:

$$(4.2) \qquad \mathbf{A} = \begin{bmatrix} -1 & \alpha & & \\ & -1 & \ddots & \\ & & \ddots & \alpha \\ & & & -1 \end{bmatrix}.$$

The departure of $\mathbf{A}$ from normality is small when α is small. In this case $\mathbf{A} \approx -\mathbf{I}$, so all columns of $\mathbf{K}$ will be nearly the same: $\varsigma_k(\mathbf{K})$ will be small for all $k \ge 2$, and (4.1) captures the fast decay of the singular values of $\mathbf{X}$. For large α, the matrix $\mathbf{K}$ will be severely graded; specifically, the norm of the kth column of $\mathbf{K}$ will be on the order of α^{k-1}. Thus for large α, the singular values of $\mathbf{K}$ must also decay rapidly.[2] Since $\|\mathbf{A}\|$

---

[2]One could apply results on the singular values of graded matrices (see, e.g., [22]) to obtain quantitative estimates.
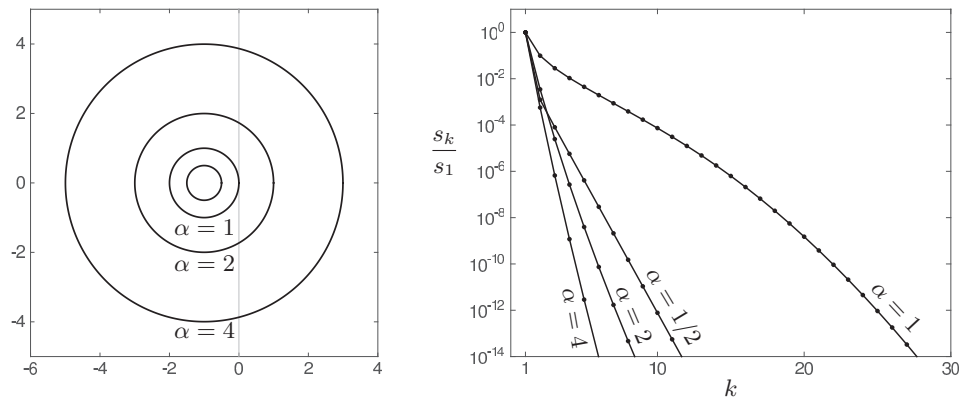
FIG. 3. *Boundaries of* $W(\mathbf{A})$ *(left) and decay of singular values of* $\mathbf{X}$ *(right) for Jordan blocks* (4.2) *of dimension* $n = 64$ *with off-diagonal* $\alpha = 1/2, 1, 2, 4$*. As* α *increases, the numerical range enlarges. For small and large* α*, the singular values of* $\mathbf{X}$ *decay quickly; the* $\alpha = 1$ *case, having an intermediate departure of* $\mathbf{A}$ *from normality, gives singular values of* $\mathbf{X}$ *that decay much slower. (Here* $\mathbf{B} = [1, \ldots, 1]^*$*.)*

only grows linearly with α,[3] by (4.1), the singular values of $\mathbf{X}$ must decay quickly as well. The slowest decay should thus occur for values of α that are neither too small nor too large, as suggested by the two dimensional case. Indeed, this intuition is confirmed in Figure 3, which shows an example with $n = 64$ and $\alpha = 1/2, 1, 2, 4$. Of the cases shown, the singular values decay most slowly for $\alpha = 1$, when the rightmost extent of $W(\mathbf{A})$ comes closest to the imaginary axis. We next describe rigorous bounds that connect properties of $W(\mathbf{A})$ to the decay of the singular values of $\mathbf{X}$.

5. Large numerical abscissa implies fast decay. In (2.7) we defined the numerical abscissa, $\omega(\mathbf{A})$, which is both the rightmost extent of the numerical range and the rightmost eigenvalue of the Hermitian part $(\mathbf{A}+\mathbf{A}^*)/2$ of $\mathbf{A}$. The subordinate eigenvalues of the Hermitian part further inform our understanding of the departure of $\mathbf{A}$ from normality. For example, these eigenvalues have recently been used to bound the number of Ritz values of $\mathbf{A}$ that can fall in subregions of $W(\mathbf{A})$ [5, Thm. 1.2].[4] Like $\omega(\mathbf{A})$, interior eigenvalues of $(\mathbf{A} + \mathbf{A}^*)/2$ can be positive even when $\mathbf{A}$ is stable. The following theorem bounds these eigenvalues in terms of the singular values of $\mathbf{X}$. This result can be read from two different perspectives: given the singular values of $\mathbf{X}$, the bound reveals something about those $\mathbf{A}$ that can support such solutions (Theorem 5.1 and Corollary 5.2); given $\mathbf{A}$, one obtains an upper bound on the decay of singular values of $\mathbf{X}$ that requires, in a specific context, *faster* decay as the departure of $\mathbf{A}$ from normality increases (Corollary 5.3).

THEOREM 5.1. *Let* $\mathbf{X} \in \mathbb{C}^{n \times n}$ *solve the Lyapunov equation* (1.1) *with* $(\mathbf{A}, \mathbf{B})$ *controllable. Then for all* $k = 1, \ldots, n$,

$$(5.1) \qquad \frac{s_k}{s_1} - 1 - \frac{\|\mathbf{B}\|^2}{2 s_1 \|\mathbf{A}\|} \leq \frac{\omega_k}{\|\mathbf{A}\|} \leq 1 - \frac{s_{n-k+1}}{s_1},$$

where ω_k *denotes the* kth *rightmost eigenvalue of* $\frac{1}{2}(\mathbf{A} + \mathbf{A}^*)$ *and* s_k *denotes the* kth *singular value of* $\mathbf{X}$.

---

[3]Gerschgorin's theorem applied to $\mathbf{A}^*\mathbf{A}$ gives $\alpha - 1 \leq \|\mathbf{A}\| \leq \alpha + 1$ for $\alpha > 3$.

[4]In the context of moment-matching model reduction algorithms [1, Chap. 11], these results relating Ritz values to the eigenvalues of $(\mathbf{A} + \mathbf{A}^*)/2$ restrict the number of poles of a reduced-order model that can fall in the right half-plane.

Proof. Write the solution $\mathbf{X} = \xi(\mathbf{I} - \mathbf{E})$ for $\xi > 0$ and $\mathbf{E}$ Hermitian. Then since $\mathbf{X}$ solves the Lyapunov equation (1.1),

$$
(5.2) \qquad \frac{\mathbf{A} + \mathbf{A}^*}{2} = -\frac{1}{2\xi}\mathbf{B}\mathbf{B}^* + \frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}.
$$

Let $\lambda_k(\cdot)$ denote the kth eigenvalue of a Hermitian matrix, labeled from right to left, and let $\varsigma_k(\cdot)$ denote the *kth* singular value of a matrix, again labeled from largest to smallest. Weyl's inequalities for the eigenvalues of sums of Hermitian matrices (see, e.g., [15, Thm. 4.3.1]) imply that

$$
\lambda_n\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^*\right) + \lambda_k\left(\frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right) \leq \lambda_k\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^* + \frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right)
$$

and

$$
\lambda_k\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^* + \frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right) \leq \lambda_1\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^*\right) + \lambda_k\left(\frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right).
$$

Since $-\mathbf{B}\mathbf{B}^*/(2\xi)$ is Hermitian negative semidefinite,

$$
\lambda_n\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^*\right) = -\frac{\|\mathbf{B}\|^2}{2\xi}, \qquad \lambda_1\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^*\right) \leq 0.
$$

Now by (5.2),

$$
\lambda_k\left(-\frac{1}{2\xi}\mathbf{B}\mathbf{B}^* + \frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right) = \lambda_k\left(\frac{\mathbf{A} + \mathbf{A}^*}{2}\right) =: \omega_k.
$$

Together, these pieces imply that

$$
(5.3) \qquad -\frac{\|\mathbf{B}\|^2}{2\xi} + \lambda_k\left(\frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right) \leq \omega_k \leq \lambda_k\left(\frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right).
$$

Note that $(\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*)/2$ is the Hermitian part of $\mathbf{A}\mathbf{E}$. The kth singular value of a matrix gives an upper bound on the kth rightmost eigenvalue of its Hermitian part [14, Cor. 3.1.5]. Applying this bound to both $\mathbf{A}\mathbf{E}$ and $-\mathbf{A}\mathbf{E}$ gives

$$
-\varsigma_{n-k+1}(\mathbf{A}\mathbf{E}) \leq \lambda_k\left(\frac{\mathbf{A}\mathbf{E} + \mathbf{E}\mathbf{A}^*}{2}\right) \leq \varsigma_k(\mathbf{A}\mathbf{E}).
$$

Using the singular value inequality [14, Thm. 3.3.16(d)],

$$
\varsigma_k(\mathbf{A}\mathbf{E}) \leq \varsigma_1(\mathbf{A})\,\varsigma_k(\mathbf{E}) = \|\mathbf{A}\|\,\varsigma_k(\mathbf{E});
$$

now obtain from (5.3) that

$$
(5.4) \qquad -\frac{\|\mathbf{B}\|^2}{2\xi\|\mathbf{A}\|} - \varsigma_{n-k+1}(\mathbf{E}) \leq \frac{\omega_k}{\|\mathbf{A}\|} \leq \varsigma_k(\mathbf{E}).
$$

Since $\mathbf{E} = \mathbf{I} - \mathbf{X}/\xi$, the eigenvalues of $\mathbf{E}$, labeled from right to left, are

$$
\lambda_k(\mathbf{E}) = 1 - s_{n-k+1}/\xi, \qquad k = 1, \dots, n.
$$

The form $\mathbf{X} = \xi(\mathbf{I} - \mathbf{E})$ allows for various choices of ξ and $\mathbf{E}$. Taking $\xi = s_1$ gives $\mathbf{E} = \mathbf{I} - \mathbf{X}/s_1$, and hence $0 = \lambda_n(\mathbf{E}) \le \cdots \le \lambda_1(\mathbf{E})$ and

$$\varsigma_k(\mathbf{E}) = 1 - s_{n-k+1}/s_1.$$

Thus (5.4) implies that

$$\frac{s_k}{s_1} - 1 - \frac{\|\mathbf{B}\|^2}{2 s_1 \|\mathbf{A}\|} \le \frac{\omega_k}{\|\mathbf{A}\|} \le 1 - \frac{s_{n-k+1}}{s_1}. \qquad \square$$

Remark 5.1. In the proof of Theorem 5.1, the choice $\xi = s_1$ for the scaling factor ξ is usually suboptimal. Smaller values of $\xi > 0$ can give tighter bounds but usually at the expense of more intricate formulas (since then the eigenvalues of $\mathbf{E}$ can be positive and negative). As a special case, we can take $\xi = (s_1 + s_n)/2$ to optimize (5.4) for $k = 1$, giving $\lambda_1(\mathbf{E}) = -\lambda_n(\mathbf{E}) = (s_1 - s_n)/(s_1 + s_n)$ and

$$\omega(\mathbf{A}) \le \frac{s_1 - s_n}{s_1 + s_n} \|\mathbf{A}\|.$$

This expression has a nice interpretation: if the smallest singular value s_n of $\mathbf{X}$ is on the same order as s_1, then $\omega(\mathbf{A})$ must be quite a bit smaller than $\|\mathbf{A}\|$. When combined with the $k = 1$ lower bound from Theorem 5.1 (with $\xi = s_1$), we obtain bounds on the rightmost extent of any numerical range that can support a solution $\mathbf{X}$ with extreme singular values s_1 and s_n.

COROLLARY 5.2. *For controllable $(\mathbf{A}, \mathbf{B})$, the numerical abscissa $\omega(\mathbf{A})$ is bounded by the extreme singular values of the solution $\mathbf{X} \in \mathbb{C}^{n \times n}$ to the Lyapunov equation* (1.1):

$$(5.5) \qquad -\frac{\|\mathbf{B}\|^2}{2 s_1} \le \omega(\mathbf{A}) \le \frac{s_1 - s_n}{s_1 + s_n} \|\mathbf{A}\|.$$

Figure 4 provides a schematic illustration of this corollary. When the singular values decay slowly (as described in the caption), the numerical range cannot extend to the right of the gray strip. Note that the converse need not hold: the singular values can decay quickly regardless of $W(\mathbf{A})$, depending on $\mathbf{B}$ and finer spectral properties of $\mathbf{A}$. Also notice that the singular values can decay slowly when $W(\mathbf{A})$ is far to the left of the imaginary axis but the rank of $\mathbf{B}$ is large, consistent with the bound (2.6). (An extreme example is $\mathbf{A} = -\mathbf{I}$ and $\mathbf{B} = \mathbf{I}$, for which $\omega(\mathbf{A}) = -1$ but $s_n/s_1 = 1$.)

Rearranging the upper bound in Theorem 5.1 gives an upper bound on the decay of the trailing singular values of $\mathbf{X}$ that is distinguished by being independent of the rank of $\mathbf{B}$.

COROLLARY 5.3. *For controllable $(\mathbf{A}, \mathbf{B})$, the singular values of the solution $\mathbf{X} \in \mathbb{C}^{n \times n}$ to the Lyapunov equation* (1.1) *satisfy*

$$(5.6) \qquad \frac{s_{n-k+1}}{s_1} \le 1 - \frac{\omega_k}{\|\mathbf{A}\|}, \qquad k = 1, \ldots, n.$$

Remark 5.2. As observed in section 2, when $\mathrm{rank}(\mathbf{B}) < n$ the case of no decay ($s_1 = s_n$) implies that $\omega_1 \equiv \omega(\mathbf{A}) = 0$, in which case Corollary 5.3 with $k = 1$ is sharp. On the other hand, in the highly nonnormal case where $0 < \omega_k \approx \|\mathbf{A}\|$, Corollary 5.3 requires that the kth lowest singular value be small, *regardless of* $\mathbf{B}$. This stands in contrast to the traditional bounds surveyed in section 2 for two reasons: higher
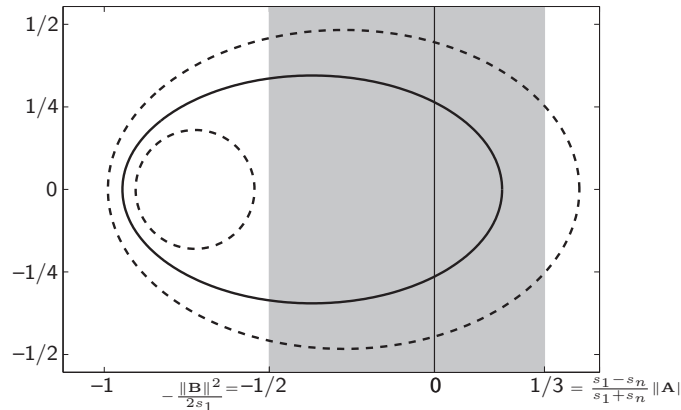
FIG. 4. *Visualization of Corollary* 5.2 *in the complex plane, with* $\|\mathbf{A}\| = \|\mathbf{B}\| = s_1 = 1$ *and* $s_n = 1/2$. *The three curves show boundaries* $\partial W(\mathbf{A})$ *of potential numerical ranges* $W(\mathbf{A})$. *The smaller dashed* $\partial W(\mathbf{A})$ *falls entirely to the left of the gray strip, violating the first inequality in* (5.5). *The rightmost extent of the solid* $\partial W(\mathbf{A})$ *falls inside the gray strip, so this* $W(\mathbf{A})$ *is consistent with* (5.5). *The larger dashed* $\partial W(\mathbf{A})$ *extends to the right of the gray strip, in violation of the second inequality of* (5.5): *the singular values of* $\mathbf{X}$ *must decay more rapidly for any* $\mathbf{A}$ *with such a numerical range.*

nonnormality implies faster decay, rather than slower decay; the rank of $\mathbf{B}$ does not feature in the bound on s_{n-k+1}/s_1, whereas the other bounds predict slower decay as the rank of $\mathbf{B}$ increases.

Corollary 5.3 is designed to show that decay must occur in this specific highly nonnormal scenario. The result is not useful when $\|\mathbf{A}\|$ is controlled by eigenvalues far in the left half-plane, rather than being dominated by the departure of $\mathbf{A}$ from normality. In this case, s_{n-k+1}/s_1 can be quite small while the right-hand side of (5.6) is not. In particular, when $\omega_k < 0$ (as must occur for all k when $\mathbf{A}$ is stable and normal), the bound in (5.6) is vacuous.

Remark 5.3. Note that the rate of decay could be even stronger than indicated by Corollary 5.3. For the 2×2 Jordan block considered in section 3,

$$\omega_1 = \alpha/2 - 1, \qquad \|\mathbf{A}\| = \sqrt{1 + \alpha^2/2 + \alpha\sqrt{\alpha^2/4 + 1}},$$

so Corollary 5.3 gives the bound

$$\frac{s_2}{s_1} \le 1 - \frac{\omega_1}{\|\mathbf{A}\|} \to 1/2, \qquad \alpha \to \infty,$$

whereas we saw in section 3 that $s_2/s_1 \to 0$ as $\alpha \to \infty$ for this example. Thus, while the results of this section are a marked improvement over previously existing bounds in some highly nonnormal regimes, they cannot be the last word on the subject.

6. Conclusions. We have illustrated a regime of stable matrices $\mathbf{A}$ for which all previous bounds on the decay of singular values of Lyapunov solutions fail to even qualitatively capture the correct behavior. This shortcoming is clear from specific examples; Theorem 5.1 and Corollary 5.3 provide a contrasting perspective on this phenomenon. While these results are not entirely sharp, they clearly illustrate that, beyond a threshold, an increased departure of $\mathbf{A}$ from normality can lead to faster decay of the singular values of $\mathbf{X}$. Sharper results will require a more complete understanding of the role of nonnormal coefficients on Lyapunov solutions.

REFERENCES

[1] A. C. ANTOULAS, *Approximation of Large-Scale Dynamical Systems*, SIAM, Philadelphia, 2005.

[2] A. C. ANTOULAS, D. C. SORENSEN, AND Y. ZHOU, *On the decay rate of Hankel singular values and related issues*, Systems Control Lett., 46 (2002), pp. 323–342.

[3] B. BECKERMANN, *An error analysis for rational Galerkin projection applied to the Sylvester equation*, SIAM J. Numer. Anal., 49 (2011), pp. 2430–2450.

[4] P. BENNER AND J. SAAK, *Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: A state of the art survey*, GAMM-Mitt., 36 (2013), pp. 32–52.

[5] R. L. CARDEN AND M. EMBREE, *Ritz value localization for non-Hermitian matrices*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1320–1338.

[6] M. CROUZEIX, *Numerical range and functional calculus in Hilbert space*, J. Funct. Anal., 244 (2007), pp. 668–690.

[7] V. DRUSKIN, L. KNIZHNERMAN, AND V. SIMONCINI, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1875–1898.

[8] N. S. ELLNER AND E. L. WACHSPRESS, *New ADI model problem applications*, in Proceedings of the 1986 ACM Fall Joint Computer Conference (Los Alamitos, CA, 1986), IEEE Computer Society Press, Los Angeles, CA, 1986, pp. 528–534.

[9] H. C. ELMAN, K. MEERBERGEN, A. SPENCE, AND M. WU, *Lyapunov inverse iteration for identifying Hopf bifurcations in models of incompressible flow*, SIAM J. Sci. Comput., 34 (2012), pp. A1584–A1606.

[10] L. ELSNER AND M. H. C. PAARDEKOOPER, *On measures of nonnormality of matrices*, Linear Algebra Appl., 92 (1987), pp. 107–123.

[11] M. EMBREE AND D. C. SORENSEN, *An Introduction to Model Reduction for Linear and Nonlinear Differential Equations*, manuscript.

[12] L. GRUBIŠIĆ AND D. KRESSNER, *On the eigenvalue decay of solutions to operator Lyapunov equations*, Systems Control Lett., 73 (2014), pp. 42–47.

[13] N. J. HIGHAM, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.

[14] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[15] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2013.

[16] M. MARCUS AND B. N. SHURE, *The numerical range of certain 0,1-matrices*, Linear Multilinear Algebra, 7 (1979), pp. 111–120.

[17] T. PENZL, *A cyclic low-rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418.

[18] T. PENZL, *Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case*, Systems Control Lett., 40 (2000), pp. 139–144.

[19] J. SABINO, *Solution of Large-Scale Lyapunov Equations via the Block Modified Smith Method*, Ph.D. thesis, Rice University, Houston, TX, 2006.

[20] V. SIMONCINI, *Computational Methods for Linear Matrix Equations*, preprint, 2014.

[21] D. C. SORENSEN AND Y. ZHOU, *Bounds on Eigenvalue Decay Rates and Sensitivity of Solutions to Lyapunov Equations*, Tech. Rep. TR 02-07, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 2002.

[22] G. W. STEWART, *On the eigensystems of graded matrices*, Numer. Math., 90 (2001), pp. 349–370.

[23] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.

[24] L. N. TREFETHEN, A. E. TREFETHEN, S. C. REDDY, AND T. A. DRISCOLL, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.

[25] N. TRUHAR, Z. TOMLJANOVIĆ, AND R.-C. LI, *Analysis of the solution of the Sylvester equation using low-rank ADI with exact shifts*, Systems Control Lett., 59 (2010), pp. 248–257.

[26] N. TRUHAR AND K. VESELIĆ, *Bounds on the trace of a solution to the Lyapunov equation with a general stable matrix*, Systems Control Lett., 56 (2007), pp. 493–503.

[27] K. ZHOU, *Robust and Optimal Control*, Prentice–Hall, Upper Saddle River, NJ, 1995.