Theory and Applications of the Polynomial Numerical Hull

A. Greenbaum
Mathematics Dept.
University of Washington
greenbau@math.washington.edu

Let A by an n by n matrix.

Question: What properties of A determine (or approximately determine) $\|f(A)\|$ for given functions f?

If A is a **normal** matrix and $\|\cdot\|$ is the 2-norm, then *eigenvalues* determine the behavior of A:

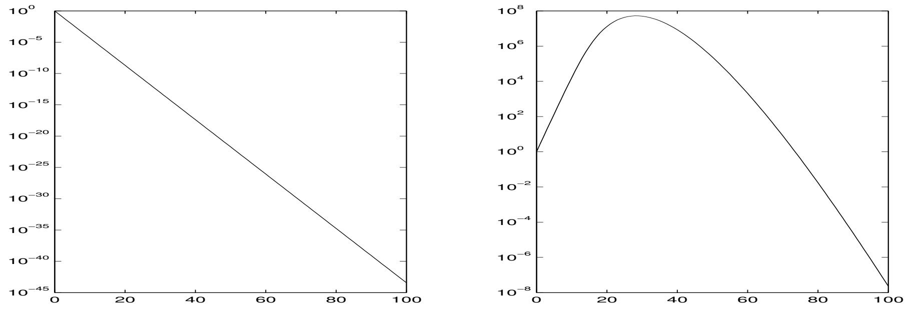
$$\|f(A)\| = \max_{\lambda \in \sigma(A)} |f(\lambda)|.$$

Suppose A is highly **nonnormal** (or $\|\cdot\|$ is very different from a unitarily invariant norm).

Functions of Interest

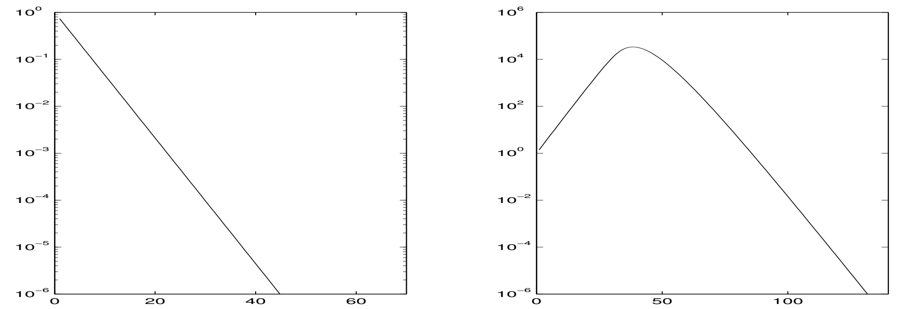
- e^{tA}: Determines the stability of $y' = Ay$.

  $\lim_{t\to\infty} \|e^{tA}\| = 0$ if and only if the eigenvalues of A have real parts less than 0. But eigenvalues alone cannot distinguish:



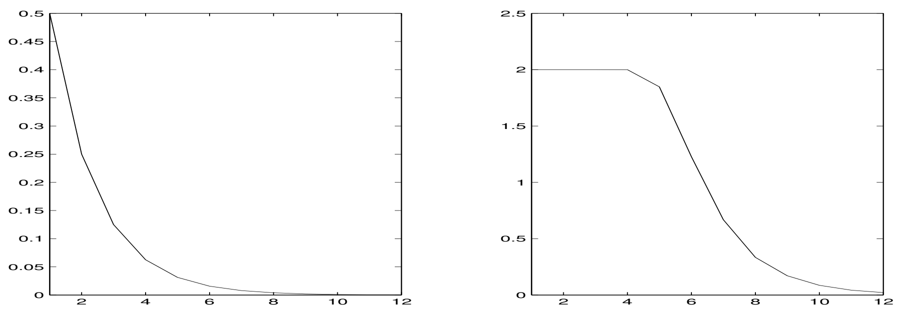
- A^k: Determines stability of finite difference schemes; determines the convergence of stationary iterative methods for linear systems.

  $\lim_{k\to\infty} \|A^k\| = 0$ if and only if $\rho(A) < 1$. But eigenvalues alone cannot distinguish:



- A^k: Markov chains.

  y_0 = initial state; $A^k y_0$ = state after k steps. $A^k y_0 \to v$ = eigenvector corresponding to eigenvalue 1. For k large, convergence rate is determined by second largest eigenvalue. But eigenvalues cannot distinguish:



- $\min_{p \in \mathcal{P}_k(0)} \|p(A)\|$: Residual norm in ideal GMRES.

  Since asymptotic behavior is not of interest (for $k \geq n$, residual is 0), eigenvalues alone may give little information.

Eigenvalues determine asymptotic behavior but not transient. What does?

- Field of values or numerical range:

$$\mathcal{F}(A) = \{q^*Aq : \quad q^*q = 1\}.$$

- ϵ-Pseudospectrum:

$$\begin{aligned} \Lambda_\epsilon(A) &= \{z \in \mathbf{C} : \quad \|(zI - A)^{-1}\| \geq \epsilon^{-1}\} \\ &= \{z \in \mathbf{C} : \quad z \text{ is an eigenvalue of } A + E \\ &\qquad \text{for some } E \text{ with } \|E\| \leq \epsilon\}. \end{aligned}$$

- Polynomial numerical hull of degree k (Nevanlinna, G.):

$$\begin{aligned} \mathcal{G}_k(A) &= \{z \in \mathbf{C} : \quad \|p(A)\| \geq |p(z)| \ \ \forall p \in \mathcal{P}_k\} \\ &= \{z \in \mathbf{C} : \quad \min_{p \in \mathcal{P}_k(0)} \|p(A - zI)\| \geq 1\}. \end{aligned}$$

Polynomial Numerical Hull of Degree k

$$\mathcal{G}_k(A) = \{z \in \mathbf{C} : \quad \|p(A)\| \geq |p(z)|$$
$$\text{for all } p \text{ of degree } k \text{ or less}\}$$

- For all k, $\mathcal{G}_k(A)$ contains the spectrum of A.
- For $k \geq$ degree of minimal polynomial of A, $\mathcal{G}_k(A) = \sigma(A)$.
- $\mathcal{G}_1(A) = \mathcal{F}(A)$.
- For $1 < k <$ degree of minimal polynomial, $\mathcal{G}_k(A)$ is intermediate between the field of values and the spectrum.

Properties of the PNH of Degree k

$$\mathcal{G}_k(A) = \{z \in \mathbf{C} : \quad \|p(A)\| \geq |p(z)|$$
$$\text{for all } p \text{ of degree } k \text{ or less}\}$$

- $\mathcal{G}_k(Q^*AQ) = \mathcal{G}_k(A)$, if Q is unitary and $\|\cdot\|$ is unitarily invariant.
- If $AQ = QB$, where Q has orthonormal columns, then $p(A)Q = Qp(B)$ and $\mathcal{G}_k(B) \subset \mathcal{G}_k(A)$.
- $\mathcal{G}_k(\alpha A + \beta I) = \alpha \mathcal{G}_k(A) + \beta$.
- Closed bounded set. $\mathcal{G}_k(A) \subset \{z : \quad |z| \leq \|A\|\}$. Not convex.

Computing the PNH of Degree k

- Test all points z in a region known to contain $\mathcal{G}_k(A)$ by computing:

  $$\min_{p \in \mathcal{P}_k(0)} \|p(A - zI)\|$$

  and seeing if this is greater than or equal to 1.

- For $\|\cdot\| = \|\cdot\|_2$, express the problem of minimizing $\|p(A - zI)\|$ as a semidefinite programming problem, and use software by Toh, Todd, and Tutuncu:

$$\min_{x \in \mathbf{C^k}} \|B_0 - \sum_{j=1}^{k} x_j B_j\| \quad \equiv \quad -\max(\lambda) \text{ s.t.}$$

$$\begin{pmatrix} 0 & B(x) \\ B(x)^* & 0 \end{pmatrix} + \lambda \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + Z = 0, \quad Z \geq 0.$$

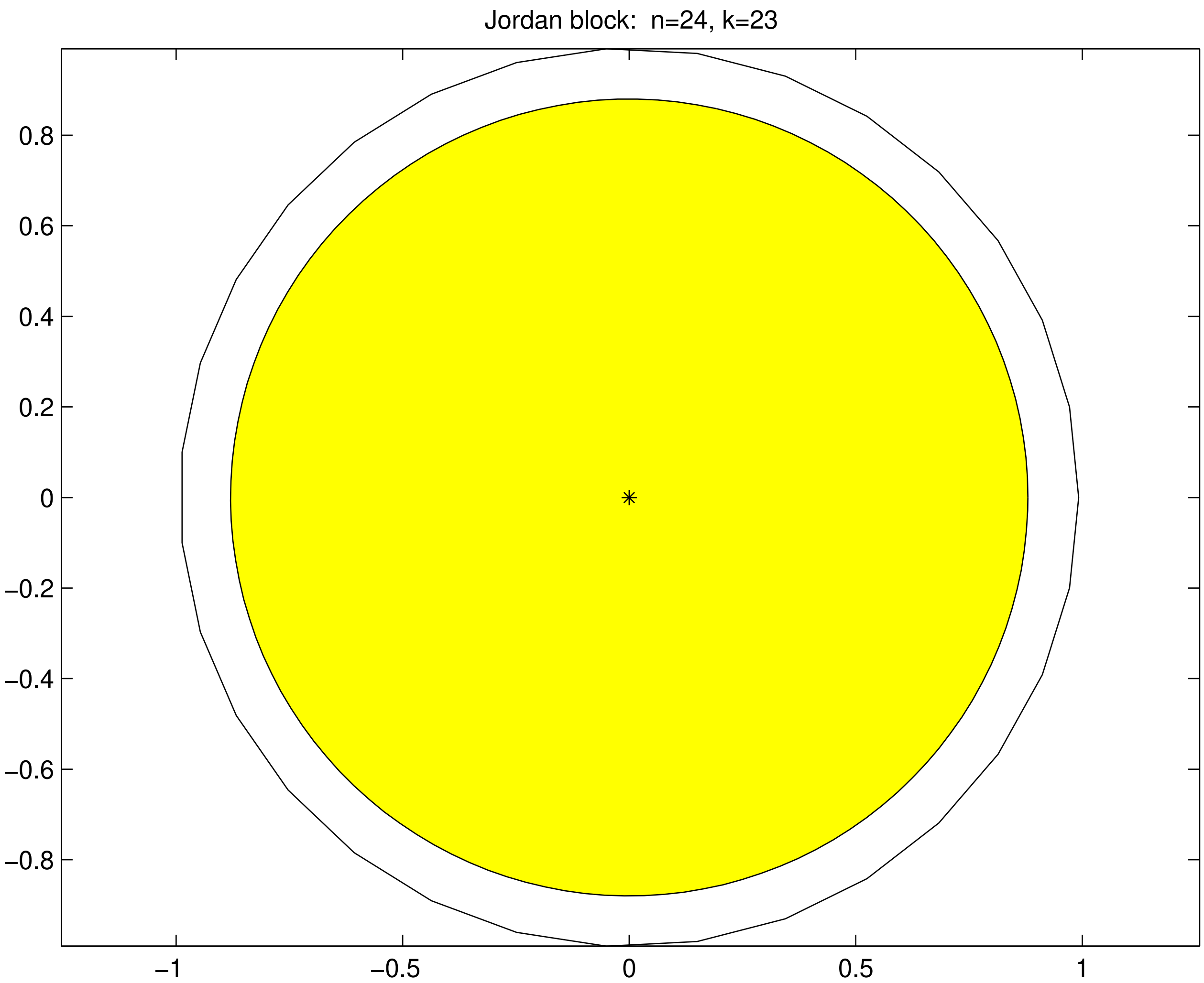


Figure 1: Field of values and polynomial numerical hull of degree 23 for a 24 by 24 Jordan block.

Analytic Expressions for the PNH of Degree k

Theorem: (Faber, G., Marshall) The 2-norm pnh of degree $n-1$ for an n by n Jordan block is a disk of radius $r_{n-1,n}$ about the eigenvalue, where: for n even, $r_{n-1,n}$ is the positive root of

$$2r^n + r - 1 = 0,$$

and for n odd, $r_{n-1,n}$ satisfies a similar formula. For large n,

$$r_{n-1,n} \approx 1 - \ln(2n)/n + \ln(\ln(2n))/n.$$

The pnh of degree 1 for an n by n Jordan block (i.e., the field of values) is a disk about the eigenvalue of radius

$$\cos(\pi/(n+1)) \approx 1 - .5(\pi/(n+1))^2.$$

The polynomial numerical hulls of all degrees for a large Jordan block are sandwiched together just inside the unit disk.

Banded Triangular Toeplitz Matrices

Since an n by n banded triangular Toeplitz matrix:

$$T = \begin{pmatrix} t_0 & \dots & t_b & & \\ & \ddots & & \ddots & \\ & & \ddots & & t_b \\ & & & \ddots & \vdots \\ & & & & t_0 \end{pmatrix}$$

is a polynomial in J $(T = t_0 I + t_1 J + \ldots + t_b J^b)$,

$$\|p_k(T)\| = \|p_k(q_b(J))\| \geq \max_{z \in \mathcal{G}_{kb}(J)} |p_k(q_b(z))|.$$

For $kb < n$, the pnh of degree k for T contains the image, under q_b, of the disk of radius $r_{kb,n}$.

The pnh of degree k for T is contained in that of T_∞, which is the image, under q_b, of the unit disk.

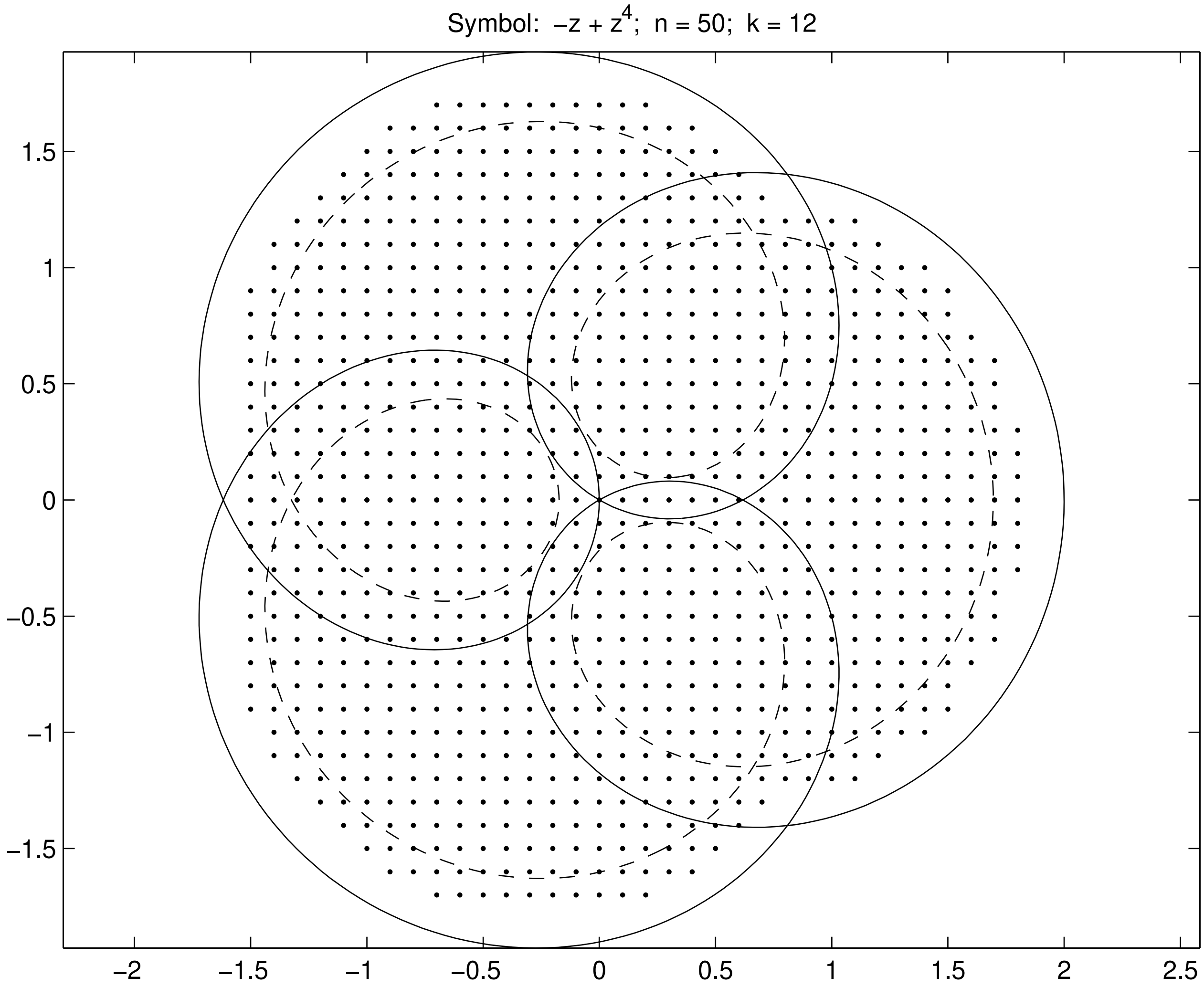


Figure 2: Image of the unit circle (solid) and of the circle of radius .934 (dashed) under the mapping $q_b(z) = -z + z^4$. Region with dots is the numerically computed polynomial numerical hull of degree 12 for the 50 by 50 Toeplitz matrix with symbol q_b.

Estimating the PNH of Degree k Using the Arnoldi Algorithm

- Outer bound: Run Arnoldi or GMRES and plot the curve $\{z \in \mathbf{C} : \ \|p_k(A)\| = |p_k(z)|\}$.

- Estimate: $\mathcal{G}_k(H_j) \approx \mathcal{G}_k(A)$.

- Inner bound: Use $j \times j$ piece of products of rectangular Hessenberg matrices: $H_{j+1,j}$, $H_{j+2,j+1}H_{j+1,j}$, etc., to obtain inner bound for $\mathcal{G}_k(A)$.

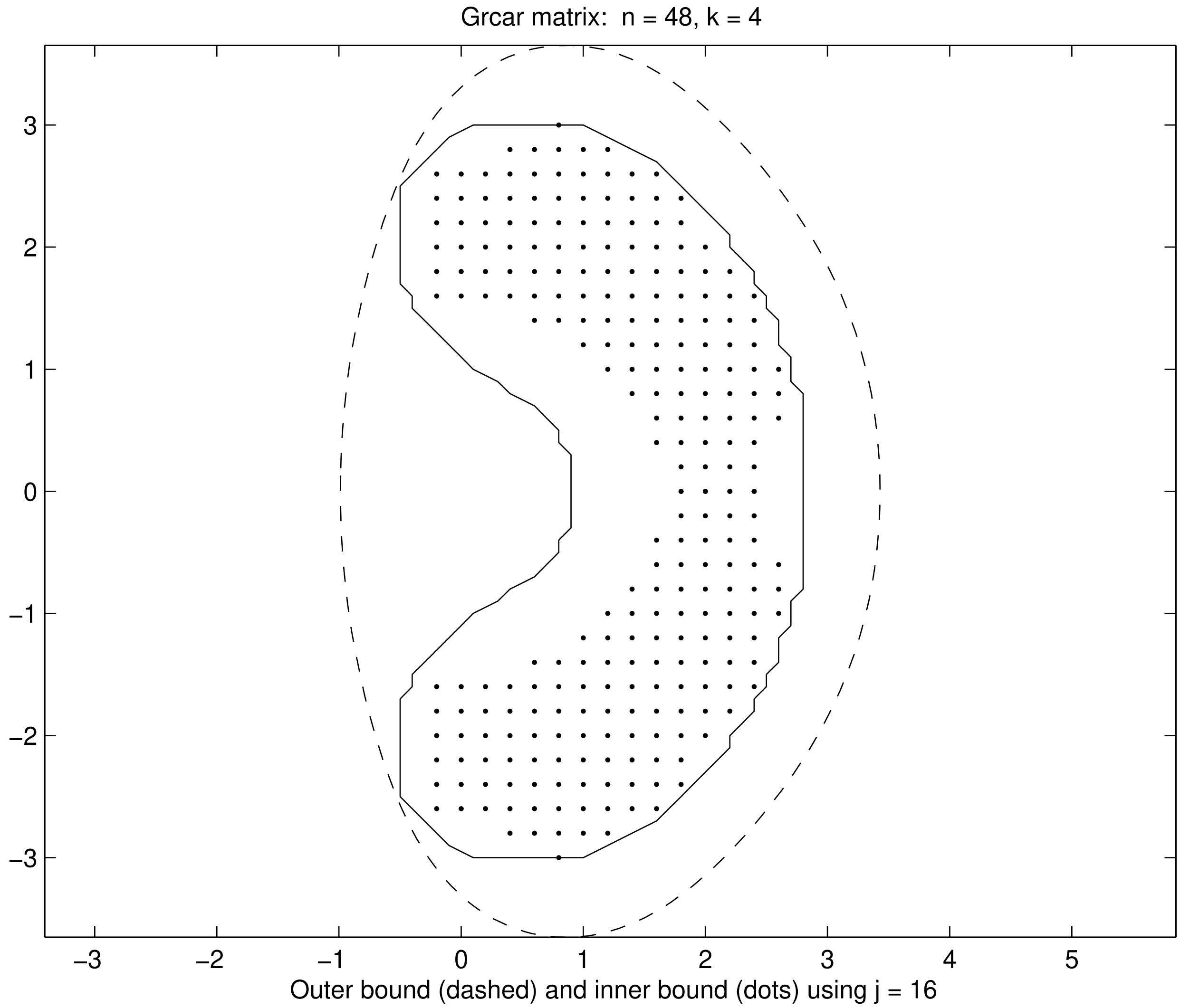


Figure 3: Inner and outer bounds on the pnh of degree 4 produced by the Arnoldi algorithm.

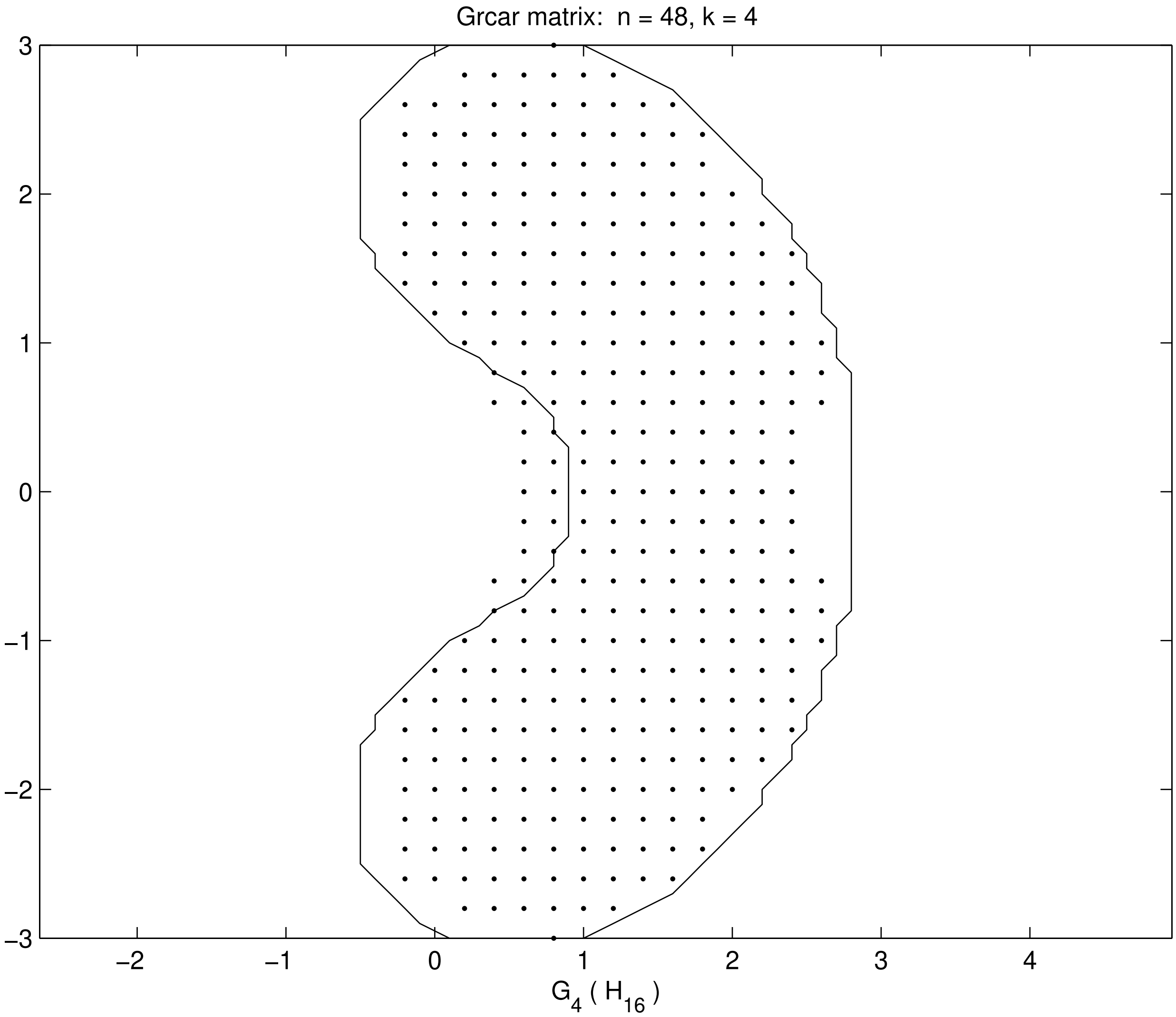


Figure 4: Estimate of the pnh of degree 4 produced by the Arnoldi algorithm.

Cutoff Phenomena in Markov Chains

- Probability transition matrix has a single eigenvalue of maximum modulus 1.
- *Asymptotic* rate of convergence to the steady state is governed by the second largest eigenvalue.
- Eigenvalues do **not** determine transient behavior. May take a number of steps before distance from steady state decreases appreciably, and then there may be a sudden drop.

Markov Chain Model of Riffle Shuffle

A deck of N cards is first cut into two heaps with the probability of a heap containing k cards being given by a binomial distribution: $\binom{N}{k}/2^N$. Cards from the left and right heaps are then riffled together into a single pile, with the probability of dropping the next card from a given heap being proportional to the number of cards in that heap.

Number of possible orderings $= N!$.

Rising Sequences

Bayer and Diaconis showed that the transition probabilities are completely determined by the number of *rising sequences*.

A rising sequence is a series of consecutively numbered cards found during one pass through the deck; for example, the arrangement 1452637 has two rising sequences, 123 and 4567.

Now states are the number of rising sequences: 1—N.

Probability Transition Matrix

(Jónsson and Trefethen)

$$P_{ji} = 2^{-n} \binom{n+1}{2i-j} \frac{\alpha_j}{\alpha_i},$$

where α_j denotes the number of permutations of $\{1, \ldots, n\}$ that have exactly j rising sequences. The numbers α_j are called *Eulerian numbers*.

Eigenvalues: 2^{-j}, $j = 0, 1, \ldots, n$.

Eigenvector corresponding to eigenvalue 1:

$$v = (\alpha_1, \ldots, \alpha_n)^T / n!,$$

This is the steady-state to which the system converges after infinitely many shuffles.

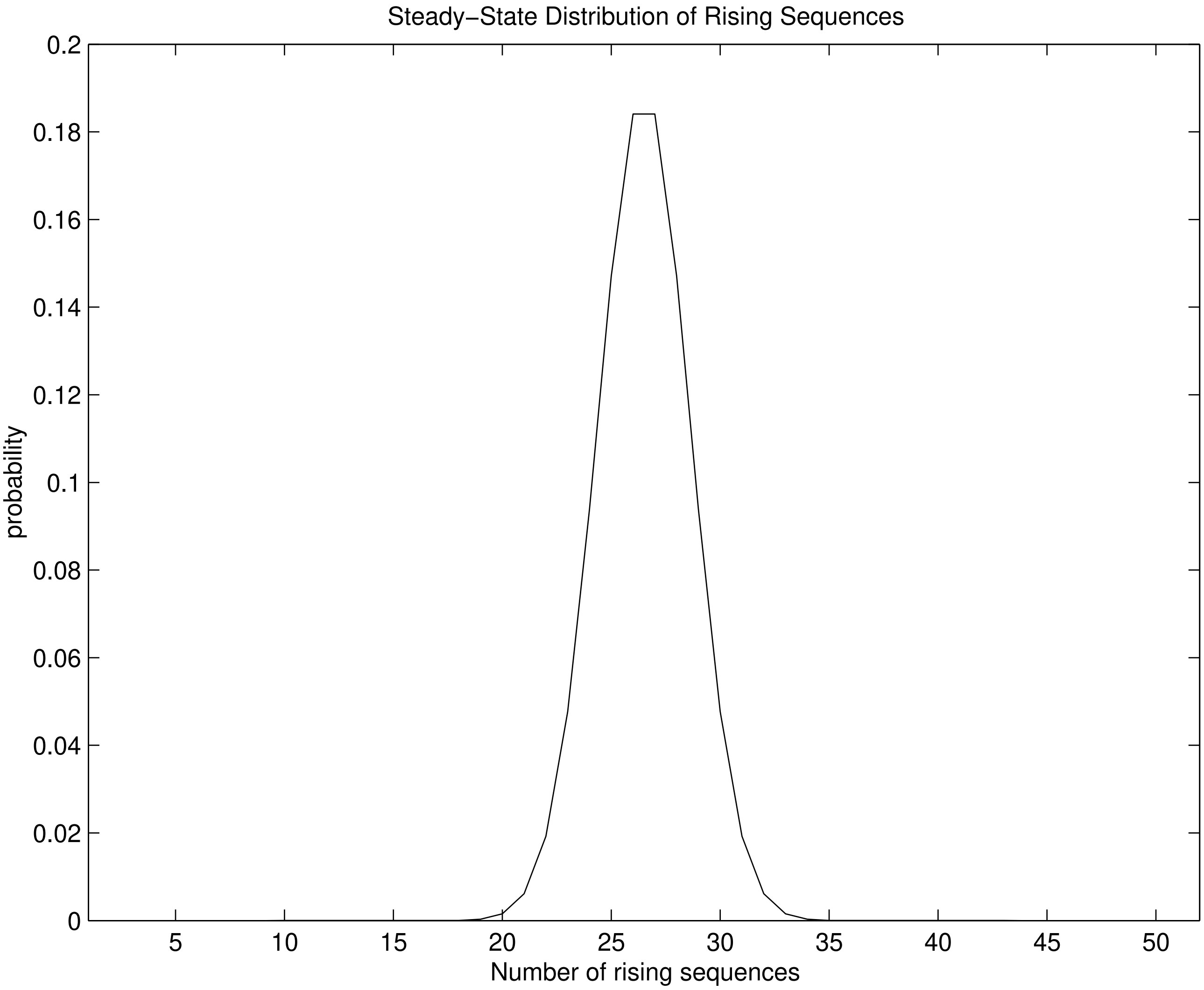


Figure 5: Steady-state distribution of rising sequences.

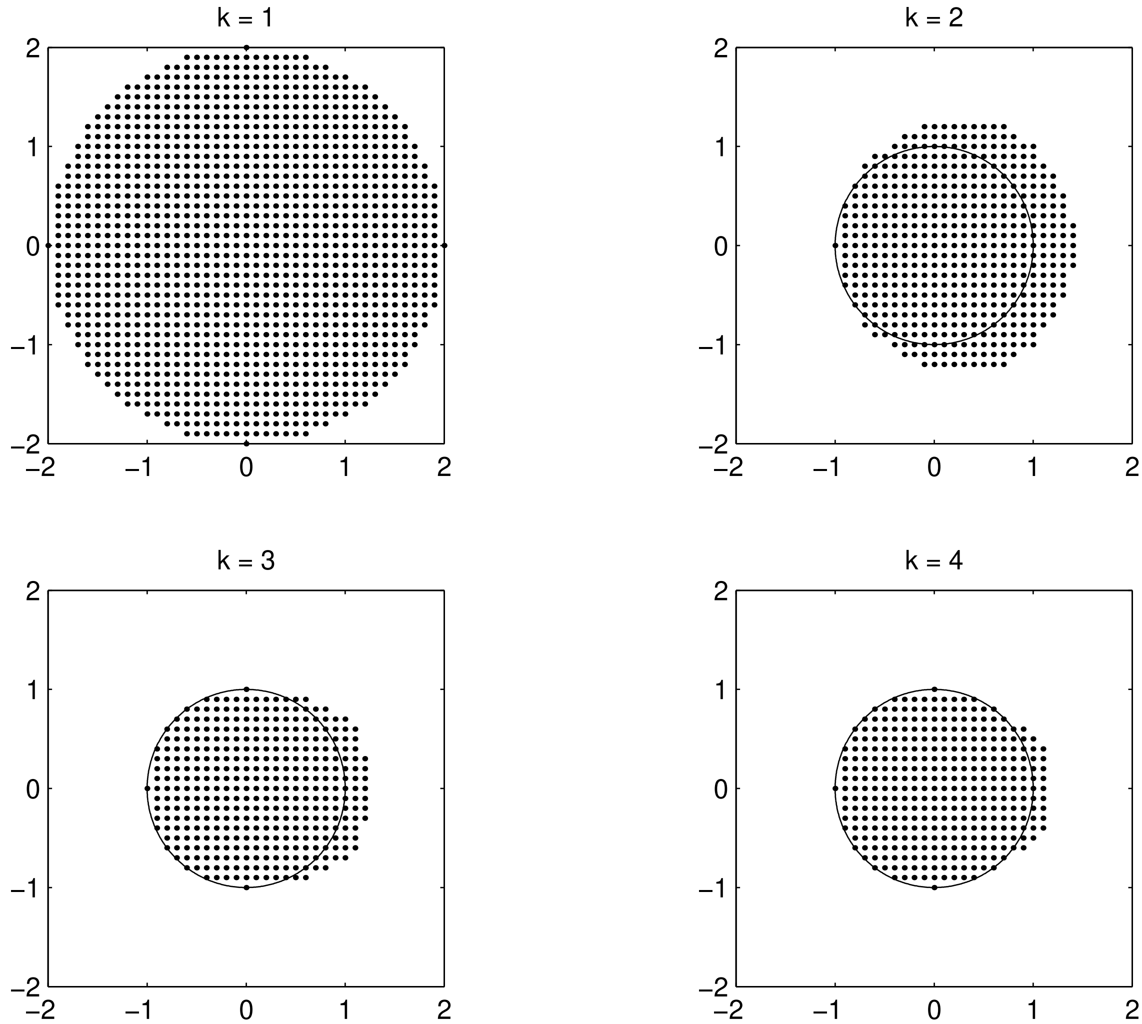


Figure 6: Polynomial numerical hulls of degrees $1-4$ for the card shuffling matrix.

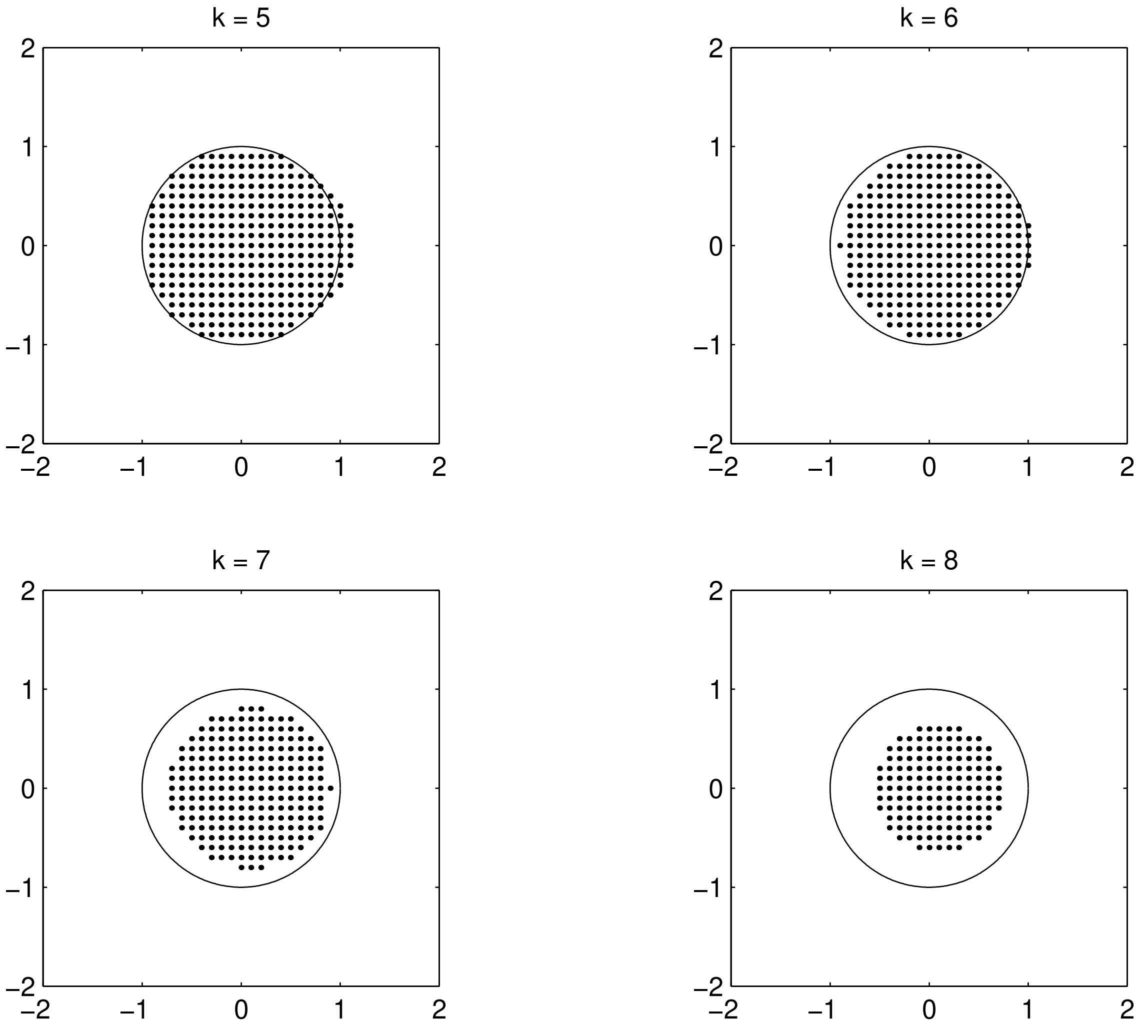


Figure 7: Polynomial numerical hulls of degrees $5 - 8$ for the card shuffling matrix. The first step at which the hull is inside the unit disk is step $k = 7$: $\nu_7 = .904$, $\nu_7^7 = .49$.

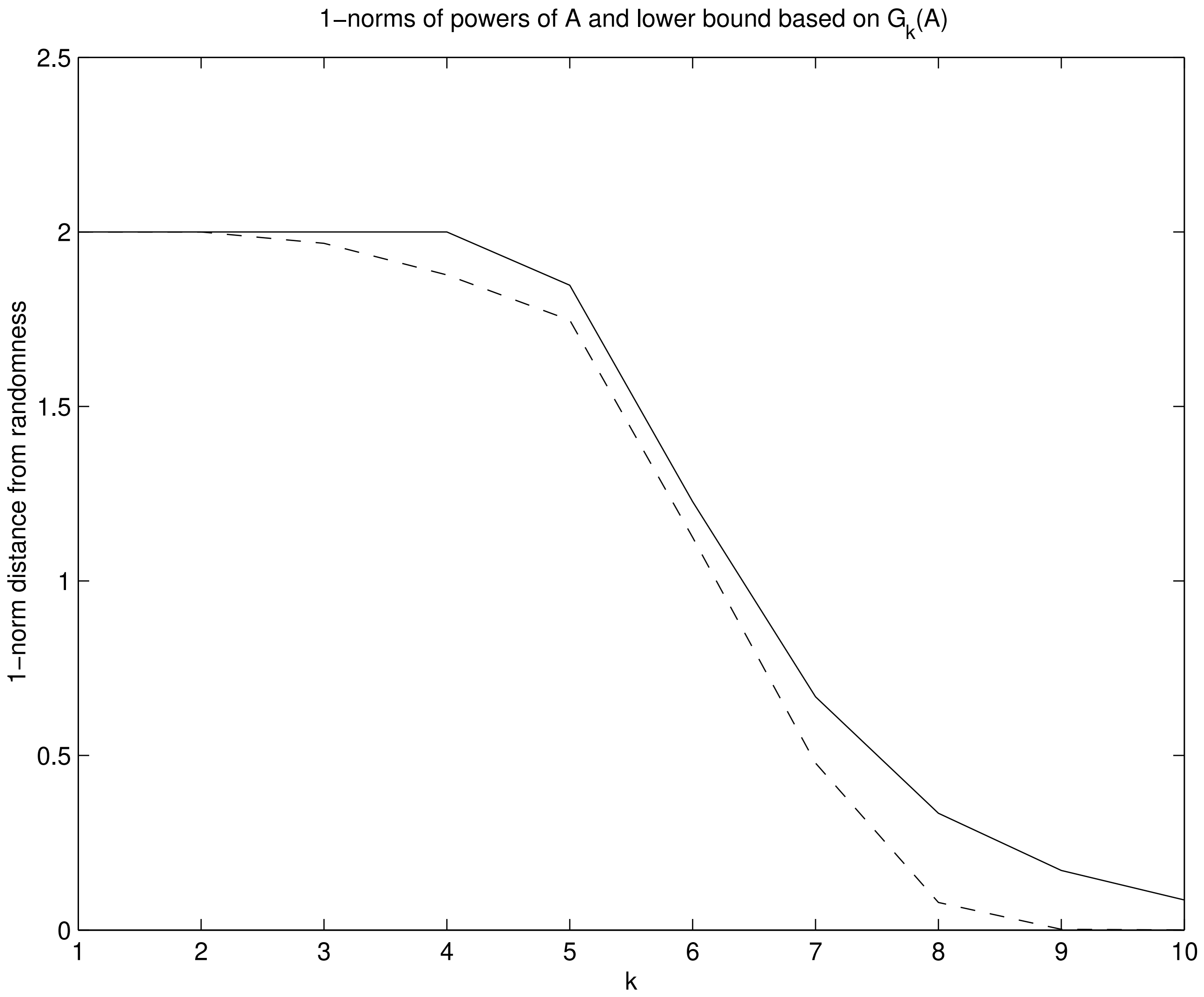


Figure 8: 1-Norm distance from randomness (solid) and lower bound based on polynomial numerical hulls of various degrees (dashed).