Electronic Transactions on Numerical Analysis.
Volume 28, pp. 16-39, 2007.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

DECAY BOUNDS AND $O(n)$ ALGORITHMS FOR APPROXIMATING FUNCTIONS OF SPARSE MATRICES*

MICHELE BENZI† AND NADER RAZOUK†

Dedicated to Gene Golub on the occasion of his 75th birthday

Abstract. We establish decay bounds for the entries of $f(A)$, where A is a sparse (in particular, banded) $n \times n$ diagonalizable matrix and f is smooth on a subset of the complex plane containing the spectrum of A. Combined with techniques from approximation theory, the bounds are used to compute sparse (or banded) approximations to $f(A)$, resulting in algorithms that under appropriate conditions have linear complexity in the matrix dimension. Applications to various types of problems are discussed and illustrated by numerical examples.

Key words. Matrix functions, sparse and banded matrices, decay rates, linear time algorithms, Chebyshev polynomials, Faber polynomials, density matrix, trace, determinant

AMS subject classifications. Primary 65F10, 65F30. Secondary 15A.

1. Introduction.

This paper is devoted to the problem of computing approximations to matrix functions $f(A)$, where A is a large, sparse (in particular, banded) matrix of order n and f is smooth on a region containing the eigenvalues of A. We are motivated by the current need for fast algorithms for approximating a wide class of matrix functions that occur in important areas of computational science and engineering, like electronic structure calculations in quantum chemistry and nanoscience. Our approach is based on the study and careful exploitation of certain decay properties exhibited by the entries of $f(A)$. We show that in many cases the entries of $f(A)$ decay very quickly in magnitude outside some band or sparsity pattern. In other words, the entries of $f(A)$ tend to be strongly localized, an observation that opens the door to the possibility of approximating relevant matrix functions in $O(n)$ time. Localization phenomena of this kind are well-known to physicists and chemists, especially those working on electronic structure calculations. With few exceptions, numerical analysts have so far largely neglected this property and its impact on the design of fast approximation algorithms. There is also a need for numerical analysts to analyze existing $O(n)$ algorithms developed by physicists and to develop new, more efficient ones. This paper intends to be a first step towards the realization of such a program.

The paper is organized as follows. In section 2 we provide some background on matrix functions and make a few observations on related work. In section 3, after recalling some basic notions from approximation theory, we study localization phenomena in matrix functions. In particular, we prove very general decay bounds for the entries of matrix functions. Our results generalize and unify several previous results known in the literature. This theory can be used, in principle, to determine the bandwidth or sparsity pattern outside of which the entries of $f(A)$ are guaranteed to be so small that they can be neglected without exceeding a prescribed error tolerance. In section 4 we discuss approximation methods based on classical approximation theory, including Chebyshev and Newton polynomials, and related approaches. Together with truncation strategies for neglecting small elements, these techniques form the basis for the approximation algorithms we use. Preliminary numerical experiments illustrating the promising behavior of the approximation schemes are presented in section 5. Finally, in section 6 we give our conclusions.

*Received March 12, 2007. Accepted for publication May 10, 2007. Recommended by A. Wathen. Work supported by National Science Foundation grant DMS-0511336.

†Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia 30322, USA (`{benzi,nrazouk}@mathcs.emory.edu`).

ETNA
Kent State University
etna@mcs.kent.edu

2. Functions of matrices. Unless otherwise specified, all matrices considered in this paper are square and complex. Given a matrix A and a function f defined on a subset of the complex plane containing the spectrum of A, it is possible to give several (equivalent) definitions of the matrix function $f(A)$; see, e.g., [31, 39, 43, 46, 68]. For instance, when f is analytic, $f(A)$ can be defined by a power series expansion, or by a contour integral (using Cauchy's Theorem). Other definitions of $f(A)$ can be given in terms of canonical forms, like the Jordan canonical form or the Schur triangular form. In this paper we limit ourselves to the case where A is diagonalizable, with $A = XDX^{-1}$ where $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, so that $f(A) = Xf(D)X^{-1}$ with $f(D) = \mathrm{diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n))$. Another useful definition is given by $f(A) = r(A)$, where r denotes the Lagrange–Hermite interpolation polynomial that interpolates f (and, in case of multiple eigenvalues, certain of its derivatives) at the eigenvalues of A. This last definition shows that *every matrix function of A is a polynomial in A*; the coefficients of the polynomial, of course depending on A (and not just on f). For the purpose of the theoretical part of this paper, the most useful definitions are those in terms of spectral decompositions and of power series expansions.

We are concerned here with the fundamental problem of computing approximations to matrix functions $f(A)$, where A is a large, sparse (in particular, banded) $n \times n$ matrix. Although the problem of computing matrix functions has received considerable attention in the literature on numerical linear algebra and differential equations, we intend to address problems and issues that so far have been largely neglected in the numerical analysis literature and yet are of considerable importance for applications and also in their own right. To explain how our approach differs from the large body of existing work, we point out that much of the literature on the approximation of matrix functions falls into two broad classes:

1. Algorithms designed for the "exact" (i.e., high-accuracy) computation of $f(A)$, where A is a dense matrix of moderate size; here f is often some elementary function such as the exponential, the sine (or cosine), the p-th root, or the sign function. See, e.g., [12, 23, 39, 41, 43, 53, 62].
2. Algorithms designed for the computation of the vector $f(A)v$, where v is a given vector and A is a large, sparse matrix. This problem is especially important in the numerical integration of large systems of time-dependent differential equations, parabolic PDEs, etc.; see, e.g., [10, 11, 13, 19, 20, 21, 25, 26, 28, 35, 44, 45, 51, 57, 69]. Most of these papers are concerned with the computation of the action of the exponential operator $\exp(-tA)$ on a vector v, where $t > 0$ and the eigenvalues of A have positive real part.

In contrast, we focus on the problem of finding approximations (to within a prescribed tolerance and for an appropriate choice of the error metric, usually dictated by the underlying application) to the actual entries of $f(A)$ when A is large and banded or sparse. The ultimate goal is to develop reliable algorithms with guaranteed accuracy and optimal complexity, when feasible. For a sparse or banded $n \times n$ matrix A, a lower bound for the best possible complexity is $O(n)$ arithmetic work and storage: i.e., the cost of the approximation, in terms of time and storage, scales at least linearly in the size of the problem. Generally speaking, attainment of this lower bound rules out the computation of $f(A)$ by reduction to diagonal or triangular (Schur) form. Also, the straightforward use of algorithms for computing $f(A)v$ applied to the case where $v = e_i$ (the ith column of the identity matrix, with $i = 1 : n$) is likely to result in $O(n^2)$ complexity, although careful implementations of inexact ("sparse-sparse") variants of this approach may work in some cases.

Clearly, $O(n)$ algorithms are feasible only if $f(A)$ can be well-approximated, in some representation, by a banded or sparse matrix, with an error that is independent of n (see section 3). It is also worth noting that in many applications not all the entries of $f(A)$ are needed, but only selected ones. For instance, it may happen that only an estimate of the trace (sum of the diagonal elements) of $f(A)$ is required. Another scalar function that is sometimes necessary to estimate is the determinant $\det[f(A)]$, and in some cases it is possible to approximate this quantity without computing all the entries of $f(A)$. Also, in many applications very high accuracy is not required or warranted.

One approach that has been used in the literature to approximate selected entries of matrix functions is the method based on Gaussian quadrature rules developed by Golub and collaborators over the years; see [38], and [4, 5, 8] for further details and some applications. While the original method was developed for real symmetric matrices, nonsymmetric extensions have been since developed; see, e.g., [15] for an approach based on the Arnoldi process. This methodology is quite different from that used in the present paper.

The key idea of the paper is to exploit the fact that for large classes of matrices A and functions f, the "important" entries of $f(A)$ are highly localized; for instance, the entries of $f(A)$ exhibit exponentially fast off-diagonal decay. We emphasize that the vast majority of the algorithms currently in vogue among numerical analysts for the computation of $f(A)$ do not exploit this property. Among the few exceptions, see [4, 6, 7, 21, 47, 67]; however, these papers all address some very special problems, and a systematic treatment of sparsity in matrix functions is lacking. We also mention that chemists and physicists who work in the field of electronic structure calculations have observed the exponential decay property in the density matrix (a function of the Hamiltonian) for non-metallic systems, see [3, 37, 42, 52, 59]. (For metallic systems, the decay is algebraic rather than exponential, see [36].) Physicists have exploited this property by developing "sparsified" iterative algorithms for the approximation of density matrices. Partial results and a few algorithms tailored to specific matrix functions can be found scattered in the literature, but no general theory or rigorous foundations for many of the algorithms developed by physicists exist. There is, in other words, little available theory for such approximation algorithms; moreover, error estimates and safe stopping criteria have yet to be developed, and optimality results are also missing. Finally, the existing $O(n)$ algorithms are applicable to the Hermitian case only, and there is an interest for methods that can handle non-Hermitian matrices as well. In this paper we begin to address some of these questions.

3. Decay bounds.

Many physical phenomena are characterized by strong localization, that is, rapid decay outside of a small spatial or temporal region. Frequently, such localization can be expressed as decay in the entries of a function $f(A)$ of an appropriate sparse or banded matrix A that encodes a description of the system under study. In this section we survey the current state of knowledge on this topic and contribute some new results that extend and unify previously known ones. We begin with a simple graph-theoretical result on the sparsity pattern of $f(A)$, according to which such a matrix is generally full. Then we consider banded approximations to matrices that exhibit exponential off-diagonal decay of the entries and discuss situations where $O(n)$ approximations are feasible. Next, we provide some necessary background on approximation theory and finally we prove new exponential decay bounds for a wide class of functions f and matrices A.

3.1. A structural result.

Let $A = [a_{ij}]$ be an $n \times n$ matrix. We associate to A a digraph $G(A) = (V, E)$ in the standard way (we refer to [24] for basic background on graph theory). The vertex set V consists of the integers from 1 to n and the edge set E contains all ordered pairs (i, j) with $a_{ij} \neq 0$. Recall that for $i, j \in V$ the *distance* $d(i, j)$ between i and j is the length of the shortest directed path connecting node i to node j; the maximum distance between any two nodes in $G(A)$ is the *diameter* of the digraph. Proposition 3.1 states a simple

ETNA
Kent State University
etna@mcs.kent.edu

structural result about $f(A)$ for a large class of analytic functions f.

PROPOSITION 3.1. *Let f be an analytic function of the form*

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k \qquad \left(a_k = \frac{f^{(k)}(z_0)}{k!} \right),$$

where $z_0 \in \mathbb{C}$ and the power series expansion has radius of convergence $r > 0$. Let A have an irreducible sparsity pattern with nonzero diagonal entries and let l ($1 \le l \le n-1$) be the diameter of $G(A)$. Assume further that there exists $k \ge l$ such that $f^{(k)}(z_0) \ne 0$. Then it is possible to assign values to the nonzero entries of A in such a way that $f(A)$ is defined and structurally full.

Proof. We first observe that it is always possible to assign numerical values to the entries of A so that the spectrum of A lies inside the circle of convergence of the power series. Indeed, any matrix can be shifted and scaled so that the eigenvalues λ_i of the resulting matrix satisfy $|\lambda_i - z_0| < r$, for $i = 1 : n$. Note that shifting and scaling does not affect the irreducibility of the sparsity pattern. Thus, the power series

$$a_0 I + a_1(A - z_0 I) + a_2(A - z_0 I)^2 + \cdots + a_k(A - z_0 I)^k + \cdots$$

with $a_k = f^{(k)}(z_0)/k!$ converges to $f(A)$. Since A is irreducible, there is a directed path in the digraph of A from any node i to any other node j. From the assumption that the maximum distance between any two nodes has length l we conclude that $(A - z_0 I)^k$ for $k \ge l$ is structurally full, since there is an entry for each i, j with $d(i,j) \le l$ in $G(A)$. Finally, the assumption that $f^{(k)}(z_0) \ne 0$ for at least one $k \ge l$ implies that $f(A)$ is structurally full. $\square$

The above result means that barring fortuitous cancellation, all entries of $f(A)$ are nonzero when A is irreducible (with a nonzero diagonal) and f is a "generic" analytic function, by which we mean an analytic function with at least one non-vanishing derivative of sufficiently high order. We note that this result generalizes a well-known fact about the inverse of an irreducible matrix [27].

Thus, it would seem that the quest for $O(n)$ algorithms for approximating $f(A)$ is doomed from the start. However, as we will see, it is often the case that most of the entries in $f(A)$ are very small in magnitude, thus making banded or sparse approximations possible.

3.2. Banded approximations to highly localized matrices. Here we consider approximating a dense matrix with certain decay properties by means of a banded matrix. The problem is important for several reasons. First, there are situations where the matrix A itself is completely dense, but has rapidly decaying entries; matrices of this kind are sometimes called *pseudosparse*. For example, a matrix representing a differential operator with respect to a basis made of globally supported but highly localized functions will typically be pseudosparse. It is common practice to replace A with a banded matrix $\tilde{A}$ obtained by "sparsifying" (or "truncating") the matrix outside a certain band, or by dropping entries whose magnitude falls under a certain (small) threshold. As a preliminary condition for developing linear scaling algorithms, it is essential to show that the *sparsification error* $\|A - \tilde{A}\|$ may be bounded, for an appropriate norm, by a quantity independent of the order n of the matrix; here we show that this is indeed the case when the entries of A satisfy a certain exponential decay condition. Also, in order to approximate a matrix function $f(A)$, we actually compute an approximation of $f(\tilde{A})$. How far this is from the actual $f(A)$ depends on several factors, including of course on how far $\tilde{A}$ is from A; see [50] for a general theory of condition estimation for matrix functions. It is therefore desirable to bound the sparsification error in terms of the rate of

ETNA
Kent State University
etna@mcs.kent.edu

decay of the entries of A. Furthermore, the bounds are useful in estimating the error in the approximation of the generally full matrix $f(\tilde{A})$ by a sparse or banded matrix.

We say that an $n \times n$ matrix $A = [a_{ij}]$ has the *exponential off-diagonal decay property* if there is a constant $c > 0$ such that

$$|a_{ij}| \leq c\lambda^{|i-j|}, \quad \text{where} \quad 0 < \lambda < 1, \tag{3.1}$$

for all $i, j = 1 : n$. Corresponding to the matrix A we then define for a nonnegative integer m the matrix $A^{(m)} = \left[a_{ij}^{(m)}\right]$ defined as follows:

$$a_{ij}^{(m)} = \begin{cases} a_{ij}, & \text{if } |i-j| \leq m; \\ 0, & \text{otherwise.} \end{cases}$$

The matrix $A^{(m)}$ is m-banded, with m being the bandwidth of $A^{(m)}$; according to this definition, diagonal matrices have zero bandwidth, tridiagonal ones have bandwidth $m = 1$, etc. Note that we do not require the matrix to be symmetric here, we only assume (for simplicity) that the same pattern of nonzero off diagonals is present on either side of the main diagonal. The following simple result provides an estimate of the rate at which the sparsification error decreases as the bandwidth m of the approximation increases. In addition, it establishes n-independence when the constants c and λ in (3.1) are independent of n. For convenience, we work with the 1-norm.

PROPOSITION 3.2. *Let A be a matrix with entries a_{ij} satisfying (3.1) and let $A^{(m)}$ be the corresponding m-banded approximation. Then for any $\varepsilon > 0$ there is an $\overline{m}$ such that $\|A - A^{(m)}\|_1 \leq \varepsilon$ for $m \geq \overline{m}$.*

Proof. We have

$$\|A - A^{(m)}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij} - a_{ij}^{(m)}| \leq \sum_{k=m+1}^{n} c\lambda^k \leq \sum_{k=m+1}^{\infty} c\lambda^k = \frac{c\lambda^{m+1}}{1-\lambda}.$$

Since $0 < \lambda < 1$, for any given $\varepsilon > 0$ we can always make

$$\frac{c\lambda^{m+1}}{1-\lambda} \leq \varepsilon$$

by taking m sufficiently large. $\square$

The result of Proposition 3.2 also holds for matrices with other types of decay behavior, for instance algebraic decay of the form $|a_{ij}| \leq c|i-j|^{-p}$ with $p > 1$. Moreover, the argument can be adapted to handle sparse matrices rather than banded ones, along the lines to be discussed in section 3.4.

Clearly, Proposition 3.2 is not very useful when applied to a single $n \times n$ matrix A since it is obvious that taking $m = n$ always results in a zero error. The result, however, shows that if we have a sequence $\{A_n\}$ of matrices of increasing dimension n, with entries $[A_n]_{ij}$ satisfying a *uniform* exponential off-diagonal decay property of the type

$$|[A_n]_{ij}| \leq c\lambda^{|i-j|}, \quad \text{where} \quad 0 < \lambda < 1, \quad \text{with } c \text{ and } \lambda \text{ independent of } n,$$

then for any given $\varepsilon > 0$ it is possible to find a bandwidth m independent of n such that the m-banded approximations $A_n^{(m)}$ satisfy $\|A_n^{(m)} - A_n\|_1 < \varepsilon$, for all n.

We mention that although in this section we have limited our discussion to absolute errors relative to a certain matrix norm, in practice we adopt relative error measures when assessing the distance of a computed approximation from a desired quantity.

ETNA
Kent State University
etna@mcs.kent.edu

3.3. Faber polynomials and series. In this section we provide some background on approximation theory that will be needed in proving general decay bounds for the entries of functions of sparse or banded matrices. Our treatment follows closely that in [58], but see also [18, 75]. In the following, F denotes a continuum containing more than one point. By a *continuum* we mean a nonempty, compact and connected subset of $\mathbb{C}$. Let G_∞ denote the component of the complement of F containing the point at infinity. Note that G_∞ is a simply connected domain in the extended complex plane $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. By the Riemann Mapping Theorem there exists a function $w = \Phi(z)$ which maps G_∞ conformally onto a domain of the form $|w| > \rho > 0$ satisfying the normalization conditions

$$\Phi(\infty) = \infty, \quad \lim_{z\to\infty} \frac{\Phi(z)}{z} = 1; \tag{3.2}$$

ρ is the *logarithmic capacity* of F. Given any integer $N > 0$, the function $[\Phi(z)]^N$ has a Laurent series expansion of the form

$$[\Phi(z)]^N = z^N + \alpha_{N-1}^{(N)} z^{N-1} + \cdots + \alpha_0^{(N)} + \frac{\alpha_{-1}^{(N)}}{z} + \cdots \tag{3.3}$$

at infinity. The polynomials

$$\Phi_N(z) = z^N + \alpha_{N-1}^{(N)} z^{N-1} + \cdots + \alpha_0^{(N)}$$

consisting of the terms with nonnegative powers of z in the expansion (3.3) are called the *Faber polynomials* generated by the continuum F.

Let Ψ be the inverse of Φ. By C_R we denote the image under Ψ of a circle $|w| = R > \rho$. The (Jordan) region with boundary C_R is denoted by $I(C_R)$. By Theorem 3.17, p. 109 of [58], every function $f(z)$ analytic on $I(C_{R_0})$ with $R_0 > \rho$ can be expanded in a series of Faber polynomials:

$$f(z) = \sum_{k=0}^{\infty} \alpha_k \Phi_k(z), \tag{3.4}$$

where the series converges uniformly inside $I(C_{R_0})$. The coefficients are given by

$$\alpha_k = \frac{1}{2\pi i} \int_{|w|=R} \frac{f(\Psi(w))}{w^{k+1}} dw,$$

where $\rho < R < R_0$. We denote the partial sums of the series in (3.4) by

$$\Pi_N(z) := \sum_{k=0}^{N} \alpha_k \Phi_k(z). \tag{3.5}$$

Each $\Pi_N(z)$ is a polynomial of degree at most N, since each $\Phi_k(z)$ is of degree k. We are now ready to state a result that will be instrumental in our proof of the decay bounds.

THEOREM 3.3. *(Bernstein's Theorem) Let f be a function defined on F. Then given any $\varepsilon > 0$ and any integer $N \geq 0$, there exists a polynomial Π_N of degree at most N and a positive constant $c(\varepsilon)$ such that*

$$|f(z) - \Pi_N(z)| < c(\varepsilon)(q+\varepsilon)^N \quad (0 < q < 1) \tag{3.6}$$

for all $z \in F$ if and only if f is analytic on the domain $I(C_{R_0})$, where $R_0 = \rho/q$. In this case, the sequence $\{\Pi_N\}$ converges uniformly to f inside $I(C_{R_0})$ as $N \to \infty$.

ETNA
Kent State University
etna@mcs.kent.edu

In the special case where F is a disk of radius ρ centered at z_0, Theorem 3.3 states that for any function f analytic on the disk of radius ρ/q centered at z_0, where $0 < q < 1$, there exists a polynomial Π_N of degree at most N and a positive constant $c(\varepsilon)$ such that for any $\varepsilon > 0$,

$$|f(z) - \Pi_N(z)| < c(\varepsilon)(q+\varepsilon)^N,$$

for all $z \in F$. We are primarily concerned with the sufficiency part of Theorem 3.3. Note that the choice of q (with $0 < q < 1$) depends on the region where the function f is analytic. If f is defined on a continuum F with logarithmic capacity ρ then we can pick q bounded away from 1 as long as the function is analytic on $I(C_{\rho/q})$. Therefore, the rate of convergence is directly related to the properties of the function f, such as the location of its poles (if there are any). Following [58], the constant $c(\varepsilon)$ can be estimated as follows. Let R_0, q and ε be given as in Theorem 3.3. Furthermore, let R' and R be chosen such that $\rho < R' < R < R_0$ and

$$\frac{R'}{R} = q + \varepsilon,$$

then we define

$$M(R) = \max_{z \in C_R} |f(z)|.$$

An estimate for the value of $c(\varepsilon)$ is asymptotically (i.e., for sufficiently large N) given by

$$c(\varepsilon) \approx \frac{3}{2} M(R) \frac{q+\varepsilon}{1-(q+\varepsilon)}\,.$$

It may be necessary to replace the above expression for $c(\varepsilon)$ by a larger one to obtain validity of the bound (3.6) for all N. However, for certain regions F it is possible to obtain an explicit constant valid for all $N \geq 0$; see [29] and section 3.7 below.

3.4. General decay bounds. Simple numerical experiments in MATLAB reveal that if A is a band symmetric matrix (e.g., tridiagonal) and f is a smooth function on an interval containing the spectrum of A, then most of the (relatively) large entries of $f(A)$ are concentrated on or near the nonzero diagonals of A; outside the band, the entries of $f(A)$ tend to drop off rather quickly, albeit perhaps non-monotonically, and far from the main diagonal their magnitude becomes negligible. In other words, $f(A)$ exhibits the off-diagonal decay property. Similar behavior is observed for matrices with more general sparsity patterns, with most of the large entries in $f(A)$ located near the nonzero entries of A. This phenomenon is well-known to physicists, who have referred to it as the *nearsightedness principle*, see [52]. Further experiments suggest that a similar phenomenon holds for other types of diagonalizable matrices, including non-normal ones and often even for matrices whose eigenvectors are very far from being orthogonal. In Figs. 3.1–3.2 we display "city plots" demonstrating the rapid decay in the square root and logarithm of two sparse symmetric positive definite matrices, `nos4` and `bcsstk03`, available from the Matrix Market (http://math.nist.gov/MatrixMarket/). Prior to the computations, the matrices have been divided by their largest element and reordered with reverse Cuthill–McKee.

In the mathematical literature, the decay property has been first noticed for the inverse of banded, symmetric positive definite matrices; the seminal paper is [22], but see also [9, 40, 49, 60, 65]. The main result is an exponentially decaying upper bound of the form

$$|[A^{-1}]_{ij}| \leq c\lambda^{|i-j|}, \quad \text{where} \quad c > 0 \quad \text{and} \quad 0 < \lambda < 1 \quad (1 \leq i,j \leq n).$$

ETNA
Kent State University
etna@mcs.kent.edu

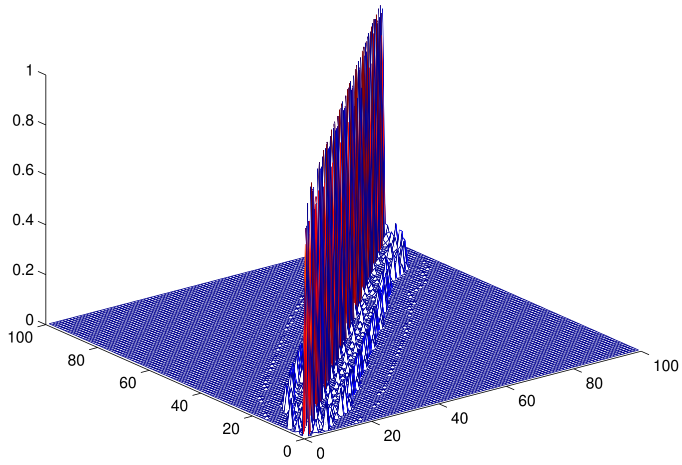

FIG. 3.1. *City plot for the square root of* `nos4`.

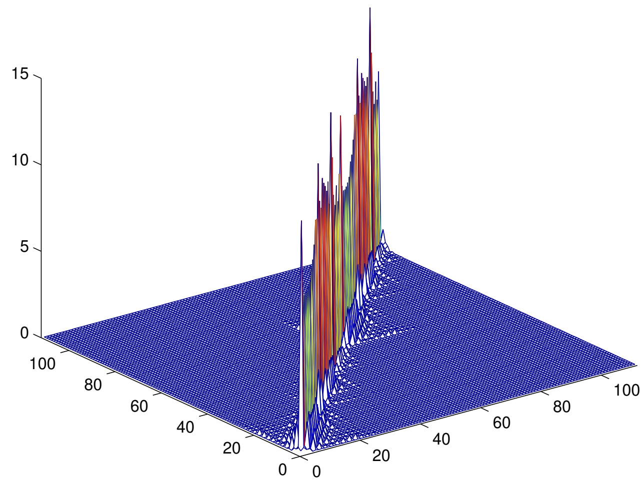

FIG. 3.2. *City plot for the logarithm of* `bcsstk03`.

Here c and λ depend on the bandwidth and extreme eigenvalues of A, with decay being faster when the bandwidth is small and the matrix well-conditioned. The decay bound can be generalized to some extent to nonsymmetric (diagonalizable) matrices as well. Similar decay results have also been obtained for the resolvent $(zI - A)^{-1}$, the inverse square root $A^{-1/2}$, and the exponential $\exp(A)$; see [48, 49, 66, 74]. In [8], Gene Golub and the first author have obtained a general decay result of the form

$$|[f(A)]_{ij}| \leq c\lambda^{|i-j|}, \quad \text{where} \quad c > 0 \quad \text{and} \quad 0 < \lambda < 1\,,$$

ETNA
Kent State University
etna@mcs.kent.edu

where A is banded symmetric and f is analytic on an ellipse containing the spectrum of A. This result is based on a classical result of Bernstein (different from Theorem 3.3) combined with the spectral theorem. The constants c and λ depend on the bandwidth of A and on the parameters of the ellipse and can be estimated explicitly. Some generalizations and applications of this result can be found in the quantum information processing literature; see [16, 17, 72]. We further mention the decay bounds for functions of banded skew-symmetric matrices given in [21, 56] which, however, were obtained using different techniques from ours.

In the next sections we provide some further generalizations of these known results that extend and unify most existing bounds, using as our basic tool Theorem 3.3 above.

3.5. Decay rates for the entries of diagonalizable matrices. The decay bounds are obtained by translating Bernstein's Theorem 3.3 in terms of matrices, via matrix diagonalization. Our treatment follows closely the ones in [8, 22]. Let $A = [a_{ij}]$ be a diagonalizable matrix and let $G(A) = (V, E)$ be the digraph associated with A. Recall that for $i, j \in V$, $d(i, j)$ denotes the length of the shortest directed path connecting node i with node j. (If there is no directed path from i to j, we set $d(i, j) = \infty$.) We note that this graph-theoretic 'distance' $d(i, j)$ is not necessarily a true distance, since in general $d(i, j) \neq d(j, i)$, unless A is structurally symmetric. Clearly, whenever $d(i, j) > 1$ we have that $a_{ij} = 0$. More generally, it is easy to see that $a_{ij}^{(k)} = 0$ whenever $d(i, j) > k$ for $k = 0, 1, \ldots$, where $a_{ij}^{(k)}$ denotes the (i, j) entry in the matrix A^k. To show this, we use induction over k and note that the situation is trivial for $k = 1$. Consider now the following identity:

$$a_{ij}^{(k+1)} = \sum_l a_{il}^{(k)} a_{lj}.$$

For pairs (i, l) such that $d(i, l) > k$ we have that $a_{il}^{(k)} = 0$, by the inductive hypothesis. For pairs with $d(i, l) \le k$ we have that

$$k + 1 < d(i, j) \le d(i, l) + d(l, j) \le k + d(l, j)$$

and therefore $d(l, j) > 1$, hence $a_{lj} = 0$.

We are now in a position to state our general decay result. In the following, $\sigma(A)$ denotes the spectrum of A.

THEOREM 3.4. *Let A be diagonalizable and assume that $\sigma(A)$ is contained in a continuum F. Let $\kappa(X)$ denote the spectral condition number of the matrix X of eigenvectors of A. Let f be a function defined on F. Furthermore, assume that f is analytic on $I(C_{R_0})$ $(\supset \sigma(A))$, where $R_0 = \frac{\rho}{q}$ with $0 < q < 1$ and ρ is the logarithmic capacity of F. Then there are positive constants K and $0 < \lambda < 1$ such that*

$$|[f(A)]_{ij}| < \kappa(X) K \lambda^{d(i,j)} \quad \text{for all} \quad i, j = 1 : n. \tag{3.7}$$

Proof. From Theorem 3.3 we know that for any $\varepsilon > 0$ there exists a sequence of polynomials Π_k of degree k which satisfies for all $z \in F$

$$|f(z) - \Pi_k(z)| < c(\varepsilon)(q + \varepsilon)^k, \quad \text{where} \quad 0 < q < 1.$$

We therefore have, since $A = XDX^{-1}$, that

$$\|f(A) - \Pi_k(A)\|_2 \le \kappa(X) \max_{z \in \sigma(A)} |f(z) - \Pi_k(z)| < \kappa(X) c(\varepsilon)(q + \varepsilon)^k,$$

ETNA
Kent State University
etna@mcs.kent.edu

where $0 < q < 1$. For $i \neq j$ we can write

$$d(i,j) = k + 1$$

and therefore, observing that $[\Pi_k(A)]_{ij} = 0$ for $d(i,j) > k$, we have

$$|[f(A)]_{ij}| = |[f(A)]_{ij} - [\Pi_k(A)]_{ij}| \leq \|f(A) - \Pi_k(A)\|_2 < \kappa(X)c(\varepsilon)(q+\varepsilon)^{d(i,j)-1}.$$

Hence, choosing $\varepsilon > 0$ such that $\lambda = q + \varepsilon < 1$ and letting $K_0 = c(\varepsilon)/(q+\varepsilon)$ we obtain

$$|[f(A)]_{ij}| < \kappa(X)K_0\lambda^{d(i,j)}.$$

If $i = j$ then $|[f(A)]_{ii}| \leq \|f(A)\|_2$ and therefore, letting $K = \max\{K_0, \|f(A)\|_2\}$, we see that inequality (3.7) holds for all i, j. ☐

It is clear that when the matrix A is dense, the bound (3.7) is not very meaningful since the digraph $G(A)$ will be dense as well. Hence, in practice the result is of some interest only in the (pseudo)sparse case. Also note that if A is normal, then X can be chosen to be unitary and therefore $\kappa(X) = 1$. This covers for example the Hermitian and skew-Hermitian cases; we have thus obtained a single proof for results that had been previously proved using different tools. It should be clear from the proof of this result that in general one should not expect the bound to be very sharp. In particular, when A is highly non-normal (in the sense that $\kappa(X) \gg 1$) the bound can be a severe overestimate of the actual size of the elements of $f(A)$, and in this case other tools, such as for example pseudospectra, may be better suited to characterize the actual decay behavior; see [78, 79].

3.6. The banded case. We now specialize the result to the case of band matrices. As A is not assumed to have symmetric structure, it is necessary to account for the fact that for a nonsymmetric matrix, the rate of decay can be different above and below the main diagonal. For instance, A could have different numbers of nonzero diagonals below and above the main diagonal. An extreme case is when A is an upper (lower) Hessenberg matrix, in which case $f(A)$ typically exhibits fast decay below (above) the main diagonal, and generally no decay above (below) it. In Figure 3.3 we display on the left the magnitude of the entries of $f(A) = \cos(A)$ where A is a band matrix with one nonzero diagonal below and four nonzero diagonals above the main diagonal. The matrix is diagonalizable, but the condition number $\kappa(X)$ of the eigenvector matrix is of the order of 10^{10}. The plot on the right shows the magnitude of the entries on the 50th row of $\cos(A)$.

We establish the appropriate decay results next. We give the proof of the bound for entries in the lower part of $f(A)$ only; the proof for the entries in the upper part of $f(A)$ is identical, with the obvious changes.

We say that a matrix A has *lower bandwidth* $p > 0$ if $a_{ij} = 0$ whenever $i - j > p$ and *upper bandwidth* $s > 0$ if $a_{ij} = 0$ whenever $j - i > s$. We note that if A has lower bandwidth p then A^k has lower bandwidth kp for $k = 0, 1, 2, \ldots$, and similarly for the upper bandwidth s.

THEOREM 3.5. *Let A be diagonalizable with upper bandwidth p and lower bandwidth s. Assume that $\sigma(A)$ is contained in a continuum F. Let f be a function defined on F. Furthermore, assume that f is analytic on $I(C_{R_0})$ $(\supset \sigma(A))$, where $R_0 = \frac{\rho}{q}$ with $0 < q < 1$ and ρ is the logarithmic capacity of F. Then there are positive constants K, C and $0 < \lambda, \eta < 1$ such that for $i \geq j$,*

$$|[f(A)]_{ij}| < \kappa(X)K\lambda^{i-j}, \tag{3.8}$$

and for $i < j$,

$$|[f(A)]_{ij}| < \kappa(X)C\eta^{j-i}. \tag{3.9}$$

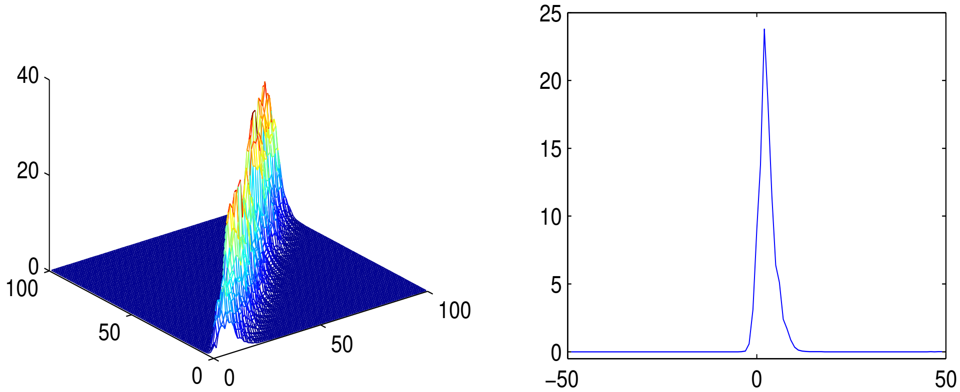

FIG. 3.3. *Cosine of a nonsymmetric band matrix.*

Proof. Observe first that for $k = 0, 1, \ldots$ the matrix $\Pi_k(A)$ has lower bandwidth kp for all polynomials $\Pi_k(z)$ of degree at most k. From Bernstein's Theorem we know that for any $\varepsilon > 0$ there exists a sequence of polynomials $\Pi_k(z)$ of degree k which satisfies for all $z \in F$

$$|f(z) - \Pi_k(z)| < c(\varepsilon)(q + \varepsilon)^k, \quad \text{where} \quad 0 < q < 1.$$

We therefore have since $A = XDX^{-1}$ that

$$\|f(A) - \Pi_k(A)\|_2 \leq \kappa(X) \max_{z\in\sigma(A)} |f(z) - \Pi_k(z)| < \kappa(X)c(\varepsilon)(q + \varepsilon)^k,$$

where $0 < q < 1$. For $i > j$ write

$$i - j = kp + l \quad \text{for} \quad l = 1, \ldots, p;$$

then we have that

$$\frac{i - j}{p} - 1 \leq k$$

and hence observing that $[\Pi_k(A)]_{ij} = 0$ for $i - j > pk$,

$$|[f(A)]_{ij}| = |[f(A)]_{ij} - [\Pi_k(A)]_{ij}| \leq \|f(A) - \Pi_k(A)\|_2 < \kappa(X)c(\varepsilon)(q + \varepsilon)^{\frac{i-j}{p}-1}.$$

Hence, letting $\lambda = (q + \varepsilon)^{\frac{1}{p}}$ and $K_0 = c(\varepsilon)/(q + \varepsilon)$ we obtain

$$|[f(A)]_{ij}| < \kappa(X)K_0\lambda^{i-j}.$$

If $i = j$ then $|[f(A)]_{ii}| \leq \|f(A)\|_2$ and therefore, letting $K = \max\{K_0, \|f(A)\|_2\}$, we see that inequality (3.8) holds for all $i \geq j$. ◻

An immediate corollary of Theorem 3.5 is the following.

COROLLARY 3.6. *Let A be a banded diagonalizable matrix. Assume that $\sigma(A)$ is contained in a continuum F. Let f be a function defined on F. Furthermore, assume that f is analytic on $I(C_{R_0})$ ($\supset \sigma(A)$), where $R_0 = \frac{\rho}{q}$ with $0 < q < 1$ and ρ is the logarithmic capacity of F. Then there are positive constants K and $0 < \lambda < 1$ such that for all i,j=1:n,*

$$|[f(A)]_{ij}| < \kappa(X)K\lambda^{|i-j|}.$$

We note that for A real and symmetric, this is essentially Theorem 2.2 in [8].

ETNA
Kent State University
etna@mcs.kent.edu

3.7. More refined bounds. The bounds obtained in the previous section can be improved by using sharper estimates on the Faber polynomials, provided that some (fairly reasonable) additional conditions are imposed on the continuum F. As already mentioned at the end of section 3.3, in [29] the author presents more quantitative estimates on the Nth Faber polynomial for certain choices of F. Specifically, assume that F is a closed *Jordan region*. By a Jordan region we mean a region F that is bounded and whose boundary Γ consists of pairwise disjoint closed Jordan curves. If Γ is rectifiable, there exists at almost every point $z \in \Gamma$ a tangent vector that makes an angle $\Theta(z)$ with the positive real axis. We say that Γ has *bounded total rotation* V if

$$V = \int_\Gamma |d\Theta(z)| < \infty.$$

It is apparent that $V \geq 2\pi$ and equality holds if F is convex. In the case where F degenerates to a *Jordan arc*, the following results remain valid; we refer the reader to [29] for details.

THEOREM 3.7. *Let F be a Jordan region whose boundary Γ is of bounded total rotation V. Then the following holds:*

(a) For $N \geq 1$,

$$\|\Phi_N\|_\infty \leq \rho^N V/\pi.$$

This bound holds with equality when F coincides with the interval $[-1, 1]$.

(b) Let f be analytic on the closed region C_R, $R > \rho$, then for $N \geq 0$ we have

$$\|f - \Pi_N\|_\infty \leq \frac{M(R)V}{\pi} \frac{(\rho/R)^{N+1}}{1 - \rho/R},$$

where $M(R) = \max_{z \in C_R} |f(z)|$ and $\Pi_N(z)$ are the polynomials defined by equation (3.5).

These bounds can be used to improve our earlier estimates in the case when F and f satisfy the conditions of Theorem 3.7. If in addition A satisfies the conditions of Theorem 3.4, we can obtain the following bound for the entries of $f(A)$:

$$|[f(A)]_{ij}| \leq \kappa(X) K \lambda^{d(i,j)}, \tag{3.10}$$

where now $K = \max\left\{\|f(A)\|_2, \frac{M(R)V}{\pi(1-\rho/R)}\right\}$ and $\lambda = \rho/R$.

We now compare the above bound on the entries of $f(A)$ with the magnitude of the actual entries $[f(A)]_{ij}$, for $i < j$, in a special case. The matrix A is tridiagonal, with spectrum in $F = [-1, 1]$. In this case $\rho = 1/2$ and since F is convex, we have that $V = 2\pi$. The function f must be analytic on the ellipse

$$\frac{x^2}{\left(\frac{R}{2\rho} + \frac{\rho}{2R}\right)^2} + \frac{y^2}{\left(\frac{R}{2\rho} - \frac{\rho}{2R}\right)^2} = 1.$$

This is an ellipse with foci at ± 1, the sum of whose semiaxes is R/ρ. For our experiment we take $f(z) = \mathrm{e}^z$, which is entire and therefore satisfies our assumptions. The constant $M(R)$ is then given by

$$M(R) = \mathrm{e}^{R + \frac{1}{4R}}.$$

ETNA
Kent State University
etna@mcs.kent.edu

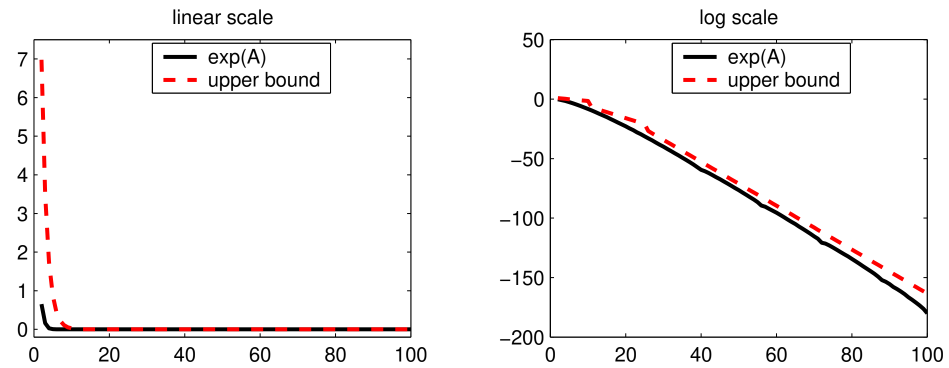

FIG. 3.4. *Upper bound (3.10) vs. the entries of* $\exp(A)$*, first row.*

It is important to note that (3.10) gives us a bound that changes with R. We can use small values of R to estimate the elements on the first few off-diagonals, and increase R as we move away from the diagonal. This adaptive strategy allows us to obtain sharper bounds than those described in the previous section. In Figure 3.4 we show the upper bound compared to the actual magnitude of the off-diagonal entries in the first row of $f(A) = \exp(A)$, in both linear and logarithmic scales. Clearly, the bound does a good job capturing the decay behavior, especially as one moves away from the main diagonal.

The decay bounds obtained so far can be used to predict a minimum bandwidth (more generally, sparsity pattern) for the computed approximation to $f(A)$ to satisfy a prescribed error tolerance; if the rate of decay, expressed in terms of the various quantities entering the upper bound, does not depend on n, then an $O(n)$ approximation is feasible. A sufficient condition is that the $n \times n$ matrix sequence $\{A_n\}$ is such that for all n, the spectra $\sigma(A_n)$ are all contained in a region that is bounded and bounded away from the singularities of f (if any); furthermore, the condition numbers of the corresponding eigenvector matrices must be bounded independent of n. Although these conditions seem rather restrictive, they are satisfied by large classes of matrices and functions that arise in applications, for example when f is entire and $\{A_n\}$ is a sequence of normal (e.g., Hermitian) matrices with bounded spectra. If f is not entire (for example if $f(z) = z^{-1}$, $f(z) = z^{-1/2}$, or $f(z) = \log(z)$) then the spectra of the $\{A_n\}$ must remain bounded away from the singularities of f. As an example, $O(n)$ approximation of the inverse A_n^{-1} is possible if there is a compact set that contains all the spectra $\sigma(A_n)$ and has positive distance from $z = 0$, provided that the eigenvector matrices have condition numbers that remain bounded as $n \to \infty$.

4. Approximation algorithms. In this section we give a description of the algorithms we use to compute approximations to functions of sparse or banded matrices. As an integral part of the approach used, we also describe a procedure that can be used to determine a sparsity pattern or bandwidth for the approximation that captures all the relevant entries to a prescribed tolerance. In the following, when we write $f(A)$ we are implicitly assuming that the function $f(z)$ is defined on the spectrum of the matrix A and therefore the matrix function $f(A)$ is well defined.

To approximate $f(A)$ we consider algorithms that are based upon polynomial approximation. The problem is thus reduced to that of finding a rapidly converging polynomial approximation to the scalar function $f(z)$ on a domain in the complex plane. This problem has been extensively studied in the approximation theory and numerical analysis literature; see, e.g., [30, 33, 73]. Depending on the location of the spectrum, different approaches can be used. If the eigenvalues of A lie on a line segment in $\mathbb{C}$, and estimates for the endpoints of the segment are available, it is possible to scale and shift the matrix A so that the resulting matrix has its spectrum in the interval $[-1, 1]$. For instance, let A have real spectrum with $a = \lambda_{\min}(A)$ and $b = \lambda_{\max}(A)$. The affine linear function

$$\xi : \mathbb{C} \longrightarrow \mathbb{C} \quad \text{given by} \quad \xi(\lambda) = \frac{2\lambda - (a + b)}{b - a}$$

satisfies $\xi([a, b]) = [-1, 1]$, and therefore the spectrum of $B = \xi(A) = \frac{2}{b-a}A - \frac{a+b}{b-a}I$ is contained in $[-1, 1]$. Hence, we can use the well known three-term recurrence relation to generate the (optimal) Chebyshev polynomials and obtain an approximation to $f(B)$ by truncating the Chebyshev series. The approximation for the original matrix function $f(A)$ can then be easily obtained from that of $f(\xi^{-1}(B))$ by inverting ξ. For simplicity of notation, we will assume that the matrix A has already been scaled and shifted so that its spectrum is contained in the interval $[-1, 1]$.

For more general matrices, whose eigenvalues lie within a region $F \subset \mathbb{C}$ more complicated than a line segment, it is possible to generalize the method above to some extent by means of Faber polynomials; see, e.g., [10, 63, 64]. For certain domains F the Faber polynomials satisfy a three-term recurrence relation similar to that of the Chebyshev polynomials; for instance, this is the case when the spectrum lies in the interior of an ellipse on which f is analytic. However, for more general regions F the kth Faber polynomial satisfies a k-term recurrence relation. From the point of view of storage requirements, this is a serious drawback. Some truncation and restarting strategies have been considered in the literature; here we consider instead an alternative approximation technique based on Newton-type interpolation, already used in [76].

4.1. Chebyshev expansion. Expansion in the Chebyshev basis has long been a popular approach for computing approximations to the density matrix in electronic structure calculations; see, e.g., [2, 37]. The method can be used to approximate a wide variety of matrix functions, as long as the matrix can be easily transformed to one whose eigenvalues lie in $[-1, 1]$; e.g., if A is Hermitian or skew-Hermitian. We start by recalling the matrix version of the classical three-term recurrence relation for the Chebyshev polynomials:

$$T_{k+1}(A) = 2AT_k(A) - T_{k-1}(A), \qquad k = 1, 2, \ldots \tag{4.1}$$

(with $T_0(A) = I$, $T_1(A) = A$). These matrices can be used to obtain an approximation

$$f(A) = \sum_{k=1}^{\infty} c_k T_k(A) - \frac{c_1}{2} I \approx \sum_{k=1}^{N} c_k T_k(A) - \frac{c_1}{2} I =: P_N(A)$$

to $f(A)$ by truncating the Chebyshev series expansion after N terms. The coefficients c_k in the expansion only depend on f (not on A) and can be easily computed numerically at a cost independent of n using the approximation

$$c_k \approx \frac{2}{M} \sum_{j=1}^{M} f(\cos(\theta_j)) \cos((k-1)\theta_j),$$

where $\theta_j = \pi(j - \frac{1}{2})/M$. Thus, most of the computational work is performed in (4.1). The basic operation in (4.1) is the matrix–matrix multiply. If the initial matrix A is m-banded, then after k iterations the matrix $T_{k+1}(A)$ will be km-banded. In order to have a linear scaling algorithm, it is essential to fix a maximum bandwidth for the approximation $P_N(A)$, which must not depend on n. Then the cost is dominated by the matrix–matrix multiplies, and this is an $O(n)$ operation provided that the maximum bandwidth remains bounded as $n \to \infty$. Similar conclusions apply for more general sparsity patterns, which can be determined by using the structure of successive powers A^k of A. Alternatively, dropping elements by size using a drop tolerance is often used, although rigorous justification of this procedure is more difficult.

Let us now consider the error incurred by the series truncation:

$$\|e_N(A)\| = \|f(A) - P_N(A)\|, \tag{4.2}$$

where $\|\cdot\|$ is, for instance, the matrix 2-norm. We limit our discussion to the banded case, but the same arguments apply in the case of general sparsity patterns as well. Since $|T_k(x)| \leq 1$ for all $x \in [-1,1]$ and $k = 1,2,\ldots$, we have that $\|T_k(A)\| \leq 1$ for all k, since $\sigma(A) \subset [-1,1]$. Using this well-known property to bound the error in (4.2), we obtain that

$$\|e_N(A)\| = \left\| \sum_{k=N+1}^{\infty} c_k T_k(A) \right\| \leq \sum_{k=N+1}^{\infty} |c_k|.$$

The last inequality shows that the error defined by (4.2) only depends on the sum of the absolute values of the coefficients c_k for $k = N+1, N+2, \ldots$, but these in turn do not depend on n, the dimension of the matrix we are approximating. Hence if we have a sequence of $n \times n$ matrices $\{A_n\}$ with $\sigma(A_n) \subset [-1,1]$ for all n, we can use an estimate of the quantity $\sum_{k=N+1}^{\infty} |c_k|$ and use that to prescribe a sufficiently large bandwidth (sparsity pattern) to ensure a prescribed accuracy of the approximation. As long as the bandwidth of the approximation does not exceed the maximum prescribed bandwidth, the error is guaranteed to be n-independent. In practice, however, we found that this strategy is too conservative. Because of the rapid decay outside of the bandwidth of the original matrix, it is usually sufficient to prescribe a much smaller maximum bandwidth than the one predicted by the truncation error. This means that numerical dropping is necessary (see section 4.3 below for our recommended dropping strategy), since the bandwidth of $P_N(A)$ rapidly exceeds the maximum allowed bandwidth. Because of dropping, the simple error estimate given above is no longer rigorously valid. Our numerical experiments, however, suggest that n-independence (and therefore linearly scaling complexity and storage requirements) is maintained.

4.2. General polynomial approximation. We now turn to the problem of approximating $f(A)$ when the spectrum of A is contained in an arbitrary continuum $F \subset \mathbb{C}$; for a more detailed description of the technique we use, see [76]. The (Newton) interpolation polynomial we use is of the form

$$P_N(A) = c_0 I + c_1(A - z_0 I) + c_2(A - z_0 I)(A - z_1 I) + \cdots + c_N(A - z_0 I)\ldots(A - z_{N-1} I),$$

where c_k is the divided difference of order k, i.e.,

$$c_k = f[z_0, \ldots, z_k], \qquad k = 0, \ldots, N.$$

For $k = 0, \ldots, N-1$, the interpolation points are chosen as $z_k = \Psi(\omega_k)$, where ω_k are the N roots of the equation $\omega^N = \rho$ and $\Psi(z)$ is the inverse of the map $\Phi(z)$ that maps the complement of F to the outside of a disk with radius ρ and satisfies the normalization conditions (3.2). This method does not require the computation of Faber polynomials and their coefficients. However, it does require knowledge of the map $\Psi(z)$. For specific domains F this map can be determined analytically, see for example [10, 76]. In addition, $\Psi(z)$ may

ETNA
Kent State University
etna@mcs.kent.edu

require information on the convex hull of the eigenvalues of A. For more general domains one may have to resort to numerical approximations to compute $\Psi(z)$; see [77]. Once again, the approximation algorithm requires mostly matrix–matrix multiplies with banded (or sparse) matrices and appropriate sparsification is generally required to keep the cost within $O(n)$ as the problem size n grows. A rigorous error analysis that takes dropping as well as truncation into account is currently being attempted.

4.3. Dropping. We combine the above approaches with a dropping strategy. The idea applies to more general sparsity patterns, but we restrict our discussion to the case where A is a banded matrix with bandwidth m. In this case we only keep elements inside a prescribed bandwidth $\hat{m}$ in every iteration. For given ρ and R we choose $\hat{m}$ a priori so as to guarantee that

$$(\rho/R)^{\hat{m}} \approx \varepsilon/C,$$

where $C > 0$ is the constant for the bounds on $[f(A)]_{ij}$ (with $i \neq j$) discussed in section 3.7 and $\varepsilon > 0$ is a prescribed tolerance. Note that C is independent of the norm of $f(A)$ since we do not expect the diagonal elements to be insignificant.

As already noted, if A is banded with bandwidth m, then A^k has bandwidth km. This means that if we want the approximation to have a fixed bandwidth $\hat{m}$, where $\hat{m}$ is (say) an integer multiple of m corresponding to a prescribed approximation error ε, then we ought to truncate the expansion at the N^*th term, with $N^* = \hat{m}/m$. It may happen, however, that this value of N is actually too small to reduce the error below the prescribed threshold. In this case it is necessary to add extra terms to the Chebyshev expansion; but this would lead to an increase of the bandwidth beyond the prescribed limit. A solution that has been used by physicists is simply to continue the recurrence but ignoring all entries in positions outside the prescribed bandwidth; see, e.g., [2, 37, 71]. The important fact is that by restricting all the terms in the three-term recurrence (4.1) to have a fixed bandwidth (independent of n and N) we obtain an approximation scheme whose cost scales linearly in the size n of the problem.

4.4. Stopping criteria. In alternative to strategies that fix the number of terms in the polynomial expansion a priori (with or without dropping), one could decide to continue to add additional terms to the approximation to $f(A)$ until some convergence criterion is satisfied. In other words, a dynamic criterion rather than a static, a priori one could be used. For instance, with the Chebyshev expansion method one can monitor the size of the coefficients c_k and stop iterating when these are sufficiently small. Also, for some matrix functions a natural notion of *residual* exists. For example in the case of $f(A) = A^{-1}$, stopping criteria could be formulated on the basis of the size (in some appropriate norm) of the matrix $R = I - AM$ (or $R = I - MA$), where M is the current computed approximation. For $f(A) = A^{1/2}$, where A is Hermitian and positive definite, one can use $\|R\| = \|A - M^2\|$, and $\|R\| = \|I - MAM\|$ for $f(A) = A^{-1/2}$. Relative versions of these residuals can also be used, if an estimate of $\|f(A)\|$ is available.

For many other types of matrix functions, however, no natural notion of residual seems to be readily available. In this case one may decide to stop the approximation process if significant progress is no longer being made. A possibility would be to stop iterating when the norm of the difference $\|M_{new} - M_{old}\|$ between two successive approximations is below a prescribed threshold $\delta > 0$. Another possibility, which makes sense in some applications, could be to monitor the change in the trace; that is, the process is stopped when $|\text{trace}(M_{new}) - \text{trace}(M_{old})| < \delta$. Such criteria may fail in case of temporary stagnation; see [32] for some discussion of this issue, although in a somewhat different context. A rigorous justification of such criteria, however, is still lacking, and more work is needed to devise reliable dynamic stopping criteria.

ETNA
Kent State University
etna@mcs.kent.edu

5. Numerical experiments. In this section we present the results of a few preliminary numerical experiments aimed at illustrating the performance of the approximation algorithms discussed in section 4, with particular attention to linear scaling behavior. We performed several tests involving different types of matrices and functions. We regard our approximation schemes as iterative methods, with each new iteration consisting of the inclusion of an additional term in the polynomial expansion. In our computations we exploit the decay property of the entries in $f(A)$ to keep a fixed number $\hat{m}$ of off–diagonals in every iteration. By N we denote the number of terms in the expansion that were necessary to obtain the indicated error. The errors were obtained by comparing the computed approximation to the exact computation via MATLAB's `funm` command. Unless otherwise specified, we measure the relative error in the Frobenius norm.

As a first example we consider the matrix function known as *Fermi–Dirac density matrix* (FDM), which is fundamental to computational quantum chemistry:

$$P = f(A), \quad \text{where} \quad f(z) = \frac{1}{1 + \mathrm{e}^{\beta(z-\mu)}}.$$

Here A is the Hamiltonian and μ is a shift parameter, called the *chemical potential*, which can be chosen so as to have $\mathrm{trace}[f(A)] = N_e$, the number of electrons (occupied states) in the system. The parameter β is called the *inverse temperature*; see, e.g., [2]. The exponential off-diagonal decay property of the density matrix, which is closely related to the localization of the eigenvectors of the Hamiltonian, is well-established in the physics literature; see, e.g., [1, 37, 42, 52, 59]. Since the early 1990s, this property has been exploited in the development of $O(n)$ methods for electronic structure calculation [14, 34, 37, 52, 55, 61, 70, 71]. Here we note that since $P = f(A)$ is a smooth function of a sparse symmetric matrix A (the Hamiltonian is usually well-localized for a suitable choice of the basis set), the decay property also follows from Theorem 2.2 in [8]. The city plot in Figure 5.1 shows the rapid decay in the Fermi–Dirac density matrix corresponding to a "one-dimensional Anderson model"; see [1]. That is, the matrix function is $f(A) = (\exp(\beta(A - \mu I)) + I)^{-1}$ where A is a 200×200 symmetric tridiagonal matrix with random entries on the main diagonal (drawn from the uniform distribution on $[0, 1]$) and with the remaining nonzero entries equal to -1. In this example we used $\beta = 2$ and $\mu = 0.5$.

In Table 5.1 we show results for the approximation of the Fermi–Dirac density matrix for different values of n, μ, and β. Under "cost/n" we report the cost per unknown, i.e., the total number of arithmetic operations divided by n. The results clearly show the n-independent behavior (and hence, linear scaling complexity) of the algorithm. The total cost of the approximation as a function of problem size for the two test cases is shown in Fig. 5.2.

TABLE 5.1
Results for $f(z) = \frac{1}{1+\exp(\beta(z-\mu))}$

	$\mu = 2, \beta = 2.13$				$\mu = 0.5, \beta = 1.84$			
n	error	N	$\hat{m}$	cost / n	error	N	$\hat{m}$	cost / n
100	$9e{-}06$	18	20	6129	$6e{-}06$	18	22	6129
200	$4e{-}06$	19	20	6498	$9e{-}06$	18	22	6129
300	$4e{-}06$	19	20	6498	$5e{-}06$	20	22	6867
400	$6e{-}06$	19	20	6498	$8e{-}06$	20	22	6867
500	$8e{-}06$	19	20	6498	$8e{-}06$	20	22	6867

We now discuss how the values of N and $\hat{m}$ were determined. Note that since the Fermi–Dirac density matrix is an approximation of an orthogonal projector of known rank, special

ETNA
Kent State University
etna@mcs.kent.edu

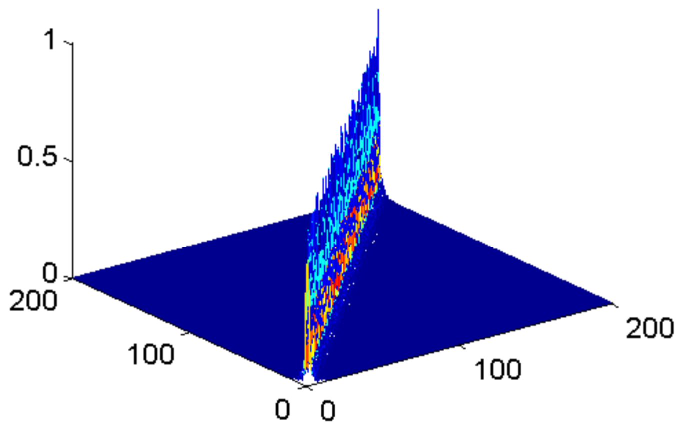

FIG. 5.1. *City plot for a Fermi–Dirac density matrix.*

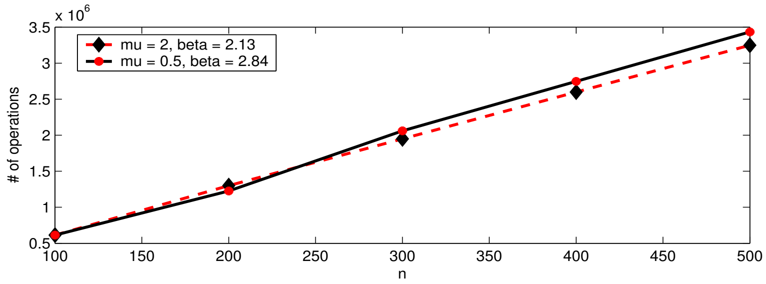

FIG. 5.2. *Flop count as a function of n for Fermi–Dirac matrix functions.*

stopping criteria can be used in this case. However, rather than using such criteria we adopted a more general strategy. The choice of the number N of terms to include in the Chebyshev series expansion is an easy matter, since the magnitude of the coefficients can be monitored. Note that if f is analytic, the coefficients c_k decay exponentially fast. Once the sum of the absolute values of three consecutive coefficients is sufficiently small, no additional terms are added; here we used the rather conservative tolerance of $1e{-}15$. Although not very sophisticated, this criterion gave good results in practice. We obtained the a priori bandwidth $\hat{m} = 20$ as follows. We first note that $|f(z)| \leq 1$ for all z, therefore the constant $M(R)$ in the bounds for $|[f(A)]_{ij}|$ can be neglected. To perform the computation in the Chebyshev basis we scale and shift the matrix A so that its spectrum lies in $[-1, 1]$. To obtain the approximation to $f(A)$ with the original A we shift and scale back the coefficients in the Chebyshev series. The logarithmic capacity of $[-1, 1]$ is given by $\rho = 1/2$, and $V \frac{1}{\pi(1-\rho/R)} = 4$ when $R = 1$. The entries (i, j) in $f(A)$, where i and j are such that $|i - j| \geq 20$, have magnitude at most $3e{-}06$ and can therefore be neglected. For the second case ($\mu = 0.5$, $\beta = 1.84$) we had to slightly increase N and $\hat{m}$ to achieve a comparable *a priori* error. As a rule, it is a good idea to use values of $\hat{m}$ that are slightly greater than those predicted by the theory, also in view of the effects of dropping. We further note that in this example, the maximum bandwidth $\hat{m}$ was actually never reached, so no dropping was performed. This shows, incidentally, that for

ETNA
Kent State University
etna@mcs.kent.edu

this example the *a priori* estimate for $\hat{m}$ was quite accurate, since the actual final bandwidth is very close to $\hat{m}$.

We performed many more experiments with the Chebyshev method and obtained excellent results whenever f is an entire function. If f has singularities, however, we found that the method can converge quite slowly, especially when the singularities are close to the spectrum of A. In these cases, the approach described in section 4.2 was found to be much better. Therefore, in the remaining tables we use the interpolation approach. The results in Table 5.2 were obtained by interpolating the "entropy" function $f(z) = z\log(z)$ to obtain an approximation to $f(A) = A\log(A)$. As before, the values of N and $\hat{m}$ that are guaranteed to result in a given error can be estimated a priori from knowledge of the parameters ρ and R. Note that in this case, the estimated bandwidth is about twice the final one, suggesting that the decay bounds used to estimate $\hat{m}$ can sometimes be conservative. The matrix A is a tridiagonal matrix with random diagonal entries, shifted so as to be diagonally dominant (and therefore positive definite, so that the logarithm is well-defined). We also report the absolute error in the trace of the matrix function. Once again, the linear scaling behavior is clear.

TABLE 5.2
Results for $f(z) = z\log(z)$

	$A\log(A)$	trace$[A\log(A)]$		
n	error	error	$\hat{m}$	N
100	$5e{-}07$	$3e{-}04$	20	9
200	$6e{-}07$	$8e{-}04$	20	9
300	$1e{-}07$	$3e{-}04$	20	10
500	$2e{-}07$	$5e{-}04$	20	10

For Table 5.3 we use a pseudosparse matrix A with entries given by $a_{ij} = \mathrm{e}^{-\alpha|i-j|}$, where $\alpha > 0$. This is a convenient way to generate matrices with a given rate of decay. Note that in the limiting case $\alpha = 0$, the matrix is "flat" (there is no decay). As α is increased, the rate of decay also increases. It is known that A^{-1} is a tridiagonal M-matrix, hence A is positive definite. To compute $f(A) = \log(A)$, we first sparsify A by keeping enough diagonals to make the sparsification error $\|A - \tilde{A}\|$ smaller than a prescribed accuracy and then we approximate $\log(\tilde{A})$. The reported errors are relative to the matrix function $\log(A)$ computed on the full A. For the first test we use $\alpha = 2$ and we keep $m = 15$ off-diagonals above and below the main diagonal. Again, the linear scaling behavior is apparent.

TABLE 5.3
Results for $f(z) = \log(z)$

	$\log(A)$	trace$[\log(A)]$		
n	error	error	$\hat{m}$	N
100	$4e{-}07$	$7e{-}06$	10	9
200	$4e{-}07$	$1e{-}05$	10	9
300	$4e{-}07$	$2e{-}05$	10	9
500	$4e{-}07$	$3e{-}05$	10	9

Next we consider some non-Hermitian examples. In Table 5.4 we show results for the exponential, the sine and the cosine. The matrix A has entries $a_{ij} = \mathrm{e}^{-\alpha(i-j)}$ for $i \geq j$ and $a_{ij} = \mathrm{e}^{-\beta(j-i)}$ for $i < j$ with $\alpha, \beta > 0$. To obtain a non-symmetric matrix we choose $\alpha = 1$ and $\beta = 1.5$. We sparsify the matrix by keeping 25 off–diagonals on either side of the main diagonal. The reported errors are relative to the "exact" computation using the full A.

TABLE 5.4
Results for $f(z) = \exp(z)$, $\cos(z)$, $\sin(z)$

	$\exp(A)$			$\cos(A)$			$\sin(A)$		
n	error	$\hat{m}$	N	error	$\hat{m}$	N	error	$\hat{m}$	N
100	$6e{-}08$	30	12	$4e{-}07$	30	11	$9e{-}07$	30	11
300	$6e{-}08$	30	12	$4e{-}07$	30	11	$3e{-}08$	30	12
500	$4e{-}08$	30	11	$2e{-}08$	30	11	$2e{-}07$	30	10

In Table 5.5 we show results for the exponential function for a non-symmetric matrix with A generated as above, with $\alpha = 1$ and $\beta = 2$. Now we keep only 15 off–diagonals on either side of the main diagonal. Again, we observe n-independence of the results.

TABLE 5.5
Results for $f(z) = \exp(z)$

	$\exp(A)$		
n	error	$\hat{m}$	N
100	$3e{-}07$	25	10
300	$4e{-}07$	25	9
500	$4e{-}07$	25	9
1000	$4e{-}07$	25	9

For the results in Table 5.6 we use two examples of banded non-symmetric matrices from the Matrix Market. Again, we show results for the exponential, the sine and the cosine. The matrices where scaled by dividing through by the largest eigenvalue.

TABLE 5.6
Results for $f(z) = \exp(z)$, $\cos(z)$, $\sin(z)$

$f(A)$	A	n	error	$\hat{m}$	N
$\exp(A)$	*Bai/olm100*	100	$3e{-}06$	12	8
$\exp(A)$	*Bai/olm500*	500	$2e{-}06$	12	9
$\sin(A)$	*Bai/olm100*	100	$3e{-}06$	12	8
$\sin(A)$	*Bai/olm500*	500	$1e{-}06$	12	9
$\cos(A)$	*Bai/olm100*	100	$4e{-}06$	12	8
$\cos(A)$	*Bai/olm500*	500	$2e{-}06$	12	9

Additional possible applications of the theory and methods discussed in the present paper include fast algorithms for approximating the determinant of a large, sparse symmetric positive definite matrix; see [6, 47, 54, 67] for motivation and background. One possible approach is the reduction of the computation of the determinant to that of the trace of a matrix function; for instance, one could exploit the straightforward identity $\det(A) = \exp(\text{trace}[\log(A)])$. If the smallest eigenvalue of A is not too close to zero, approximating the trace of $\log(A)$ is not difficult and good results can be expected. In Table 5.7 we give results for the approximation of the determinant of two symmetric positive definite matrices obtained using the identity $\det(A) = \exp(\text{trace}[\log(A)])$. For the first example we use the matrix A with entries $a_{ij} = e^{-2|i-j|}$ as above. Prior to computing the approximation of the logarithm we truncate A symmetrically to fifteen off-diagonals. For the second example we use a banded diagonally dominant matrix with sixteen nonzero diagonals. The (relative) errors reported in the table compare the computed approximation with the actual value of the determinant of A.

ETNA
Kent State University
etna@mcs.kent.edu

TABLE 5.7
Approximation of the determinant

	Toeplitz matrix			Banded matrix		
n	error	N	$\hat{m}$	error	N	$\hat{m}$
100	$7e{-}06$	10	9	$2e{-}06$	14	45
200	$1e{-}05$	10	9	$5e{-}06$	14	45
300	$2e{-}05$	10	9	$7e{-}06$	14	45
500	$4e{-}05$	10	9	$1e{-}05$	14	45

Finally, we consider banded approximations to $f(A) = A^{-1}$ as preconditioners for the conjugate gradient method applied to linear systems $Ax = b$ where $A = T + D$ is a symmetric positive definite "Toeplitz-plus-diagonal" matrix. We note that no fast solvers are available for this problem unless D is nearly a multiple of the identity, but see the recent paper [66] for an approach that yields reasonable results. Here T has entries $t_{ij} = \mathrm{e}^{-0.1|i-j|}$ and D is a random diagonal matrix with entries in the interval $(5, 6)$. The right-hand side b is a random vector and $x_0 = 0$ is used as the initial guess. We stop the preconditioned conjugate gradient (PCG) iteration when the residual norm is reduced to at least $1e{-}07$. The preconditioners are constructed as follows. We first truncate A to obtain a matrix with bandwidth $m = 20$, then we run a fixed number $N = 10$ of iterations of the polynomial approximation method for $f(z) = z^{-1}$ to construct a banded approximate inverse to the truncated matrix. We fix the bandwidth of the approximate inverse to $\hat{m} = 40$, i.e., we allow the approximate inverse to have twice the bandwidth of the truncated matrix. Of course, the original matrix A is used for the matrix–vector multiplies in the PCG iteration. The results in Table 5.8 show that this simple preconditioner works quite well. Note that with an FFT-based implementation the cost of each matrix–vector multiply with A in the conjugate gradient method is $O(n \log n)$, whereas the cost of forming and applying the preconditioner in each iteration is just $O(n)$. In contrast, application of a circulant-plus-diagonal preconditioner costs $O(n \log n)$ in general.

TABLE 5.8
Number of PCG iterations

	n						
Precond.	100	200	600	800	1000	2000	3000
None	37	56	73	82	85	86	95
$M \approx A^{-1}$	8	12	13	13	14	14	16

6. Conclusions. In this paper we have proved some general theoretical results on decay phenomena in functions of sparse matrices, and we have investigated some algorithms, based on polynomial approximation, for approximating matrix functions. Our decay bounds extend and unify several results that were known from the literature, and help understand various localization phenomena observed by physicists. Moreover, our results show that $O(n)$ approximation is possible in many cases of practical interest. Of course, much work remains to be done. In this paper we have illustrated the theory by some simple MATLAB experiments involving mostly band matrices of fairly small size. Although this was enough for our purpose, which was to illustrate the decay behavior of $f(A)$ and the $O(n)$ scaling of the approximation algorithms, we plan to investigate more realistic problems in the near future. This includes larger matrices with general sparsity structure, and more complicated matrix functions.

Acknowledgment. The authors are indebted to Sandro Graffi, Nick Higham and two anonymous referees for useful suggestions.

REFERENCES

[1] P. W. ANDERSON, *Absence of diffusion in certain random lattices*, Phys. Review, 109 (1958), pp. 1492–1505.

[2] R. BAER AND M. HEAD-GORDON, *Chebyshev expansion methods for electronic structure calculations on large molecular systems*, J. Chem. Phys., 107 (1997), pp. 10003–10013.

[3] ———, *Sparsity of the density matrix in Kohn–Sham density functional theory and an assessment of linear system-size scaling methods*, Phys. Rev. Lett., 79 (1997), pp. 3962–3965.

[4] Z. BAI, M. FAHEY AND G. H. GOLUB, *Some large-scale matrix computation problems*, J. Comput. Appl. Math., 74 (1996), pp. 71–89.

[5] Z. BAI, M. FAHEY, G. H. GOLUB, M. MENON AND E. RICHTER, *Computing partial eigenvalue sums in electronic structure calculations*, Tech. Report SCCM-98-03, Stanford University, 1998.

[6] Z. BAI AND G. H. GOLUB, *Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices*, Ann. Numer. Math., 4 (1997), pp. 29–38.

[7] C. BEKAS, E. KOKIOPOULOU, AND Y. SAAD, *An estimator for the diagonal of a matrix*, Appl. Numer. Math., to appear.

[8] M. BENZI AND G. H. GOLUB, *Bounds for the entries of matrix functions with applications to preconditioning*, BIT, 39 (1999), pp. 417–438.

[9] M. BENZI AND M. TŮMA, *Orderings for factorized sparse approximate inverse preconditioners*, SIAM J. Sci. Comput., 21 (2000), pp. 1851–1868.

[10] L. BERGAMASCHI, M. CALIARI AND M. VIANELLO, *Efficient approximation of the exponential operator for discrete 2D advection-diffusion problems*, Numer. Linear Algebra Appl., 10 (2003), pp. 271–289.

[11] L. BERGAMASCHI AND M. VIANELLO, *Efficient computation of the exponential operator for large, sparse, symmetric matrices*, Numer. Linear Algebra Appl., 7 (2000), pp. 27–45.

[12] D. A. BINI, N. J. HIGHAM AND B. MEINI, *Algorithms for the matrix pth root*, Numer. Algorithms, 39 (2005), pp. 349–378.

[13] A. BORIÇI, *A Lanczos approach to the inverse square root of a large and sparse matrix*, J. Comput. Phys., 162 (2000), pp. 123–131.

[14] D. R. BOWLER AND M. J. GILLAN, *Density matrices in $O(N)$ electronic structure calculations: theory and applications*, Comp. Phys. Comm., 120 (1999), pp. 95–108.

[15] D. CALVETTI, S.-M. KIM AND L. REICHEL, *Quadrature rules based on the Arnoldi process*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 765–781.

[16] M. CRAMER AND J. EISERT, *Correlations, spectral gap and entanglement in harmonic quantum systems on generic lattices*, New J. Phys., 8 (2006), Art. No. 71.

[17] M. CRAMER, J. EISERT, M. B. PLENIO AND J. DREISSIG, *Entanglement-area law for general bosonic harmonic lattice systems*, Phys. Rev. A, 73 (2006), Art. No. 012309.

[18] J. H. CURTISS, *Faber polynomials and the Faber series*, Amer. Math. Monthly, 78 (1971), pp. 577–596.

[19] P. I. DAVIES AND N. J. HIGHAM, *Computing $f(A)b$ for matrix functions f*, in A. Boriçi, A. Frommer, B. Joo, A. Kennedy and B. Pendleton, eds., *QCD and Numerical Analysis III*, Lect. Notes Comput. Sci. Eng., Vol. 47, Springer-Verlag, Berlin, 2005, pp. 15–24.

[20] N. DEL BUONO AND L. LOPEZ, *A survey on methods for computing matrix exponentials in numerical schemes for ODEs*, Lecture Notes in Comput. Sci., 2658 (2003), pp. 111–120.

[21] N. DEL BUONO, L. LOPEZ AND R. PELUSO, *Computation of the exponential of large sparse skew-symmetric matrices*, SIAM J. Sci. Comput., 27 (2005), pp. 278–293.

[22] S. DEMKO, W. F. MOSS AND P. W. SMITH, *Decay rates for inverses of band matrices*, Math. Comp., 43 (1984), pp. 491–499.

[23] L. DIECI, B. MORINI AND A. PAPINI, *Computational techniques for real logarithms of matrices*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 570–593.

[24] R. DIESTEL, *Graph Theory*, Springer-Verlag, Berlin, 2000.

[25] V. L. DRUSKIN AND L. A. KNIZHNERMAN, *Two polynomial methods for computing functions of symmetric matrices*, Comput. Math. Math. Phys., 29 (1989), pp. 112–121.

[26] ———, *Extended Krylov subspaces: approximation of the matrix square root and related functions*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 755–771.

[27] I. S. DUFF, A. M. ERISMAN, C. W. GEAR, AND J. K. REID, *Sparsity structure and Gaussian elimination*, SIGNUM Newsletter, Association for Computing Machinery, New York, 23 (1988), pp. 2–8.

[28] M. EIERMANN AND O. ERNST, *A restarted Krylov subspace method for the evaluation of matrix functions*, SIAM J. Numer. Anal., 44 (2006), pp. 2481–2504.

[29] S. W. ELLACOTT, *Computation of Faber series with application to numerical polynomial approximation in the complex plane*, Math. Comp., 40 (1983), pp. 575–587.

[30] L. FOX AND I. B. PARKER, *Chebyshev Polynomials in Numerical Analysis*, Oxford University Press, London, 1968.
[31] A. FROMMER AND V. SIMONCINI, *Matrix functions*, Tech. Report, Dipartimento di Matematica, Università di Bologna, 2006.
[32] ———, *Stopping criteria for rational matrix functions of Hermitian and symmetric matrices*, Tech. Report, Dipartimento di Matematica, Università di Bologna, 2007.
[33] D. GAIER, *Lectures on Complex Approximation*, Birkhäuser, Boston, 1987.
[34] G. GALLI, *Large-scale electronic structure calculations using linear scaling methods*, Phys. Stat. Sol. B, 217 (2000), pp. 231–249.
[35] E. GALLOPULOS AND Y. SAAD, *Efficient solution of parabolic equations by Krylov subspace methods*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 1236–1264.
[36] S. GOEDECKER, *Decay properties of the finite-temperature density matrix in metals*, Phys. Rev. B, 58 (1998), pp. 3501–3502.
[37] ———, *Linear scaling electronic structure methods*, Rev. Mod. Phys., 71 (1999), pp. 1085–1123.
[38] G. H. GOLUB AND G. MEURANT, *Matrices, moments, and quadratures*, in Numerical Analysis 1993, D. F. Griffiths and G. A. Watson, eds., Pitman Research Notes in Mathematics, vol. 303, Essex, England, UK, 1994, pp. 105–156.
[39] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd edition, Johns Hopkins University Press, Baltimore, MD, 1996.
[40] K. GRÖCHENIG AND M. LEINERT, *Symmetry and inverse-closedness of matrix algebras and functional calculus for infinite matrices*, Trans. Amer. Math. Soc., 358 (2006), pp. 2695–2711.
[41] G. I. HARGREAVES AND N. J. HIGHAM, *Efficient algorithms for the matrix cosine and sine*, Numer. Algorithms, 40 (2005), pp. 383–400.
[42] L. HE AND D. VANDERBILT, *Exponential decay properties of Wannier functions and related quantities*, Phys. Rev. Lett., 86 (2001), pp. 5341–5344.
[43] N. J. HIGHAM, *Functions of matrices*, in L. Hogben, ed., *Handbook of Linear Algebra*, Chapman and Hall/CRC, Boca Raton, FL, 2006.
[44] M. HOCHBRUCK AND C. LUBICH, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.
[45] ———, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574.
[46] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1994.
[47] I. C. F. IPSEN AND D. J. LEE, *Determinant approximations*, Tech. Report, Department of Mathematics, North Carolina State University, Raleigh, NC, 2006.
[48] A. ISERLES, *How large is the exponential of a banded matrix?*, New Zealand J. Math., 29 (2000), pp. 177–192.
[49] S. JAFFARD, *Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 7 (1990), pp. 461–476.
[50] C. S. KENNEY AND A. J. LAUB, *Condition estimates for matrix functions*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 191–209.
[51] L. A. KNIZHNERMAN, *Calculation of functions of unsymmetric matrices using Arnoldi's method*, Comput. Math. Math. Phys., 31 (1991), pp. 5–9.
[52] W. KOHN, *Density functional and density matrix method scaling linearly with the number of atoms*, Phys. Rev. Lett., 76 (1996), pp. 3168–3171.
[53] P. LAASONEN, *On the iterative solution of the matrix equation* $AX^2 - I = 0$, Math. Tables Other Aids Comp., 12 (1958), pp. 109–116.
[54] D. J. LEE AND I. C. F. IPSEN, *Zone determinant expansions for nuclear lattice simulations*, Phys. Rev. C, 68 (2003), pp. 064003-1/8.
[55] X.-P. LI, R. W. NUNES AND D. VANDERBILT, *Density-matrix electronic-structure method with linear system-size scaling*, Phys. Rev. B, 47 (1993), pp. 10891–10894.
[56] L. LOPEZ AND A. PUGLIESE, *Decay behaviour of functions of skew-symmetric matrices*, in Proceedings of HERCMA 2005, 7th Hellenic-European Conference on Computer Mathematics and Applications, September 22-24, 2005, Athens, E. A. Lipitakis, ed., Electronic Editions LEA, Athens.
[57] L. LOPEZ AND V. SIMONCINI, *Analysis of projection methods for rational function approximations to the matrix exponential*, SIAM J. Numer. Anal., 44 (2006), pp. 613–635.
[58] A. I. MARKUSHEVICH, *Theory of Functions of a Complex Variable, Vol. III*, Prentice-Hall, Englewood Cliffs, NJ, 1967.
[59] P. E. MASLEN, C. OCHSENFELD, C. A. WHITE, M. S. LEE AND M. HEAD-GORDON, *Locality and sparsity of ab initio one-particle density matrices and localized orbitals*, J. Phys. Chem. A, 102 (1998), pp. 2215–2222.
[60] G. MEURANT, *A review on the inverse of symmetric tridiagonal and block tridiagonal matrices*, SIAM J. Ma-

trix Anal. Appl., 13 (1992), pp. 707–728.

[61] J. M. MILLAM AND G. E. SCUSERIA, *Linear scaling conjugate gradient density matrix search as an alternative to diagonalization for first principles electronic structure calculations*, J. Chem. Phys., 106 (1997), pp. 5569–5577.

[62] C. MOLER AND C. VAN LOAN, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Rev., 45 (2003), pp. 3–49.

[63] I. MORET AND P. NOVATI, *The computation of functions of matrices by truncated Faber series*, Numer. Funct. Anal. Optim., 22 (2001), pp. 697–719.

[64] ———, *Interpolating functions of matrices on zeros of kernel polynomials*, Numer. Linear Algebra Appl., 11 (2005), pp. 337–353.

[65] R. NABBEN, *Decay rates of the inverse of nonsymmetric tridiagonal and band matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 820–837.

[66] M. K. NG AND J. PAN, *Approximate inverse circulant-plus-diagonal preconditioners for Toeplitz-plus-diagonal matrices*, Tech. Report, Department of Mathematics, Hong Kong Baptist University, 2006.

[67] A. REUSKEN, *Approximation of the determinant of large sparse symmetric positive definite matrices*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 799–818.

[68] R. F. RINEHART, *The equivalence of definitions of a matric function*, Amer. Math. Monthly, 62 (1955), pp. 395–414.

[69] Y. SAAD, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228.

[70] V. E. SACKSTEDER, *Linear Algebra with Disordered Sparse Matrices that have Spatial Structure: Theory and Computation*, PhD thesis, Department of Physics, Università degli Studi di Roma "La Sapienza," Rome, Italy, 2004.

[71] ———, *$O(N)$ algorithms for disordered systems*, Numer. Linear Algebra Appl., 12 (2005), pp. 827–838.

[72] N. SCHUCH, J. I. CIRAC AND M. M. WOLF, *Quantum states on harmonic lattices*, Comm. Math. Phys., 267 (2006), pp. 65–92.

[73] V. I. SMIRNOV AND N. A. LEBEDEV, *Functions of a Complex Variable: Constructive Theory*, MIT Press, Cambridge, MA, 1968.

[74] D. T. SMITH, *Exponential decay of resolvents and discrete eigenfunctions of banded infinite matrices*, J. Approx. Theory, 66 (1991), pp. 83–97.

[75] P. K. SUETIN, *Fundamental properties of Faber polynomials*, Russian Math. Surveys, 19 (1964), p. 121–149.

[76] H. TAL-EZER, *Polynomial approximation of functions of matrices and applications*, J. Sci. Comput., 4 (1989), pp. 25–60.

[77] L. N. TREFETHEN, *Numerical computation of the Schwarz–Christoffel transformation*, SIAM J. Sci. Stat. Comput., 1 (1980), pp. 82–102.

[78] L. N. TREFETHEN, M. CONTEDINI AND M. EMBREE, *Spectra, pseudospectra, and localization for random bidiagonal matrices*, Comm. Pure Appl. Math., 54 (2001), pp. 595–623.

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