Electronic Transactions on Numerical Analysis.
Volume 55, pp. 438–454, 2022.
Copyright © 2022, Kent State University.
ISSN 1068–9613.
DOI: 10.1553/etna_vol55s438

ETNA
Kent State University and
Johann Radon Institute (RICAM)

DECAY BOUNDS FOR BERNSTEIN FUNCTIONS OF HERMITIAN MATRICES WITH APPLICATIONS TO THE FRACTIONAL GRAPH LAPLACIAN*

MARCEL SCHWEITZER†

Abstract. For many functions of matrices $f(A)$, it is known that their entries exhibit a rapid—often exponential or even superexponential—decay away from the sparsity pattern of the matrix A. In this paper, we specifically focus on the class of Bernstein functions, which contains the fractional powers A^α, $\alpha \in (0,1)$, as an important special case, and derive new decay bounds by exploiting known results for the matrix exponential in conjunction with the Lévy-Khintchine integral representation. As a particular special case, we find a result concerning the power law decay of the strength of connection in nonlocal network dynamics described by the fractional graph Laplacian, which improves upon known results from the literature by doubling the exponent in the power law.

Key words. matrix functions, Bernstein functions, off-diagonal decay, graph Laplacian, fractional powers, nonlocal dynamics

AMS subject classifications. 05C82, 15A16, 65F50, 65F60

1. Introduction. In this paper, we investigate the decay of the entries of matrix functions $f(A)$, where $A \in \mathbb{C}^{n\times n}$ is a Hermitian matrix and f is a *Bernstein function*, i.e., a nonnegative function $f\colon (0,\infty) \to \mathbb{R}$ that is infinitely many times continuously differentiable and satisfies

$$(-1)^{n-1} f^{(n)}(z) \geq 0 \quad \text{for all } n \in \mathbb{N} \text{ and } z \in (0,\infty). \tag{1.1}$$

As the condition (1.1) means that f' is a completely monotonic function, the class of Bernstein functions is intimately related to the completely monotonic function classes of Laplace-Stieltjes and Cauchy-Stieltjes transforms. While the latter classes have received considerable interest in the analysis of matrix functions in recent years (see, e.g., [3, 19, 25, 26, 32] and the references therein), the class of Bernstein functions has not been investigated as thoroughly, although it also frequently occurs in applications. Our present study is motivated by the fact that fractional powers L_G^α, $\alpha \in (0,1)$, where L_G is the Laplacian of an undirected graph G, have recently emerged as a useful tool in modeling non-local diffusion processes on graphs and in the efficient exploration of large networks; see, e.g., [5, 12, 36, 37]. Clearly, z^α, $\alpha \in (0,1)$, is nonnegative on $(0,\infty)$ and fulfills the condition (1.1), so that it is a Bernstein function, whereas it is neither a Laplace-Stieltjes nor a Cauchy-Stieltjes function.

The off-diagonal decay in matrix functions is a topic that has been intensively studied in the past, in particular for the case of the matrix inverse $f(A) = A^{-1}$ (see, e.g., [14, 16, 18, 21, 22]) and for entire functions like the exponential [6, 8, 29, 31, 35]. Quite recently, some of these bounds were also extended to Cauchy-Stieltjes and Laplace-Stieltjes functions of matrices; see, e.g., [8, 21] and also the recent survey [4]. A priori knowledge of the decay in matrix functions has many different applications, e.g., the efficient construction of sparse approximations [24, 39], the design of linearly scaling algorithms for certain linear algebra problems [7, 13], and the analysis of probing methods for trace estimation [23]. In the context of fractional powers of the graph Laplacian L_G, decay estimates can give insight into transition probabilities of non-local random walks on G; see [5, 12].

The remainder of this paper is organized as follows. In Section 2, we recall some basic facts about Bernstein functions and several special functions that appear in the bounds that we derive in later sections and facts on functions of matrices in general. Our main results for the

*Received November 11, 2021. Accepted March 3, 2022. Published online on May 9, 2022. Recommended by M. Benzi.
†School of Mathematics and Natural Sciences, Bergische Universität Wuppertal, 42097 Wuppertal, Germany (marcel@uni-wuppertal.de).

decay of Bernstein functions of matrices are presented in Section 3, where we consider both the case of positive semidefinite as well as the case of positive definite matrices A. Section 4 deals with the special case of the fractional graph Laplacian, where we compare our new decay estimates to previously found estimates from the literature. Concluding remarks are given in Section 5.

2. Basics. In this section, we introduce the basic concepts and notations needed for the derivations in later sections of the paper.

2.1. Bernstein functions. A *Bernstein function* is a nonnegative function $f\colon (0,\infty)\to\mathbb{R}$ that is infinitely many times continuously differentiable and satisfies (1.1). As already mentioned in Section 1, this implies that f' is a completely monotonic function. Bernstein functions can thus be characterized as nonnegative primitives of completely monotonic functions. An important result for Bernstein functions is that they exhibit the *Lévy-Khintchine integral representation*

$$f(z) = a + bz + \int_0^\infty (1 - e^{-tz})\,\mathrm{d}\mu(t), \tag{2.1}$$

where $a, b \geq 0$ and μ is a positive measure (the *Lévy measure*) on $(0,\infty)$ such that

$$\int_0^\infty \min\{t, 1\}\,\mathrm{d}\mu(t) < \infty,$$

see, e.g., [10, 38]. Also note that any Bernstein function admits a continuous extension to the origin (which we also denote by f for convenience) for which the Lévy-Khintchine representation remains valid; see, e.g., [38, Proof of Proposition 3.6].

Important examples of Bernstein functions are, e.g.,

- $f(z) = z^\alpha$, $\alpha \in (0,1)$,
- $f(z) = 1 - e^{-tz}$, $t \geq 0$, and
- $f(z) = \log(1+z)$.

As the composition of two Bernstein functions is again a Bernstein function, we also have that, e.g., $1 - e^{-tz^\alpha}$, $t \geq 0$, $\alpha \in (0,1)$, is a Bernstein function.

Because they are of particular importance in the applications we consider in later sections, we mention that the Lévy-Khintchine representation of the fractional powers is explicitly known and given by

$$z^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - e^{-tz}) t^{-\alpha-1}\,\mathrm{d}t, \qquad \alpha \in (0,1), \tag{2.2}$$

where Γ denotes the gamma function; cf. Section 2.2.

2.2. The Gamma function and related special functions. In the following, we introduce some classical special functions that appear in the derivation of our results.

The *gamma function* is defined for $z \in \mathbb{C}$ with $\Re(z) > 0$, where $\Re(z)$ denotes the real part of z, via

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t,$$

and it has the property that $\Gamma(n) = (n-1)\cdot\Gamma(n-1) = (n-1)!$ for $n \in \mathbb{N}$. Closely related are the *upper* and *lower incomplete gamma functions*, defined by

$$\Gamma(z,s) = \int_s^\infty t^{z-1} e^{-t}\,\mathrm{d}t \quad \text{and} \quad \gamma(z,s) = \int_0^s t^{z-1} e^{-t}\,\mathrm{d}t, \tag{2.3}$$

respectively. Clearly, we have $\Gamma(z) = \Gamma(z,s) + \gamma(z,s)$ for all $s \geq 0$ and $\Gamma(z) = \Gamma(z,0) = \lim_{s\to\infty} \gamma(z,s)$. We also need the *error function*

$$\operatorname{erf}(z) = \frac{1}{\sqrt{\pi}} \int_{-z}^{z} e^{-t^2}\,\mathrm{d}t$$

and the *complementary error function* $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$, which are related to the incomplete gamma functions through the identities

$$\Gamma\left(z, \frac{1}{2}\right) = \sqrt{\pi}\operatorname{erfc}(\sqrt{z}) \quad \text{and} \quad \gamma\left(z, \frac{1}{2}\right) = \sqrt{\pi}\operatorname{erf}(\sqrt{z}).$$

2.3. Matrix functions and graphs of matrices. Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix with eigendecomposition $A = V\Lambda V^H$, where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues and V contains the corresponding orthonormal eigenvectors. Then, for a scalar function $f\colon \mathbb{C} \to \mathbb{C}$, the matrix function $f(A)$ is given by the simple relation $f(A) = Vf(\Lambda)V^H$, where $f(\Lambda) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$, provided that $f(\lambda_i)$ exists for all $i = 1, \ldots, n$. For general—not necessarily diagonalizable—A, a similar definition using the Jordan canonical form is possible (where for eigenvalues with nontrivial Jordan blocks, also the derivatives of f need to be defined, up to the block size minus 1). As we only consider Hermitian matrices in this work, we do not further pursue this topic and refer the reader to [27, Chapter 1.2] for details.

It directly follows from the definition of matrix functions given above that when f is a Bernstein function (2.1), we can insert A in place of z into the integral representation and find

$$f(A) = aI + bA + \int_0^\infty (I - e^{-tA})\,\mathrm{d}\mu(t). \tag{2.4}$$

An important concept frequently used in the formulation of decay bounds for matrix functions is that of the *graph of a sparse matrix*: Given $A \in \mathbb{C}^{n\times n}$, the graph of A is given by $G(A) = (V, E)$, where $V = \{1, \ldots, n\}$ and $E = \{(i,j) : a_{ij} \neq 0, i \neq j\}$. We recall that a *path of length* ℓ in G is a sequence $(e_1, \ldots, e_\ell)$ of edges $e_k \in E, k = 1, \ldots, \ell$, such that there exist pairwise distinct nodes $i_1, \ldots, i_{\ell+1} \in V$ with $e_k = (i_k, i_{k+1}), k = 1, \ldots, \ell$. For any two nodes i, j in the graph of A, we denote by $d(i,j)$ the *geodesic distance* of the nodes in $G(A)$, i.e., the length of the shortest path from i to j. If there is no path from i to j in G, then we set $d(i,j) = \infty$. Clearly, when A is Hermitian, $G(A)$ is undirected and thus $d(i,j) = d(j,i)$ for all i, j.

3. Decay of Bernstein functions of Hermitian matrices. This section contains our main results on the off-diagonal decay of Bernstein functions of Hermitian matrices. We first consider the case of a positive semidefinite A in Section 3.1 and then discuss how the estimates can be improved when A is positive definite in Section 3.2. As these estimates for positive definite A have the drawback that some integrals occur for which no closed form solution is available, we also derive other, more explicit, bounds for fractional powers of positive definite matrices in Section 3.3 by exploiting a connection to Cauchy-Stieltjes functions.

3.1. The positive semidefinite case. Using the representation (2.1) allows us to relate the decay of matrix Bernstein functions to the decay in the matrix exponential, a very thoroughly studied topic. In particular, our analysis in this section is based on the following theorem from [8] on the decay in the matrix exponential, which in turn is based on the well-known

convergence result of Hochbruck and Lubich [28, Theorem 2] for Lanczos approximations of the action of the matrix exponential.

THEOREM 3.1 (Theorem 4.2 in [8]). *Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian positive semidefinite matrix with eigenvalues in the interval $[0, 4\rho]$, and denote by $d(i,j)$ the geodesic distance of the nodes i and j in the graph of A. Then, for $i \neq j$,*

(i) for $\rho t \geq 1$ and $\sqrt{4\rho t} \leq d(i,j) \leq 2\rho t$,

$$|[\exp(-tA)]_{ij}| \leq 10 \exp\left(-\frac{1}{5\rho t} d(i,j)^2\right),$$

(ii) for $d(i,j) \geq 2\rho t$

$$|[\exp(-tA)]_{ij}| \leq 10 \frac{\exp(-\rho t)}{\rho t} \left(\frac{e\rho t}{d(i,j)}\right)^{d(i,j)}.$$

Note that in its original form, the result of Theorem 3.1 was stated for banded matrices, but it directly generalizes to arbitrary sparse matrices; see also [8, Section 5]. In [8], Theorem 3.1 was used by Benzi and Simoncini to analyze the decay behavior of Laplace-Stieltjes matrix functions. Similar to Bernstein functions, Laplace-Stieltjes functions can be defined using an integral transform involving exponentials. Many of the arguments that we use in the derivation of our results closely follow the techniques used in [8].

By exploiting the relation between Bernstein functions (2.1) and the matrix exponential, we can prove the following result.

LEMMA 3.2. *Let f be a Bernstein function (2.1). Let $A \in \mathbb{C}^{n\times n}$ be positive semidefinite with spectral radius $\rho(A)$, and denote by $d(i,j)$ the geodesic distance of the nodes i and j in the graph of A. Then, for all i, j with $d(i,j) \geq 2$, we have*

$$\begin{aligned} |[f(A)]_{ij}| \leq 10 &\int_0^{\frac{2d(i,j)}{\rho(A)}} \frac{4\exp(-\frac{1}{4}\rho(A)t)}{\rho(A)t} \left(\frac{e\rho(A)t}{4d(i,j)}\right)^{d(i,j)} \mathrm{d}\mu(t) \\ &+ 10 \int_{\frac{2d(i,j)}{\rho(A)}}^{\frac{d(i,j)^2}{\rho(A)}} \exp\left(-\frac{4d(i,j)^2}{5\rho(A)t}\right) \mathrm{d}\mu(t) \\ &+ \int_{\frac{d(i,j)^2}{\rho(A)}}^{\infty} |[\exp(-tA)]_{ij}| \, \mathrm{d}\mu(t). \end{aligned} \tag{3.1}$$

Proof. For $i \neq j$, we have

$$|[I - \exp(-tA)]_{ij}| = |\,[\exp(-tA)]_{ij}\,|,$$

so that (2.4) implies

$$|[f(A)]_{ij}| \leq |b \cdot a_{ij}| + \int_0^{\infty} |\,[\exp(-tA)]_{ij}\,| \, \mathrm{d}\mu(t). \tag{3.2}$$

We can therefore use bounds for the entries of the matrix exponential in order to bound (3.2). Further note that for i, j with $d(i,j) \geq 2$, it directly follows that $|b \cdot a_{ij}| = 0$, so that we can ignore this term. By recasting the conditions for t in Theorem 3.1 as $\frac{d(i,j)}{2\rho} \leq t \leq \frac{d(i,j)^2}{4\rho}$ or $t \leq \frac{d(i,j)}{2\rho}$, respectively, and writing $\rho = \frac{1}{4}\rho(A)$, the assertion of the lemma follows. ∎

The integral representation (3.1) does not give a clear picture of the actual decay behavior at first sight, and in general it can only be evaluated by numerical quadrature. For fractional

powers $f(z) = z^\alpha, \alpha \in (0,1)$, the special case that we are most interested in, and in particular for the square root $f(z) = \sqrt{z}$, we can give analytic expressions for all occurring integrals in terms of the special functions introduced in Section 2.2.

THEOREM 3.3. *Let $A \in \mathbb{C}^{n\times n}$ be positive semidefinite with spectral radius $\rho(A)$. Let $\alpha \in (0,1)$, and denote by $d(i,j)$ the distance of the nodes i and j in the graph of A. Then, for all i,j with $d(i,j) \geq 2$, we have*

(3.3)
$$|[A^\alpha]_{ij}| \leq \frac{\alpha}{\Gamma(1-\alpha)} \cdot \left(\frac{10e^{d(i,j)}\rho(A)^\alpha}{4^\alpha d(i,j)^{d(i,j)}} \cdot \gamma\left(d(i,j)-\alpha-1, \frac{d(i,j)}{2}\right) + 10\left(\frac{5\rho(A)}{4d(i,j)^2}\right)^\alpha \cdot \left(\Gamma\left(\alpha, \frac{4}{5}\right) - \Gamma\left(\alpha, \frac{2d(i,j)}{5}\right)\right) + \frac{\rho(A)^\alpha}{\alpha \cdot d(i,j)^{2\alpha}}\right).$$

In particular, for $\alpha = \frac{1}{2}$ we have

(3.4)
$$|[\sqrt{A}]_{ij}| \leq \frac{1}{2\sqrt{\pi}} \cdot \left(\frac{10e^{d(i,j)}\sqrt{\pi\rho(A)}}{2d(i,j)^{d(i,j)}} \cdot \gamma\left(d(i,j)-\frac{3}{2}, \frac{d(i,j)}{2}\right) + \frac{5\sqrt{5\pi\rho(A)}}{d(i,j)}\left(\operatorname{erfc}\left(\frac{2}{\sqrt{5}}\right) - \operatorname{erfc}\left(\sqrt{\frac{2d(i,j)}{5}}\right)\right) + \frac{2\sqrt{\rho(A)}}{d(i,j)}\right).$$

Proof. We use the representation (2.2) of fractional powers and insert the result of Lemma 3.2. For the third integral, note that

$$|[\exp(-tA)]_{ij}| \leq \|\exp(-tA)\|_2 = \exp(-t\rho(A)) \leq 1$$

because A is positive semidefinite. This way, we obtain

(3.5)
$$|[A^\alpha]_{ij}| \leq \frac{\alpha}{\Gamma(1-\alpha)} \cdot \left(10\int_0^{\frac{2d(i,j)}{\rho(A)}} \frac{4\exp(-\frac{1}{4}\rho(A)t)}{\rho(A)t}\left(\frac{e\rho(A)t}{4d(i,j)}\right)^{d(i,j)} \cdot t^{-\alpha-1}\,\mathrm{d}t + 10\int_{\frac{2d(i,j)}{\rho(A)}}^{\frac{d(i,j)^2}{\rho(A)}} \exp\left(-\frac{4d(i,j)^2}{5\rho(A)t}\right) \cdot t^{-\alpha-1}\,\mathrm{d}t + \int_{\frac{d(i,j)^2}{\rho(A)}}^{\infty} t^{-\alpha-1}\,\mathrm{d}t\right).$$

We handle the three integrals in (3.5) one after the other now. First, consider

$$\int_0^{\frac{2d(i,j)}{\rho(A)}} \frac{4\exp(-\frac{1}{4}\rho(A)t)}{\rho(A)t}\left(\frac{e\rho(A)t}{4d(i,j)}\right)^{d(i,j)} \cdot t^{-\alpha-1}\,\mathrm{d}t = \frac{4}{\rho(A)}\left(\frac{e\rho(A)}{4d(i,j)}\right)^{d(i,j)} \int_0^{\frac{2d(i,j)}{\rho(A)}} \exp\left(-\frac{1}{4}\rho(A)t\right) \cdot t^{d(i,j)-\alpha-2}\,\mathrm{d}t.$$

To find an antiderivative of this function, first note that $\frac{\partial}{\partial s}\Gamma(z,s) = -\exp(-s)\cdot s^{z-1}$, which follows directly from the integral representation (2.3). Now, for a general function of the form $\exp(-mt)\cdot t^k$, with $m > 0, k > -1$, we have

(3.6)
$$\int \exp(-mt)\cdot t^k\,\mathrm{d}t = \int \exp(-u)\cdot\left(\frac{u}{m}\right)^k m^{-1}\,\mathrm{d}u = m^{-k-1}\int \exp(-u)\cdot u^k\,\mathrm{d}u,$$

where $u = mt$. From the considerations above, we clearly have

$$\int \exp(-u) \cdot u^k \, \mathrm{d}u = -\Gamma(k+1, u) + c$$

for some constant c. Backsubstituting $u = mt$ and inserting into (3.6) yields the antiderivative

$$\int \exp(-mt) \cdot t^k \, \mathrm{d}t = -m^{-k-1} \cdot \Gamma(k+1, mt) + c. \tag{3.7}$$

Using the choice $m = \frac{\rho(A)}{4}$ and $k = d(i,j) - \alpha - 2$ and inserting the limits of integration gives

$$\begin{aligned}
&\int_0^{\frac{2d(i,j)}{\rho(A)}} \exp\left(-\frac{1}{4}\rho(A)t\right) \cdot t^{d(i,j)-\alpha-2} \, \mathrm{d}t \\
&\quad = -\left(\frac{4}{\rho(A)}\right)^{d(i,j)-\alpha-1} \left(\Gamma\left(d(i,j)-\alpha-1, \frac{d(i,j)}{2}\right) - \Gamma(d(i,j)-\alpha-1)\right) \\
&\quad = \left(\frac{4}{\rho(A)}\right)^{d(i,j)-\alpha-1} \gamma\left(d(i,j)-\alpha-1, \frac{d(i,j)}{2}\right).
\end{aligned} \tag{3.8}$$

The second integral in (3.5) is of the general form $\exp(-\frac{m}{t}) \cdot t^{-\alpha-1}$, with $m > 0$, for which we find the antiderivative

$$\int \exp\left(-\frac{m}{t}\right) \cdot t^{-\alpha-1} \, \mathrm{d}t = m^{-\alpha} \cdot \Gamma\left(\alpha, \frac{m}{t}\right) + c$$

for some constant c by proceeding similarly to the derivation of (3.7) above. With the choice $m = \frac{4d(i,j)^2}{5\rho(A)}$ and inserting the limits of integration, we obtain

$$\begin{aligned}
&\int_{\frac{2d(i,j)}{\rho(A)}}^{\frac{d(i,j)^2}{\rho(A)}} \exp\left(-\frac{4d(i,j)^2}{5\rho(A)t}\right) \cdot t^{-\alpha-1} \, \mathrm{d}t \\
&\quad = \left(\frac{5\rho(A)}{4d(i,j)^2}\right)^{\alpha} \left(\Gamma\left(\alpha, \frac{4d(i,j)^2}{5\rho(A)} \cdot \frac{\rho(A)}{d(i,j)^2}\right) - \Gamma\left(\alpha, \frac{4d(i,j)^2}{5\rho(A)} \cdot \frac{\rho(A)}{2d(i,j)}\right)\right) \\
&\quad = \left(\frac{5\rho(A)}{4d(i,j)^2}\right)^{\alpha} \left(\Gamma\left(\alpha, \frac{4}{5}\right) - \Gamma\left(\alpha, \frac{2d(i,j)}{5}\right)\right).
\end{aligned} \tag{3.9}$$

Finally, for the third integral, the antiderivative is simply

$$\int t^{-\alpha-1} \, \mathrm{d}t = -\frac{1}{\alpha} t^{-\alpha} + c$$

for a constant c. After inserting the limits of integration, we directly obtain

$$\int_{\frac{d(i,j)^2}{\rho(A)}}^{\infty} t^{-\alpha-1} \, \mathrm{d}t = \frac{\rho(A)^{\alpha}}{\alpha \cdot d(i,j)^{2\alpha}}. \tag{3.10}$$

Inserting (3.8), (3.9), and (3.10) into (3.5) now yields (3.3). The more compact formula (3.4) for the special case $\alpha = \frac{1}{2}$ directly follows by using $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ together with the relation $\Gamma(z, \frac{1}{2}) = \sqrt{\pi} \operatorname{erfc}(\sqrt{z})$. $\square$

Numerical experiments illustrating the quality of the bounds obtained from Theorem 3.3 will be given in Section 4.1.

3.2. The positive definite case. The bounds derived in Section 3.1 are obviously also valid when A is positive definite and not just semidefinite. However, in this case, the bounds can be sharpened by using the following observation.

PROPOSITION 3.4. *Let $A \in \mathbb{C}^{n\times n}$, and let $\sigma \in \mathbb{C}$. Then,*

$$\exp(A + \sigma I) = \exp(\sigma)\exp(A). \tag{3.11}$$

Using (3.11), we can first shift the smallest eigenvalue of A to zero and then apply the result of Theorem 3.1.

COROLLARY 3.5. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite with smallest and largest eigenvalue $\lambda_{\min}$ and $\lambda_{\max}$, respectively. Denote by $d(i,j)$ the geodesic distance of the nodes i and j in the graph of A, and let $\rho := (\lambda_{\max} - \lambda_{\min})/4$. Then, for $i \neq j$,*

(i) for $\rho t \geq 1$ and $\sqrt{4\rho t} \leq d(i,j) \leq 2\rho t$,

$$|[\exp(-tA)]_{ij}| \leq 10\exp(-t\lambda_{\min})\exp\left(-\frac{1}{5\rho t}d(i,j)^2\right),$$

(ii) for $d(i,j) \geq 2\rho t$

$$|[\exp(-tA)]_{ij}| \leq 10\frac{\exp(-(\rho+\lambda_{\min})t)}{\rho t}\left(\frac{e\rho t}{d(i,j)}\right)^{d(i,j)}.$$

Proof. Define the shifted matrix $\widetilde{A} = A - \lambda_{\min} I$ with eigenvalues in $[0, \lambda_{\max} - \lambda_{\min}]$. From (3.11), it then follows that

$$\exp(-t\widetilde{A}) = \exp(-tA + t\lambda_{\min} I) = \exp(t\lambda_{\min})\exp(-tA),$$

which is equivalent to $\exp(-tA) = \exp(-t\lambda_{\min})\exp(-t\widetilde{A})$. The result then follows by applying Theorem 3.1 to $\exp(-t\widetilde{A})$. ∎

Corollary 3.5 directly gives rise to a result analogous to Lemma 3.2 for the positive definite case.

LEMMA 3.6. *Let f be a Bernstein function (2.1). Let $A \in \mathbb{C}^{n\times n}$ be positive definite with smallest and largest eigenvalue $\lambda_{\min}$ and $\lambda_{\max}$, respectively. Denote by $d(i,j)$ the geodesic distance of the nodes i and j in the graph of A, and let $\rho := (\lambda_{\max} - \lambda_{\min})/4$. Then, for all i,j with $d(i,j) \geq 2$, we have*

$$\begin{aligned} |[f(A)]_{ij}| \leq\ & 10\int_0^{\frac{d(i,j)}{2\rho}} \frac{\exp(-(\rho+\lambda_{\min})t)}{\rho t}\left(\frac{e\rho t}{d(i,j)}\right)^{d(i,j)} \mathrm{d}\mu(t) \\ & + 10\int_{\frac{d(i,j)}{2\rho}}^{\frac{d(i,j)^2}{4\rho}} \exp(-t\lambda_{\min})\cdot\exp\left(-\frac{d(i,j)^2}{5\rho t}\right)\mathrm{d}\mu(t) \\ & + \int_{\frac{d(i,j)^2}{4\rho}}^{\infty} \exp(-t\lambda_{\min})\,\mathrm{d}\mu(t). \end{aligned} \tag{3.12}$$

In contrast to the integrals arising in the positive semidefinite case in Lemma 3.2, even for the special case of fractional powers z^α, there is no closed-form expression for the more complicated second integral in (3.12). Thus, in order to use Lemma 3.6 for predicting the decay in $f(A)$ for positive definite A, this integral needs to be evaluated by numerical quadrature.

3.3. Explicit decay bounds for fractional powers of positive definite matrices via Cauchy-Stieltjes functions. For fractional powers, explicit decay bounds can also be obtained in a different way when A is positive definite by employing a simple trick. For this, we write

$$A^\alpha = A \cdot A^{\alpha-1} \tag{3.13}$$

and exploit the fact that $A^{\alpha-1}$ is a Cauchy-Stieltjes function when $\alpha \in (0,1)$. Using the relation (3.13), known decay results for Cauchy-Stieltjes functions [8, 21] can easily be transferred to positive fractional powers. A similar trick is used in the context of extending the scope of restarted Krylov subspace methods for Stieltjes matrix functions in [19, 20].

THEOREM 3.7. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite with condition number $\kappa = \lambda_{\max}/\lambda_{\min}$. Let $\alpha \in (0,1)$, and denote by $d(i,j)$ the geodesic distance of the nodes i and j in the graph of A. Then, for all i,j with $d(i,j) \geq 2$,*

$$|[A^\alpha]_{ij}| \leq 2\lambda_{\min}^{\alpha-1}\|A\|_\infty \cdot q^{d(i,j)-1} \quad \textit{with} \quad q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}. \tag{3.14}$$

Proof. Define $B = A^{\alpha-1}$, so that $A^\alpha = AB$, or, written elementwise,

$$[A^\alpha]_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{3.15}$$

As $z^{\alpha-1}$ is a Cauchy-Stieltjes function, we can apply [21, Theorem 4] to B, which states that

$$|b_{kj}| \leq 2\lambda_{\min}^{\alpha-1} \cdot q^{d(k,j)}, \tag{3.16}$$

where q is as defined in (3.14). When $a_{ik} \neq 0$ we clearly have $d(k,j) \geq d(i,j) - 1$. Using this relation after inserting (3.16) into (3.15) and taking the absolute value gives

$$|[A^\alpha]_{ij}| \leq 2\lambda_{\min}^{\alpha-1} \cdot q^{d(i,j)-1} \sum_{k=1}^{n} |a_{ik}| \leq 2\lambda_{\min}^{\alpha-1} \cdot q^{d(i,j)-1}\|A\|_\infty,$$

which concludes the proof. ∎

Note that the technique used in the proof of Theorem 3.7 cannot be applied when A is only positive semidefinite, as $A^{\alpha-1}$ is not defined when A has a zero eigenvalue. Thus, the result cannot be extended to this situation.

REMARK 3.8. The bound (3.14) is stated in a rather simple form that is valid for all i,j with $d(i,j) \geq 2$. When one is interested in a specific entry $|[A^\alpha]_{ij}|$, the bound can be sharpened to

$$|[A^\alpha]_{ij}| \leq 2\lambda_{\min}^{\alpha-1} \cdot \min\{\|A_{i:}\|_1, \|A_{:j}\|_1\} \cdot q^{d(i,j)-1},$$

with q as in (3.14), where $A_{i:}, A_{:j}$ denote the ith row and jth column of A, respectively. This directly follows from the fact that both the bounds obtained from writing $A^\alpha = AA^{1-\alpha}$ and from writing $A^\alpha = A^{1-\alpha}A$ are valid for each entry, so that one can always select the smaller of the two.

EXAMPLE 3.9. To illustrate our decay bounds for the positive definite case, we examine a simple model problem that is frequently used for demonstrating the quality of decay bounds [8, 23]. Let $A = I \otimes M + M \otimes I \in \mathbb{C}^{N^2\times N^2}$, where $M = \operatorname{tridiag}(-1, 2+\sigma, -1) \in \mathbb{C}^{N\times N}$ and $\sigma \geq 0$. For the shift choice $\sigma = 0$, the matrix A corresponds to the discretization of

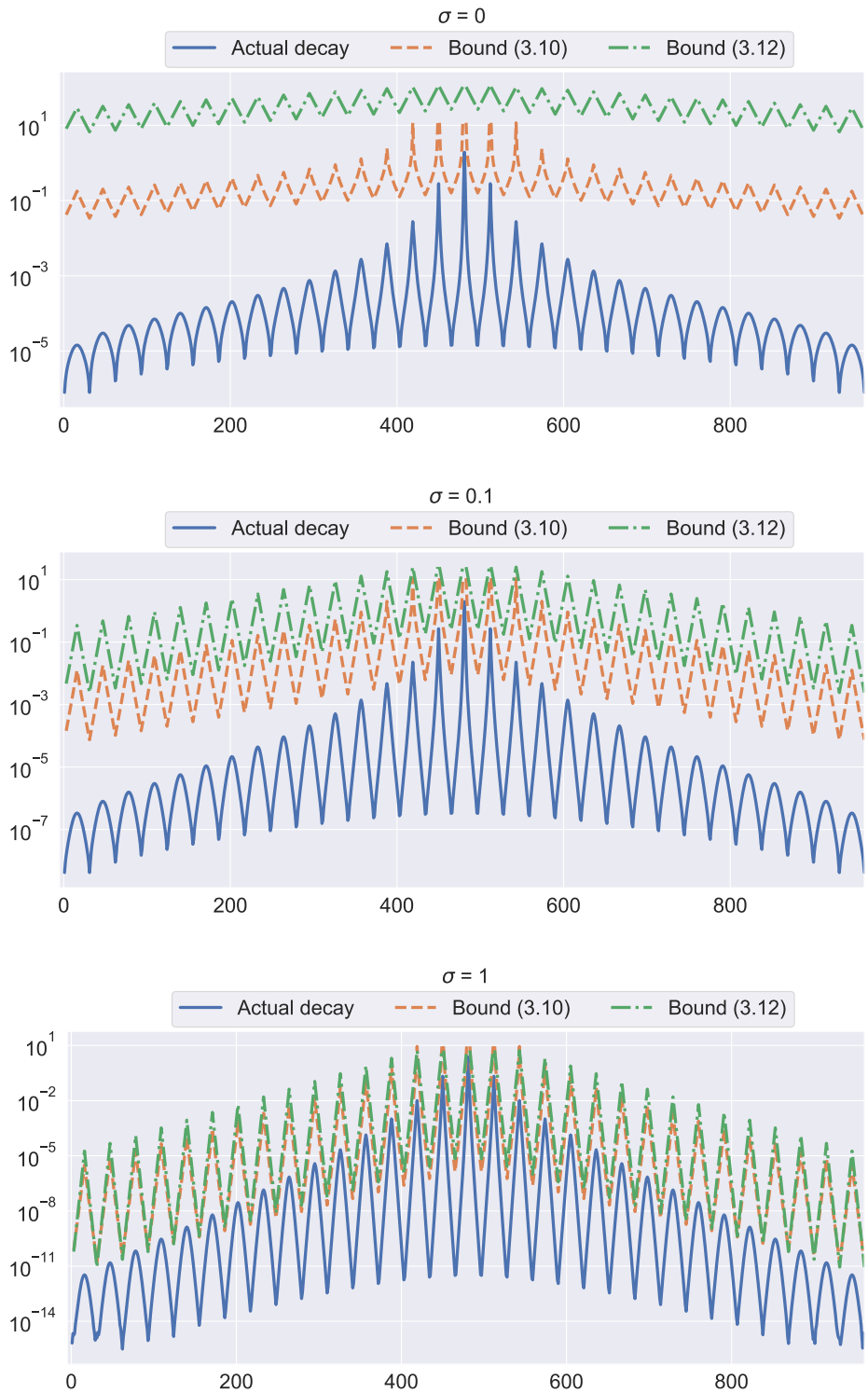

FIG. 3.1. *Decay in one column of* $\sqrt{A}$, *where* $A = I \otimes M + M \otimes I \in C^{961\times 961}$ *and where* $M = \mathrm{tridiag}(-1, 2+\sigma, -1) \in \mathbb{C}^{31\times 31}$ *for* $\sigma = 0$ *(top),* $\sigma = 0.1$ *(center), and* $\sigma = 1$ *(bottom). The graph of* A *is a regular two-dimensional grid of size* 31×31, *and the depicted column corresponds to the node in the center of this grid.*

the two-dimensional Laplace operator on a regular square grid with homogeneous Dirichlet boundary conditions (up to a scaling). Increasing the shift σ makes the resulting matrix A better conditioned. We are interested in the decay of the matrix square root $\sqrt{A}$, which plays an important role in Dirichlet-to-Neumann maps; see, e.g., [1, 2, 15].

For our experiment, we choose $N = 31$, resulting in a matrix A of size 961×961, and we compare the integral-based bound for Bernstein functions (3.12) and the bound (3.14) for A^α derived from the connection to Cauchy-Stieltjes functions. We choose three different shift parameters $\sigma = 0$, $\sigma = 0.1$, and $\sigma = 1$. We consider the magnitude of the entries of the column of $\sqrt{A}$ that belongs to the node in the center of the graph of A, i.e., at grid position $(16, 16)$. The integrals in the bound (3.12) are approximated using the general-purpose quadrature routine `quad` from `SciPy.integrate`. The results of this experiment are depicted in Figure 3.1. Note that the seemingly "oscillatory" behavior of the entries of $\sqrt{A}$ is caused by the row-wise numbering of the grid nodes. If plotted over the two-dimensional grid, then one would observe a smooth decay with respect to the geodesic graph distance, as expected.

We observe that the integral bound (3.12) always lies below the bound (3.14), with the distance between the bounds reducing when the shift σ is increased. Another interesting observation is that for $\sigma = 0$, the bound (3.12) much better resembles the actual slope of the decay for nodes nearby the center of the grid, as it is not restricted to a simple exponential decay of the form $C \cdot q^{d(i,j)}$.

4. Application to the fractional graph Laplacian. Given an undirected graph $G = (V, E)$ with $|V| = n$, the *graph Laplacian* L_G of G is defined as

$$L_G = D_G - A_G,$$

where A_G is the adjacency matrix of G and D_G is a diagonal matrix containing the degrees of the nodes of G on the diagonal, i.e., $[D_G]_{ii} = \sum_{j=1}^n a_{ij}, i = 1, \ldots, n$. The graph Laplacian has applications in modeling diffusion processes on graphs but also in spectral clustering [40], graph drawing algorithms [30], and many other areas. As all row sums of the graph Laplacian are equal to zero, it is necessarily a singular matrix, and it is well known that it is always positive semidefinite [34].

Recently, interest in the *fractional graph Laplacian* has emerged, which allows one to model nonlocal diffusion processes on graphs or to use nonlocal random walks for the exploration of large networks [5, 9, 11, 12, 17]. The fractional graph Laplacian is simply defined via a fractional power of the ordinary Laplacian, i.e., by L_G^α. As L_G^α is a singular M-matrix [5] with all entries nonzero (if G is connected), it can be interpreted as the Laplacian of a *weighted, fully-connected graph* G^α on the same set V of nodes. In this context, decay bounds for the entries of L_G^α are of interest because they give insight into the nature of the connection strength in G^α between nodes that were not connected in G.

4.1. Power law decay in the fractional Laplacian. In [5, 12], it was observed that the entries of the fractional graph Laplacian L_G^α exhibit a power-law decay away from the sparsity pattern of L_G. In particular, for instance, the following result was shown, which is based on Jackson's theorem [33].

THEOREM 4.1 (Corollary 3.1 in [5]). *Let L_G be the Laplacian of an undirected graph G, and let $\alpha \in (0, 1)$. Then, if $d(i, j) \geq 2$, we have*

$$|(L_G^\alpha)_{ij}| \leq c \cdot \left(\frac{\rho(L_G)}{2}\right)^\alpha \cdot (d(i,j) - 1)^{-\alpha} \tag{4.1}$$

with $c = 1 + \pi^2/2$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We now compare our new result, Theorem 3.3, to Theorem 4.1. An important observation concerning (3.3) and (3.4) is that the first term in the sum goes to zero exponentially in $d(i,j)$, so that asymptotically, the second and third term control the decay behavior in L_G^α. So, Theorem 3.3 gives an asymptotic decay estimate of the form

$$|[L_G^\alpha]_{ij}| \lesssim C \cdot d(i,j)^{-2\alpha}, \tag{4.2}$$

where C is a constant, indicating a power-law decay as already observed in [5, 12] but with the improved exponent -2α instead of $-\alpha$. Thus, our new bounds show that the strength of the connection between far apart nodes in G^α must actually drop off faster than known so far.

REMARK 4.2. The bound (4.2) only holds in an asymptotic sense because we ignore the influence of the first term in (3.3). It can, however, also easily be cast into an explicit, non-asymptotic form. Taking, e.g., $\alpha = \frac{1}{2}$, we have that

$$\frac{e^d}{d^d}\gamma\left(d - 3/2, \frac{d}{2}\right) \leq d^{-1}$$

holds for all $d \geq 4$. Thus, by Theorem 3.3 we directly find a rigorous, non-asymptotic bound of the form

$$|[\sqrt{L_G}]_{ij}| \leq \widetilde{C} \cdot d(i,j)^{-1} \qquad \text{for all } i,j \text{ with } d(i,j) \geq 4$$

with a modified constant $\widetilde{C}$.

EXAMPLE 4.3. To illustrate how our new decay estimates compare to those from [5, 12], we begin by considering a very simple test problem: Let G be a one-dimensional chain of length n. The corresponding graph Laplacian is the tridiagonal matrix

$$L_G = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

By the Geršgorin disk theorem, the spectral radius of L_G is bounded by $\rho(L_G) \leq 4$ independently of n. In Figure 4.1 we compare our new bounds (3.3) and (3.4) from Theorem 3.3 (based on the Lévy-Khintchine integral representation) to the bound (4.1) from Theorem 4.1 (based on Jackson's theorem) for $\alpha = \frac{1}{2}$ and $\alpha = \frac{1}{4}$. In both cases, we can observe that the slope of our new bound more closely resembles the actual decay behavior due to the additional factor 2 in the exponent. In addition, we also better predict the order of magnitude of the entries (though they are still overestimated by a quite large margin). Concerning this, it is also interesting to compare the constants involved in the two bounds. In Theorem 4.1, there is only one relevant constant, which is

$$\left(1 + \frac{\pi^2}{2}\right)\sqrt{\frac{\rho(L_G)}{2}} \approx 8.39 \qquad \text{for } \alpha = \frac{1}{2}$$

and

$$\left(1 + \frac{\pi^2}{2}\right)\left(\frac{\rho(L_G)}{2}\right)^{\frac{1}{4}} \approx 7.06 \qquad \text{for } \alpha = \frac{1}{4}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

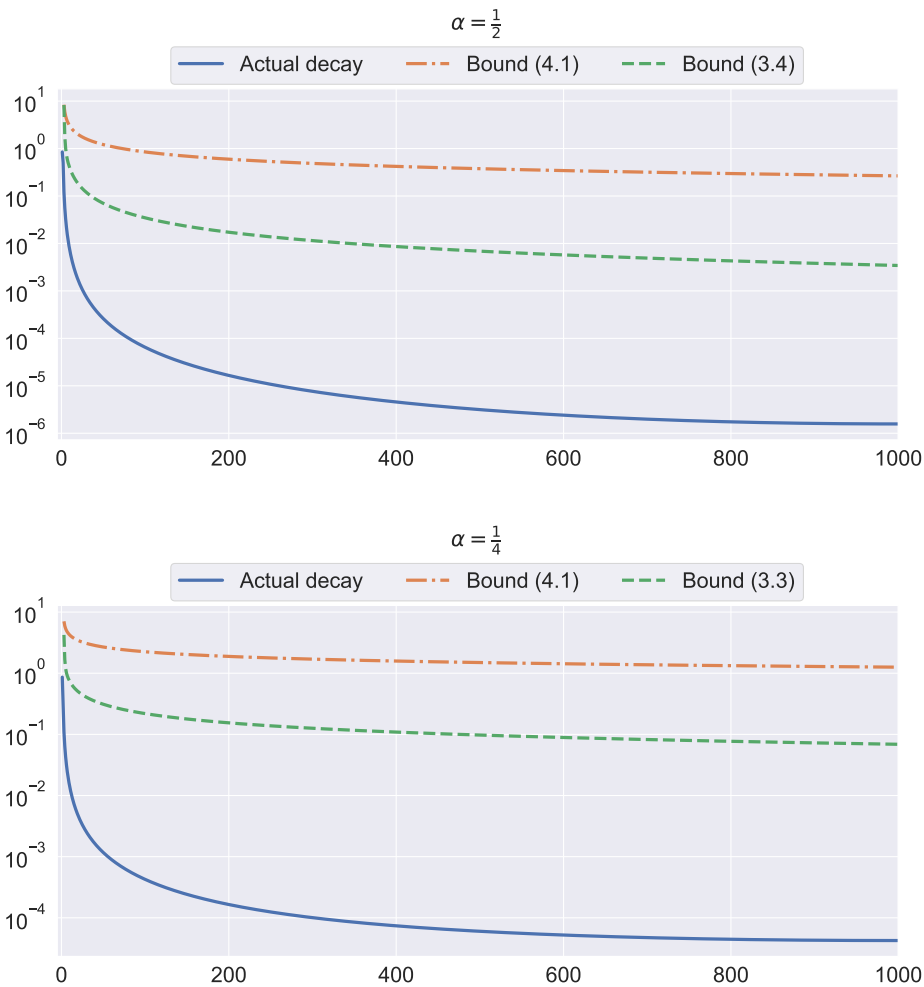

FIG. 4.1. *Decay in the first column of L_G^α, where L_G is the Laplacian corresponding to a one-dimensional chain and $\alpha = \frac{1}{2}$ (top) or $\alpha = \frac{1}{4}$ (bottom).*

In Theorem 3.3, several constants occur. As the first term does not play a role in the asymptotic behavior for growing d, we ignore it. The constant in front of the term $d(i,j)^{-1}$ can be estimated as

$$\frac{5}{2}\sqrt{5\rho(L_G)}\,\mathrm{erfc}\left(\sqrt{\frac{4}{5}}\right) + \sqrt{\frac{\rho(L_G)}{\pi}} \sim 3.43 \qquad \text{for } \alpha = \frac{1}{2}$$

because $\mathrm{erfc}\left(\sqrt{\frac{2d(i,j)}{5}}\right)$ goes to zero as $d(i,j)$ increases. For $\alpha = \frac{1}{4}$, we obtain the constant

$$\frac{\alpha}{\Gamma(1-\alpha)}\left(10\left(\frac{5\rho(L_G)}{4}\right)^{\alpha} \cdot \Gamma\left(\alpha, \frac{4}{5}\right) + \frac{\rho(L_G)^{\alpha}}{\alpha}\right) \sim 2.18,$$

again making use of the fact that $\Gamma\left(\alpha, \frac{2d(i,j)}{5}\right)$ goes to zero for increasing $d(i,j)$. Thus, the

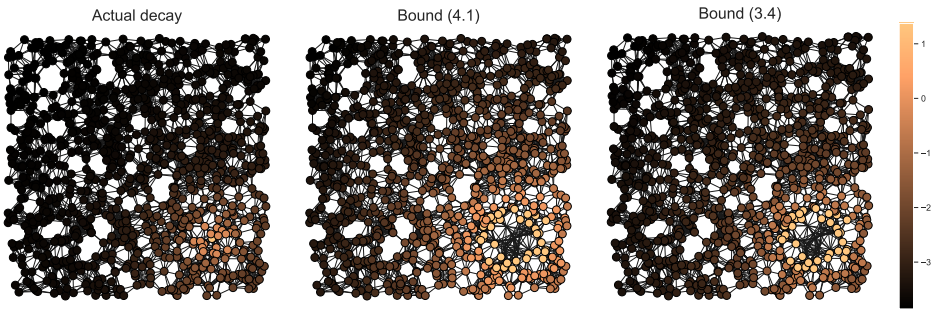

FIG. 4.2. *Decay in one column of the fractional Laplacian $\sqrt{L_G}$ for a random geometric graph G on a logarithmic scale. Left: actual decay; center: bound (4.1); right: bound (3.4). In the center and right graph, nodes for which no bound is available are not drawn.*

constants in the bounds have—at least asymptotically for large distances—also improved by a factor of roughly 2.4 and 3.2, respectively.

EXAMPLE 4.4. Next, we compare the same bounds as in the previous example but for an example graph with a more irregular structure. We construct a *random geometric graph* by sampling $n = 1000$ uniformly distributed points in the unit square and then connecting all pairs with distance below 0.075 by an edge. The resulting graph is depicted in Figure 4.2. The degrees of the nodes in this graph range from 4 to 29, and the spectral radius of its Laplacian is $\rho(L_G) \approx 31.64$. In the left-most plot in Figure 4.2, the color coding of the nodes depicts the magnitude of the entries in one column of the fractional Laplacian $\sqrt{L_G}$ of this graph (on a logarithmic scale). The column corresponds to a node near the bottom right corner of the unit square (the node with brightest color in the plot). One can nicely see how the magnitude of the entries decays the farther one moves away from the source node, which is in line with what our decay bounds predict (and what one would intuitively expect). In the center part of Figure 4.2, the decay bound (4.1) is shown, and in the right-most part we plot our new bound (3.4). In both cases, nodes for which no bound is available (i.e., nodes with distance 1 from the source node) are not drawn. Additionally, for easier comparison, we also show the magnitude of the entries and the bounds in a line plot in Figure 4.3, which also contains results for the case $\alpha = \frac{1}{4}$. As the graph has no regular underlying structure, we order the nodes for this plot according to the magnitude of the corresponding column entries, so that the entries form a monotonically decreasing sequence, which makes it easier to understand the plot. Observe the resulting "stair-case" like structure of both bounds (3.3) and (3.4) and (4.1), which predict the same order of magnitude for all entries belonging to nodes that have the same distance from the source node. While less pronounced, we can also observe a similar structure with plateau-like areas in the actual decay.

Again our new bounds (3.3) and (3.4) more accurately predict the slope of the decay and also have a smaller magnitude than (4.1). However, both bounds overestimate the actual decay by quite a large margin, and the actual decay slope is a little bit steeper than what the new bound predicts.

4.2. Investigating the sharpness of the decay bounds. An interesting question in the study of decay bounds is whether they are asymptotically optimal or whether there is a possibility for further improvement. Currently, we do not have a definitive answer to this question, but we give an illustrative example that suggests that a further improvement of the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

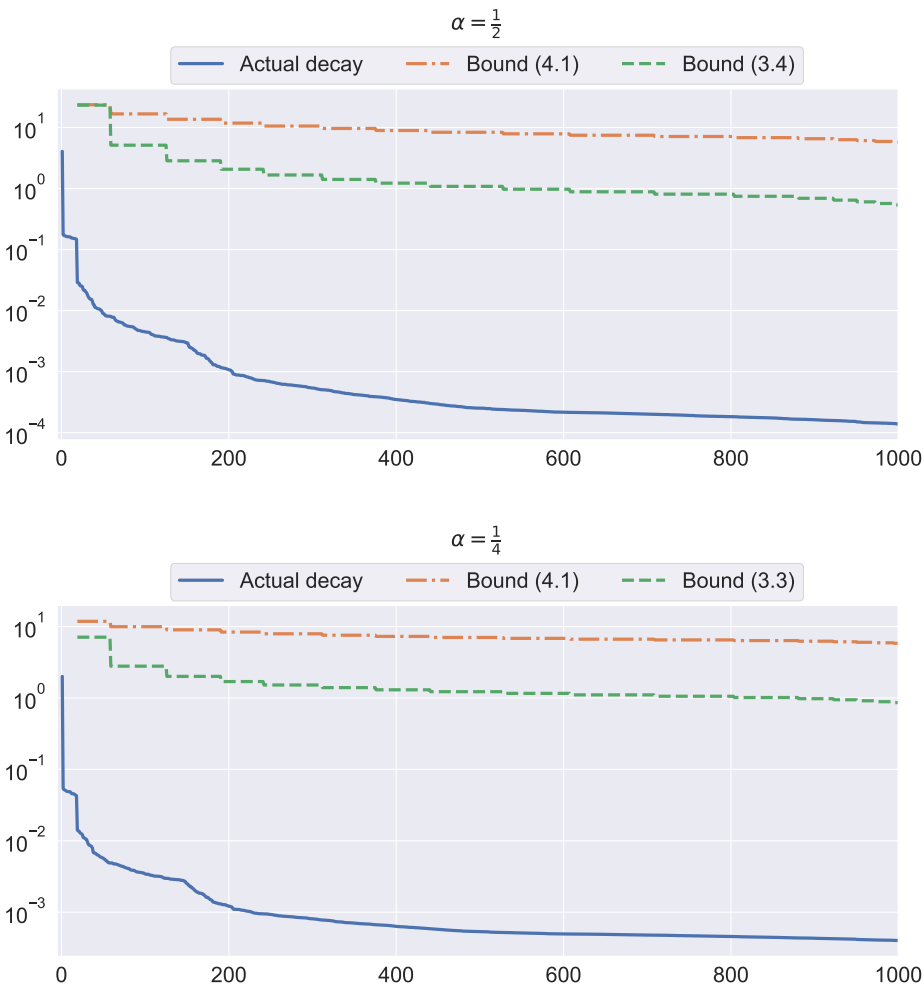

FIG. 4.3. *Decay in the first column of L_G^α, where L_G is the Laplacian corresponding to a random geometric graph G (see text for details) and $\alpha = \frac{1}{2}$ (top) or $\alpha = \frac{1}{4}$ (bottom).*

exponent in the power law might be possible.

EXAMPLE 4.5. We consider a simple example graph for which we can derive analytic formulas for the entries of the fractional Laplacian. Let G_n be a graph consisting of n nodes arranged in a circle, where we assume that n is odd. The Laplacian of this graph is given by

$$L_{G_n} = \begin{bmatrix} 2 & -1 & & & -1 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ -1 & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

The eigenvalues and eigenvectors of this matrix are analytically known and given by

$$\lambda_k = \begin{cases} 4\sin^2\left(\frac{\pi k}{2n}\right) & \text{if } k \text{ is even,} \\ 4\sin^2\left(\frac{\pi(k-1)}{2n}\right) & \text{if } k \text{ is odd,} \end{cases} \tag{4.3}$$

and

$$\boldsymbol{v}_{i,k} = \begin{cases} n^{-1/2} & \text{if } k = 1, \\ \sqrt{\frac{2}{n}}\sin\left(\frac{\pi(i-\frac{1}{2})k}{n}\right) & \text{if } k \text{ is even,} \\ \sqrt{\frac{2}{n}}\cos\left(\frac{\pi(i-\frac{1}{2})(k-1)}{n}\right) & \text{if } k \text{ is odd,} \end{cases} \tag{4.4}$$

respectively. In particular, all eigenvalues except $\lambda_1 = 0$ appear twice, and the spectrum of L_{G_n} is contained in $[0, 4]$ independent of n. Using (4.3) and (4.4), we can analytically compute the entries of $\sqrt{L_{G_n}}$. We have

$$\begin{aligned} [\sqrt{L_{G_n}}]_{ij} &= \sum\nolimits_{k=1}^{n} \sqrt{\lambda_k}\boldsymbol{v}_{i,k}\boldsymbol{v}_{j,k} \\ &= \sum\nolimits_{\ell=1}^{\frac{n-1}{2}} \sqrt{\lambda_{2\ell}}(\boldsymbol{v}_{i,2\ell}\boldsymbol{v}_{j,2\ell} + \boldsymbol{v}_{i,2\ell+1}\boldsymbol{v}_{j,2\ell+1}) \\ &= \frac{4}{n}\sum\nolimits_{\ell=1}^{\frac{n-1}{2}} \sin\left(\frac{\pi\ell}{n}\right)\left(\sin\left(\frac{2\pi(i-\frac{1}{2})\ell}{n}\right)\sin\left(\frac{2\pi(j-\frac{1}{2})\ell}{n}\right)\right. \\ &\qquad\qquad \left. + \cos\left(\frac{2\pi(i-\frac{1}{2})\ell}{n}\right)\cos\left(\frac{2\pi(j-\frac{1}{2})\ell}{n}\right)\right) \\ &= \frac{4}{n}\sum\nolimits_{\ell=1}^{\frac{n-1}{2}} \sin\left(\frac{\pi\ell}{n}\right)\cos\left(\frac{2\pi\ell(i-j)}{n}\right), \end{aligned} \tag{4.5}$$

where we used an angle sum identity for the last equality. We can resolve the summation in (4.5) using several standard trigonometric identities, yielding

$$[\sqrt{L_{G_n}}]_{ij} = \frac{1}{n}\left(\cot\left(\frac{\pi(2(i-j)+1)}{2n}\right) + \cot\left(\frac{\pi(1-2(i-j))}{2n}\right)\right). \tag{4.6}$$

For the $(\lceil\frac{n}{2}\rceil, 1)$-entry, the above formula (4.6) simplifies to

$$[\sqrt{L_{G_n}}]_{\lceil\frac{n}{2}\rceil,1} = \frac{1}{n}\cot\left(\pi\left(\frac{1}{n} - \frac{1}{2}\right)\right). \tag{4.7}$$

From L'Hôpital's rule, one can see that (4.7) implies that as n goes to infinity, $[\sqrt{L_{G_n}}]_{\lceil\frac{n}{2}\rceil,1}$ goes to zero as $(\lceil\frac{n}{2}\rceil - 1)^{-2}$. In contrast, the decay bound of Theorem 3.3 predicts a decrease of $(\lceil\frac{n}{2}\rceil - 1)^{-1}$.

In the numerical experiments in Example 4.3 and 4.4—as well as in other numerical experiments not reported here—we observed that the actual decay was even faster than predicted by our new, refined bounds, and we were not able to find a graph for which the bound was (provably or experimentally) asymptotically sharp. Thus, motivated by Example 4.5, we conjecture that

$$|[L_G^\alpha]_{ij}| \lesssim C \cdot d(i,j)^{-4\alpha}.$$

It remains an open topic for future research to prove or disprove this conjecture.

5. Conclusions. We have derived new integral-based decay bounds for Bernstein functions of Hermitian matrices A with special emphasis on the case that A is positive semidefinite and $f(z) = z^\alpha, \alpha \in (0,1)$. In this case, analytic expressions for all involved integrals are available, making the bounds particularly easy to use. In other cases, some of the integrals appearing in the bounds need to be evaluated by numerical quadrature. As a particularly important application, we have considered nonlocal network dynamics described by the fractional graph Laplacian. It is well-known that the strength of connection between far apart nodes in the network follows a power law in this case, and we were able to improve the exponent of this power law from $-\alpha$ to -2α using our new approach. Motivated by studying the closed form representation of the fractional Laplacian of a cycle graph, we conjectured that a further improvement up to an exponent of -4α could be possible.

Acknowledgement. The author wishes to thank Fabio Durastante and Andreas Frommer for helpful comments on an earlier version of the manuscript, as well as an anonymous referee for carefully reading the manuscript and suggesting several improvements of the presentation.

REFERENCES

[1] M. ARIOLI AND D. LOGHIN, *Matrix square-root preconditioners for the Steklov-Poincaré operator*, Tech. Report RAL-TR-2008-003, Rutherford Appleton Laboratory, Didcot, 2008

[2] ———, *Discrete interpolation norms with applications*, SIAM J. Numer. Anal., 47 (2009), pp. 2924–2951.

[3] B. BECKERMANN AND L. REICHEL, *Error estimation and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.

[4] M. BENZI, *Localization in matrix computations: theory and applications*, in Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications, M. Benzi and V. Simoncini, eds., vol. 2173 of Lecture Notes in Math., Springer, Cham, 2016, pp. 211–317.

[5] M. BENZI, D. BERTACCINI, F. DURASTANTE, AND I. SIMUNEC, *Non-local network dynamics via fractional graph Laplacians*, J. Complex Netw., 8 (2020), Art. cnaa017, 29 pages.

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

[7] M. BENZI AND N. RAZOUK, *Decay bounds and $O(n)$ algorithms for approximating functions of sparse matrices*, Electron. Trans. Numer. Anal., 28 (2007/08), pp. 16–39. http://etna.ricam.oeaw.ac.at/vol.28.2007-2008/pp16-39.dir/pp16-39.pdf

[8] M. BENZI AND V. SIMONCINI, *Decay bounds for functions of Hermitian matrices with banded or Kronecker structure*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1263–1282.

[9] M. BENZI AND I. SIMUNEC, *Rational Krylov methods for fractional diffusion problems on graphs*, BIT, (62) 2021, pp. 357–385.

[10] C. BERG, *Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity*, in Positive Definite Functions. From Schoenberg to Space-Time Challenges, J. Mateu and E. Porcu, eds., University Jaume I, Castellón de la Plana, 2008, pp. 15–45.

[11] D. BERTACCINI AND F. DURASTANTE, *Nonlocal diffusion of variable order on graphs*, Preprint on arXiv, 2021. https://arxiv.org/abs/2110.05424

[12] D. BIANCHI, M. DONATELLI, F. DURASTANTE, AND M. MAZZA, *Compatibility, embedding and regularization of non-local random walks on graphs*, J. Math. Anal. Appl., 511 (2022), Art. 126020, 30 pages.

[13] D. R. BOWLER AND T. MIYAZAKI, *$O(N)$ methods in electronic structure calculations*, Rep. Prog. Phys., 75 (2012), Art. 036503, 43 pages.

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

[15] V. DRUSKIN AND L. KNIZHNERMAN, *Gaussian spectral rules for the three-point second differences. I. A two-point positive definite problem in a semi-infinite domain*, SIAM J. Numer. Anal., 37 (2000), pp. 403–422.

[16] V. EIJKHOUT AND B. POLMAN, *Decay rates of inverses of banded M-matrices that are near to Toeplitz matrices*, Linear Algebra Appl., 109 (1988), pp. 247–277.

[17] E. ESTRADA, *Path Laplacians versus fractional Laplacians as nonlocal operators on networks*, New J. Phys., 23 (2021), Art. 073049, 20 pages.

[18] N. J. FORD, D. V. SAVOSTYANOV, AND N. L. ZAMARASHKIN, *On the decay of the elements of inverse triangular Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1288–1302.

[19] A. FROMMER, S. GÜTTEL, AND M. SCHWEITZER, *Convergence of restarted Krylov subspace methods for Stieltjes functions of matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1602–1624.

[20] ———, *Efficient and stable Arnoldi restarts for matrix functions based on quadrature*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 661–683.

[21] A. FROMMER, C. SCHIMMEL, AND M. SCHWEITZER, *Bounds for the decay of the entries in inverses and Cauchy-Stieltjes functions of certain sparse, normal matrices*, Numer. Linear Algebra Appl., 25 (2018), Art. e2131, 17 pages.

[22] ———, *Non-Toeplitz decay bounds for inverses of Hermitian positive definite tridiagonal matrices*, Electron. Trans. Numer. Anal., 48 (2018), pp. 362–372. http://etna.ricam.oeaw.ac.at/vol.48.2018/pp362-372.dir/pp362-372.pdf

[23] ———, *Analysis of probing techniques for sparse approximation and trace estimation of decaying matrix functions*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 1290–1318.

[24] P.-L. GISCARD, K. LUI, S. J. THWAITE, AND D. JAKSCH, *An exact formulation of the time-ordered exponential using path-sums*, J. Math. Phys., 56 (2015), Art. 053503, 18 pages.

[25] S. GÜTTEL AND L. KNIZHNERMAN, *A black-box rational Arnoldi variant for Cauchy-Stieltjes matrix functions*, BIT, 53 (2013), pp. 595–616.

[26] S. GÜTTEL AND M. SCHWEITZER, *A comparison of limited-memory Krylov methods for Stieltjes functions of Hermitian matrices*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 83–107.

[27] N. J. HIGHAM, *Functions of Matrices*, SIAM, Philadelphia, 2008.

[28] M. HOCHBRUCK AND C. LUBICH, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.

[29] A. ISERLES, *How large is the exponential of a banded matrix?*, New Zealand J. Math., 29 (2000), pp. 177–192.

[30] Y. KOREN, *On spectral graph drawing*, in Computing and Combinatorics, T. Warnow and B. Zhu, eds., vol. 2697 of Lecture Notes in Comput. Sci., Springer, Berlin, 2003, pp. 496–508.

[31] L. LOPEZ AND A. PUGLIESE, *Decay behaviour of functions of skew-symmetric matrices*, Proceedings of HERCMA 2005, AUEB, Athens, 2005, pp. 22–24.

[32] S. MASSEI AND L. ROBOL, *Rational Krylov for Stieltjes matrix functions: convergence and pole selection*, BIT, 61 (2021), pp. 237–273.

[33] G. MEINARDUS, *Approximation of Functions: Theory and Numerical Methods*, Springer, Berlin, 1967.

[34] R. MERRIS, *Laplacian matrices of graphs: a survey*, Linear Algebra Appl., 197/198 (1994), pp. 143–176.

[35] S. POZZA AND V. SIMONCINI, *Inexact Arnoldi residual estimates and decay properties for functions of non-Hermitian matrices*, BIT, 59 (2019), pp. 969–986.

[36] A. P. RIASCOS AND J. L. MATEOS, *Fractional dynamics on networks: emergence of anomalous diffusion and Lévy flights*, Phys. Rev. E, 90 (2014), Art. 032809, 7 pages.

[37] ———, *Random walks on weighted networks: a survey of local and non-local dynamics*, J. Complex Netw., 9 (2021), Art. cnab032, 39 pages.

[38] R. L. SCHILLING, R. SONG, AND Z. VONDRACEK, *Bernstein Functions. Theory and Applications*, de Gruyter, Berlin, 2012.

[39] C. SCHIMMEL, *Bounds for the Decay in Matrix Functions and its Exploitation in Matrix Computations*, PhD. Thesis, Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Wuppertal, 2020.

[40] U. VON LUXBURG, *A tutorial on spectral clustering*, Stat. Comput., 17 (2007), pp. 395–416.