Fast and accurate approximations to fractional powers of operators

Lidia Aceto*
Dipartimento di Matematica, Università di Pisa, via F. Buonarroti 1/C, 56127 Pisa, Italy
*Corresponding author: lidia.aceto@unipi.it

and

Paolo Novati
Dipartimento di Matematica e Geoscienze, Università di Trieste, via Valerio 12/1, 34127 Trieste, Italy
novati@units.it

In this paper we consider some rational approximations to the fractional powers of self-adjoint positive operators, arising from the Gauss–Laguerre rules. We derive practical error estimates that can be used to select *a priori* the number of Laguerre points necessary to achieve a given accuracy. We also present some numerical experiments to show the effectiveness of our approaches and the reliability of the estimates.

Keywords: fractional Laplacian; matrix functions; Gauss–Laguerre rule.

1. Introduction

The numerical solution of problems involving fractional diffusion can lead to the computation of fractional powers of unbounded operators. For instance, denoting by Δ the standard Laplace operator and taking $\alpha \in (0,1)$, the fractional Laplace equation

$$(-\Delta)^{\alpha} u = f \tag{1.1}$$

on a bounded Lipschitz domain subject to Dirichlet boundary conditions can be solved by computing

$$\sum_{j=1}^{+\infty} \mu_j^{-\alpha} \langle f, \varphi_j \rangle \varphi_j, \tag{1.2}$$

where μ_j and φ_j are the eigenvalues and the eigenfunctions of $-\Delta$, respectively, and $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product. In practice, in this situation the fractional derivative can be identified by the fractional power. Keeping in mind this kind of application, in this work we are interested in the numerical approximation of $\mathscr{L}^{-\alpha}, \alpha \in (0,1)$. Here $\mathscr{L}$ is a self-adjoint positive operator acting in a Hilbert space $\mathscr{H}$ in which the eigenfunctions of $\mathscr{L}$ form an orthonormal basis of $\mathscr{H}$, so that $\mathscr{L}^{-\alpha}$ can be written through the spectral decomposition of $\mathscr{L}$ as in (1.2).

In recent years, this problem has been studied by many authors. Due to the properties of the function $\lambda^{-\alpha}, \lambda \in [\ell, +\infty), \ell > 0$, the most effective approaches are those based on a rational approximation of this function. In the continuous setting of unbounded operators, methods based on the best uniform rational approximation (BURA) of functions closely related to $\lambda^{-\alpha}$ have been considered, for example, in Harizanov *et al.* (2018), Harizanov *et al.* (2019, 2020) and Harizanov & Margenov (2018) by using a modified version of the Remez algorithm. Another class of methods relies on quadrature rules for the integral representation of $\lambda^{-\alpha}$; see Aceto & Novati (2019, 2020), Aceto *et al.* (2019), Bonito & Pasciak (2015) and Vabishchevich (2018, 2020). Very recently, time-stepping methods for a parabolic reformulation of the fractional diffusion equation (1.1) given in Vabishchevich (2015) have also been interpreted in Hofreither (2020) as a rational approximation of $\lambda^{-\alpha}$.

In this paper, starting from the integral representation given in Bonito & Pasciak (2015, Eq. (4)),

$$\mathscr{L}^{-\alpha} = \frac{2\sin(\alpha\pi)}{\pi}\int_0^{+\infty} t^{2\alpha-1}(\mathscr{I}+t^2\mathscr{L})^{-1}\mathrm{d}t, \qquad \alpha \in (0,1), \tag{1.3}$$

where $\mathscr{I}$ is the identity operator in $\mathscr{H}$, after suitable changes of variables we consider an alternative rational approximation based on the truncated Gauss–Laguerre rule. In order to construct the truncated approach, we exploit the error analysis of the standard Gauss–Laguerre rule based on the theory of analytic functions originally introduced in Barrett (1961). We are able to show that in the operator norm the error decay is like

$$\exp(-cm^{1/2}),$$

where m is the number of inversions and $c = 3.6\alpha^{1/2}$ (cf. (6.10)). In this view, the formula seems to be competitive with the Sinc quadrature studied in Bonito & Pasciak (2015) in which $c = \pi(1-\alpha)^{1/2}\alpha^{1/2}$ by Remark 3.1 of the same paper. However, it appears to be slightly slower than that based on the analysis given in and related to the BURA approach in which $c = 2\pi(1-\alpha)^{1/2}$ although the approach presented here does not suffer from the instability of the Remez algorithm.

We also present a further modification of the truncated Gauss–Laguerre rule, called the equalized rule, that allows the number of inversions to be further reduced, to achieve the same accuracy, especially when $\alpha \leqslant 1/2$.

The paper is structured as follows. In Section 2 we present the Gauss–Laguerre approach. In Sections 3 and 4, starting from the error analysis based on the theory of analytic functions, we present the error estimate attainable with the Gauss–Laguerre approach for the approximation of $\lambda^{-\alpha}$. The analysis is then extended in Section 5 to the case of the operator $\mathscr{L}^{-\alpha}$. Finally, the truncated rules are proposed in Section 6.

2. The Gauss–Laguerre approach

As already said in the introduction, we start from the integral representation given in (1.3). Setting $y = \ln t$ we obtain

$$\mathscr{L}^{-\alpha} = \frac{2\sin(\alpha\pi)}{\pi}\int_{-\infty}^{+\infty} e^{2\alpha y}(\mathscr{I}+e^{2y}\mathscr{L})^{-1}\mathrm{d}y, \qquad \alpha \in (0,1). \tag{2.1}$$

Now we consider separately the two integrals

$$\int_{-\infty}^{0} e^{2\alpha y}(\mathscr{I}+e^{2y}\mathscr{L})^{-1}\mathrm{d}y, \quad \int_{0}^{+\infty} e^{2\alpha y}(\mathscr{I}+e^{2y}\mathscr{L})^{-1}\mathrm{d}y$$

and consider the changes of variable $2\alpha y = -x$ and $2(1-\alpha)y = x$, respectively, to obtain

$$\int_{-\infty}^{0} e^{2\alpha y}(\mathscr{I} + e^{2y}\mathscr{L})^{-1}\mathrm{d}y = \frac{1}{2\alpha}\int_0^{+\infty} e^{-x}(\mathscr{I} + e^{-x/\alpha}\mathscr{L})^{-1}\mathrm{d}x,$$

$$\int_0^{+\infty} e^{2\alpha y}(\mathscr{I} + e^{2y}\mathscr{L})^{-1}\mathrm{d}y = \frac{1}{2(1-\alpha)}\int_0^{+\infty} e^{-x}(e^{-x/(1-\alpha)}\mathscr{I} + \mathscr{L})^{-1}\mathrm{d}x.$$

Consequently, setting

$$I^{(1)}(\lambda) := \int_0^{+\infty} e^{-x}(1 + e^{-x/\alpha}\lambda)^{-1}\mathrm{d}x, \tag{2.2}$$

$$I^{(2)}(\lambda) := \int_0^{+\infty} e^{-x}(e^{-x/(1-\alpha)} + \lambda)^{-1}\mathrm{d}x, \tag{2.3}$$

the operator in (2.1) can be written as

$$\mathscr{L}^{-\alpha} = \frac{\sin(\alpha\pi)}{\alpha\pi} I^{(1)}(\mathscr{L}) + \frac{\sin(\alpha\pi)}{(1-\alpha)\pi} I^{(2)}(\mathscr{L}). \tag{2.4}$$

It is easy to check that $I^{(1)}(\mathscr{L}) \to \mathscr{I}$ as $\alpha \to 0$ and $I^{(2)}(\mathscr{L}) \to \mathscr{L}^{-1}$ as $\alpha \to 1$.

By applying the n-point Gauss–Laguerre rule to both integrals with respect to the weight function $\omega(x) = e^{-x}$, with weights $w_j^{(n)}$ and nodes $\vartheta_j^{(n)}$ (in ascending order), we obtain the following $(2n-1, 2n)$ rational approximation

$$\mathscr{L}^{-\alpha} \approx \frac{\sin(\alpha\pi)}{\alpha\pi} R^{(1)}_{n-1,n}(\mathscr{L}) + \frac{\sin(\alpha\pi)}{(1-\alpha)\pi} R^{(2)}_{n-1,n}(\mathscr{L}) =: R_{2n-1,2n}(\mathscr{L}), \tag{2.5}$$

where

$$R^{(1)}_{n-1,n}(\lambda) = \sum_{j=1}^{n} w_j^{(n)} \left(1 + e^{-\vartheta_j^{(n)}/\alpha}\lambda\right)^{-1},$$

$$R^{(2)}_{n-1,n}(\lambda) = \sum_{j=1}^{n} w_j^{(n)} \left(e^{-\vartheta_j^{(n)}/(1-\alpha)} + \lambda\right)^{-1}.$$

Clearly, formula (2.5) implies that using n points we have to perform $2n$ inversions.

3. Error analysis for a general function

In order to obtain an estimate of the error for the rational approximation defined in (2.5), we consider the approach introduced in Barrett (1961) and based on the theory of analytic functions. Assuming that we are working with a general function f and then considering the n-point Gauss–Laguerre rule $I_n(f)$ for

$$I(f) = \int_0^{+\infty} e^{-x} f(x)\,\mathrm{d}x,$$

we define the remainder as $E_n(f) = I(f) - I_n(f)$. For any given $R > 1$, the equation

$$\mathrm{Re}(\sqrt{-z}) = \ln R$$

represents a parabola in the complex plane, which we denote by Γ_R, symmetric with respect to the real axis, with vertex in $-(\ln R)^2$ and convexity oriented towards the positive real axis. By writing $z = a + \mathrm{i}b$, the above equation reads

$$a = \left(b^2 - 4(\ln R)^4\right)\frac{1}{4(\ln R)^2}.$$

The parabola degenerates to $[0, +\infty)$ as $R \to 1$. The theory given in Barrett (1961) states that if for a given R the function f is analytic on or within Γ_R except for a pair of simple poles, z_0 and its conjugate $\overline{z_0}$, then

$$E_n(f) \approx -4\pi \,\mathrm{Re}\left\{ r e^{-z_0} \left[\exp\left(\sqrt{-z_0}\right)\right]^{-2\sqrt{\bar{n}}} \right\}, \tag{3.1}$$

where r is the residue of $f(z)$ at z_0 and

$$\bar{n} = 4n + 2. \tag{3.2}$$

This result follows from the fact that $E_n(f)$ can be written as a contour integral

$$E_n(f) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{q_n(z)}{L_n(z)} f(z)\,\mathrm{d}z,$$

where $L_n(z)$ is the Laguerre polynomial, $q_n(z)$ is the so-called associated function defined by

$$q_n(z) = \int_0^{+\infty} \frac{e^{-x} L_n(x)}{z - x}\,\mathrm{d}x, \quad z \notin [0, +\infty),$$

and Γ is a contour containing $[0, +\infty)$ with the additional property that no singularity of $f(z)$ lies on or within this contour (for background see Davis & Rabinowitz, 1984, §4.6).

Denoting by C_1 and C_2 two arbitrary small circles surrounding the two poles the idea is then to define $\Gamma = \Gamma_R \cup C_1 \cup C_2$. In order to run this contour in the counterclockwise direction, one can artificially add three line segments as shown in Fig. 1 to connect the circles with the parabola. Then, following the black and the red arrows, the integrals along the line segments cancel and we obtain

$$E_n(f) = \frac{1}{2\pi \mathrm{i}} \left\{ \int_{\Gamma_R} - \int_{C_1} - \int_{C_2} \right\} \frac{q_n(z)}{L_n(z)} f(z)\,\mathrm{d}z. \tag{3.3}$$

At this point, the estimate is based on the relation given in Elliott (1967, Eq. (5.4)), namely

$$\frac{q_n(z)}{L_n(z)} = 2\pi e^{-z} \left[\exp\left(\sqrt{-z}\right)\right]^{-2\sqrt{\bar{n}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad z \notin [0, +\infty).$$

Since

$$\left[\exp\left(\mathrm{Re}\left(\sqrt{-z}\right)\right)\right]^{-2\sqrt{\bar{n}}} = R^{-2\sqrt{\bar{n}}} \quad \text{for } z \in \Gamma_R, \tag{3.4}$$

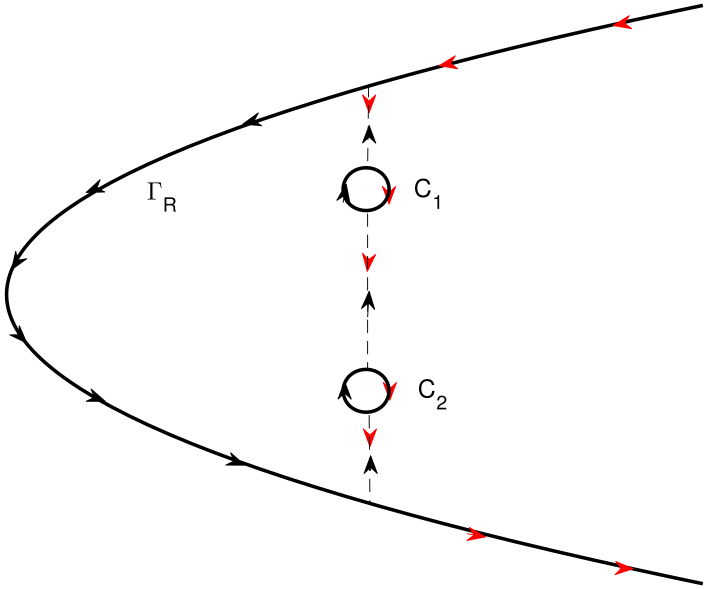

FIG. 1. Contour chosen for a function f analytic on or within the parabola Γ_R with the exception of two simple and conjugated poles located inside C_1 and C_2, respectively.

the contribution on the parabola is given by

$$\frac{1}{2\pi \mathrm{i}}\int_{\Gamma_R}\frac{q_n(z)}{L_n(z)}f(z)\,\mathrm{d}z = R^{-2\sqrt{\bar{n}}} \times \frac{1}{\mathrm{i}}\int_{\Gamma_R} e^{-z}\left[\exp\left(\mathrm{i}\,\mathrm{Im}\left(\sqrt{-z}\right)\right)\right]^{-2\sqrt{\bar{n}}} f(z)\,\mathrm{d}z\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right) := \phi(n). \tag{3.5}$$

In addition, using the residue theorem we have

$$\begin{aligned}
&\frac{1}{2\pi \mathrm{i}}\left\{\int_{C_1}+\int_{C_2}\right\} e^{-z}\left[\exp\left(\sqrt{-z}\right)\right]^{-2\sqrt{\bar{n}}} f(z)\,\mathrm{d}z \\
&= \mathrm{Res}\left(e^{-z}\left[\exp\left(\sqrt{-z}\right)\right]^{-2\sqrt{\bar{n}}} f(z), z_0\right) \\
&\quad + \mathrm{Res}\left(e^{-z}\left[\exp\left(\sqrt{-z}\right)\right]^{-2\sqrt{\bar{n}}} f(z), \overline{z_0}\right) \\
&= 2\,\mathrm{Re}\left(\mathrm{Res}\left(e^{-z}\left[\exp\left(\sqrt{-z}\right)\right]^{-2\sqrt{\bar{n}}} f(z), z_0\right)\right) \\
&= 2\,\mathrm{Re}\left(\mathrm{Res}\left(f(z), z_0\right) e^{-z_0}\left[\exp\left(\sqrt{-z_0}\right)\right]^{-2\sqrt{\bar{n}}}\right).
\end{aligned}$$

Therefore, from (3.3), by taking into account (3.5), we obtain

$$E_n(f) = -4\pi \,\mathrm{Re}\left(\mathrm{Res}\left(f(z), z_0\right) e^{-z_0} \left[\exp\left(\sqrt{-z_0}\right)\right]^{-2\sqrt{n}}\right)\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right)$$
$$+ \phi(n).$$

Obviously, this implies formula (3.1) whenever the contribution from the parabola Γ_R (i.e., $\phi(n)$) can be considered negligible. As for the modulus of the error, observing that

$$\left|\mathrm{Re}\left(\mathrm{Res}\left(f(z), z_0\right) e^{-z_0} \left[\exp\left(\sqrt{-z_0}\right)\right]^{-2\sqrt{n}}\right)\right| \leqslant \left|\mathrm{Res}\left(f(z), z_0\right) e^{-z_0}\right| \left[\exp\left(\mathrm{Re}\left(\sqrt{-z_0}\right)\right)\right]^{-2\sqrt{n}},$$

we have

$$|E_n(f)| \leqslant 4\pi \left|\mathrm{Res}\left(f(z), z_0\right) e^{-z_0}\right| \left[\exp\left(\mathrm{Re}\left(\sqrt{-z_0}\right)\right)\right]^{-2\sqrt{n}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right)$$
$$+ |\phi(n)|.$$

Since hereafter we assume that

$$\int_{\Gamma_R} |e^{-z} f(z)| \,\mathrm{d}z$$

is bounded, from (3.5) we obtain (see (3.2) and (3.4))

$$\frac{|\phi(n)|}{\left[\exp\left(\mathrm{Re}\left(\sqrt{-z_0}\right)\right)\right]^{-2\sqrt{n}}} \leqslant \frac{cR^{-2\sqrt{n}}}{\left[\exp\left(\mathrm{Re}\left(\sqrt{-z_0}\right)\right)\right]^{-2\sqrt{n}}} = \mathcal{O}\left(\exp(-n^{1/2})\right)$$

and then

$$|E_n(f)| \leqslant 4\pi \left|\mathrm{Res}\left(f(z), z_0\right) e^{-z_0}\right| \left[\exp\left(\mathrm{Re}\left(\sqrt{-z_0}\right)\right)\right]^{-2\sqrt{n}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \tag{3.6}$$

4. Error analysis for $\lambda^{-\alpha}$

From (2.4) and (2.5) and defining

$$\varepsilon_n^{(i)}(\lambda) = \left|I^{(i)}(\lambda) - R_{n-1,n}^{(i)}(\lambda)\right|, \quad i = 1, 2, \tag{4.1}$$

we can write

$$\left|\lambda^{-\alpha} - R_{2n-1,2n}(\lambda)\right| \leqslant \frac{\sin(\alpha\pi)}{\alpha\pi}\varepsilon_n^{(1)}(\lambda) + \frac{\sin(\alpha\pi)}{(1-\alpha)\pi}\varepsilon_n^{(2)}(\lambda). \tag{4.2}$$

Hence, using the results of the previous section we can develop the error analysis by working separately on the two integrals $I^{(i)}(\lambda)$, $i = 1, 2$.

4.1 *First integral* $I^{(1)}(\lambda)$

The function involved in (2.2) is

$$f(z) = (1 + e^{-z/\alpha}\lambda)^{-1}, \tag{4.3}$$

whose poles are given by

$$z_k = \alpha \ln \lambda + \mathrm{i}(2k+1)\alpha\pi, \quad k \in \mathbb{Z}.$$

They are equally spaced along the line $\mathrm{Re}(z) = \alpha \ln \lambda$, symmetric with respect to the real axis, and the closest to the real axis are $z_0 = \alpha \ln \lambda + \mathrm{i}\alpha\pi$ and $z_{-1} = \overline{z_0}$. It can immediately be verified that there exists $R > 1$ such that the corresponding parabola $\mathrm{Re}((-z)^{1/2}) = \ln R$ contains only the poles z_0 and $\overline{z_0}$ in its interior and that such an R satisfies

$$\frac{\alpha}{2}\left(\sqrt{(\ln \lambda)^2 + \pi^2} - \ln \lambda\right) < (\ln R)^2 < \frac{\alpha}{2}\left(\sqrt{(\ln \lambda)^2 + 9\pi^2} - \ln \lambda\right).$$

These bounds follow by imposing $z_0 \in \Gamma_R$ (the left-hand one) and $z_1 = \alpha \ln \lambda + \mathrm{i}3\alpha\pi \in \Gamma_R$ (the right-hand one).

In order to apply (3.6), first we observe that

$$\begin{aligned}(-z_0)^{1/2} &= [-(\alpha \ln \lambda + \mathrm{i}\alpha\pi)]^{1/2} \\ &= \sqrt{\frac{\alpha}{2}}\left(\gamma^-(\lambda) - \mathrm{i}\gamma^+(\lambda)\right),\end{aligned}$$

where

$$\gamma^{\pm}(\lambda) = \sqrt{\sqrt{(\ln \lambda)^2 + \pi^2} \pm \ln \lambda}. \tag{4.4}$$

Then, recalling that $z_0/\alpha = \ln \lambda + \mathrm{i}\pi$, we write

$$1 + e^{-z/\alpha}\lambda = 1 - e^{-(z-z_0)/\alpha} = \frac{z - z_0}{\alpha}\sum_{j=0}^{+\infty}\frac{(-1)^j (z-z_0)^j}{\alpha^j (j+1)!}.$$

In this case, the residue of the function given in (4.3) at the simple pole z_0 is given by

$$\mathrm{Res}\left(f(z), z_0\right) = \lim_{z \to z_0}\frac{z - z_0}{1 + e^{-z/\alpha}\lambda} = \alpha.$$

Therefore, from (3.6) we have

$$\varepsilon_n^{(1)}(\lambda) \leqslant 4\pi\alpha\lambda^{-\alpha}\exp\left(-\gamma^-(\lambda)\,(2\alpha\bar{n})^{1/2}\right)\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \tag{4.5}$$

4.2 *Second integral $I^{(2)}(\lambda)$*

The function to consider in this case is

$$f(z) = (e^{-z/(1-\alpha)} + \lambda)^{-1},$$

whose poles are given by

$$z_k = -(1-\alpha)\ln \lambda + \mathrm{i}(2k+1)(1-\alpha)\pi, \quad k \in \mathbb{Z}.$$

The only difference with respect to the integral $I^{(1)}(\lambda)$ is that the poles now have a negative real part. Anyway, as before we can easily find a parabola containing in its interior only the poles $z_0 = -(1-\alpha)\ln\lambda + \mathrm{i}(1-\alpha)\pi$ and its conjugate. We now have

$$(-z_0)^{1/2} = \sqrt{\frac{1-\alpha}{2}}\left(\gamma^+(\lambda) + \mathrm{i}\gamma^-(\lambda)\right),$$

where $\gamma^\pm(\lambda)$ are defined in (4.4). As for the residue at z_0 we easily find that $\operatorname{Res}\left(f(z), z_0\right) = (1-\alpha)/\lambda$. Using (3.6) again, we have

$$\varepsilon_n^{(2)}(\lambda) \leqslant 4\pi(1-\alpha)\lambda^{-\alpha}\exp\left(-\gamma^+(\lambda)\,(2(1-\alpha)\bar{n})^{1/2}\right)\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right). \tag{4.6}$$

Finally, plugging in (4.2) the bounds (4.5) and (4.6) we have the following result.

PROPOSITION 4.1 Let $\gamma^\pm(\lambda)$ be defined in (4.4) and $\bar{n} = 4n+2$. Denoting by

$$g_n^{(1)}(\lambda) := \lambda^{-\alpha}\exp(-\gamma^-(\lambda)\,(2\alpha\bar{n})^{1/2}), \tag{4.7}$$

$$g_n^{(2)}(\lambda) := \lambda^{-\alpha}\exp(-\gamma^+(\lambda)\,(2(1-\alpha)\bar{n})^{1/2}) \tag{4.8}$$

the λ-dependent factors of $\varepsilon_n^{(1)}(\lambda)$ and $\varepsilon_n^{(2)}(\lambda)$, respectively, then we have

$$\left|\lambda^{-\alpha} - R_{2n-1,2n}(\lambda)\right| \leqslant 4\sin(\alpha\pi)\left[g_n^{(1)}(\lambda) + g_n^{(2)}(\lambda)\right]\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right). \tag{4.9}$$

In order to verify the estimate provided in (4.9), in Fig. 2 we consider an example with $\lambda = 10$. Here and below, nodes and weights of the Gauss–Laguerre rule have been computed using the MATLAB function `GaussLaguerre.m` given in Van Damme (2020).

5. Error analysis for $\mathcal{L}^{-\alpha}$

For simplicity, from now on we assume that $\sigma(\mathcal{L}) \subseteq [1,+\infty)$. Since $\mathcal{L}$ is self-adjoint and positive, regarding the error we have

$$\left\|\mathcal{L}^{-\alpha} - R_{2n-1,2n}(\mathcal{L})\right\| \leqslant \max_{\lambda\geqslant 1}\left|\lambda^{-\alpha} - R_{2n-1,2n}(\lambda)\right|, \tag{5.1}$$

where $\|\cdot\|$ denotes the operator norm in $\mathcal{H}$. By (4.9) we must therefore study the functions $g_n^{(i)}(\lambda)$, $i = 1,2$, for $\lambda \geqslant 1$. In particular, this means we study the functions $\gamma^\pm(\lambda)$ (see (4.7) and (4.8)). By (4.4) it is immediate to see that $\gamma^-(\lambda) \to 0$ and $\gamma^+(\lambda) \to +\infty$ as $\lambda \to +\infty$. As a consequence, the function $g_n^{(1)}(\lambda)$ has exactly one maximum at a certain $\lambda_n > 1$, whereas $g_n^{(2)}(\lambda)$ is monotone decreasing, independently of α and n. At this point, in order to compute the right-hand side in (5.1) the first step consists in finding the point of maximum λ_n.

PROPOSITION 5.1 Let λ_n be the maximum of the function $g_n^{(1)}(\lambda)$. Then for n large enough,

$$\lambda_n = \widetilde{\lambda}_n(1+\mathcal{O}(n^{-1/3})),$$

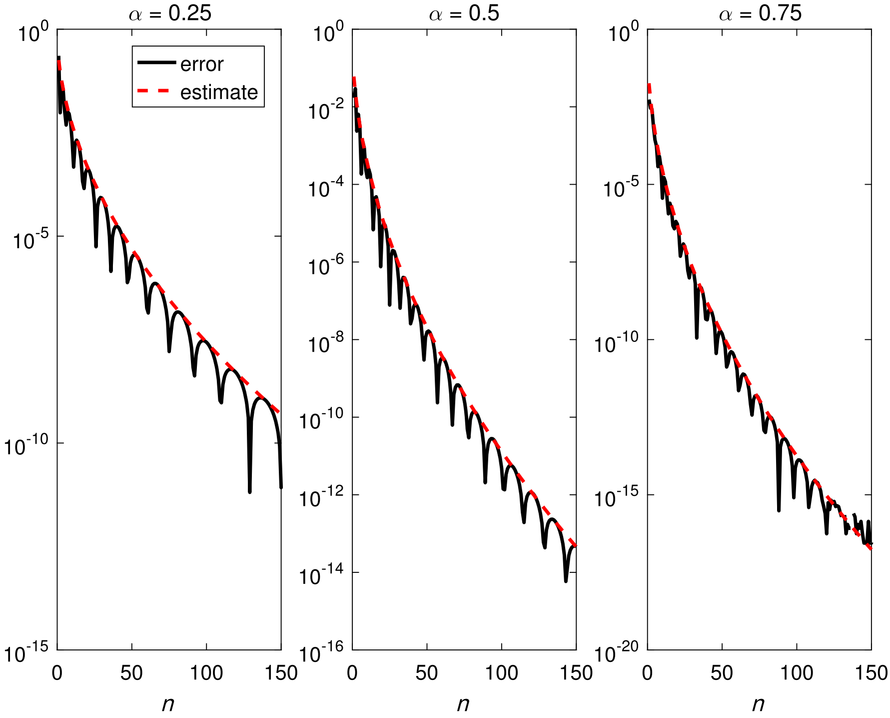

FIG. 2. Absolute error and its estimate given by (4.9) for $\lambda = 10$.

where

$$\widetilde{\lambda}_n = \exp\left(\left(\left(\frac{\bar{n}\pi^2}{4\alpha}\right)^{2/3} - \pi^2\right)^{1/2}\right).$$

Proof. By imposing $\frac{\mathrm{d}}{\mathrm{d}\lambda} g_n^{(1)}(\lambda) = 0$, after some manipulation we arrive at the equation

$$\frac{\sqrt{(\ln\lambda)^2 + \pi^2} - \ln\lambda}{(\ln\lambda)^2 + \pi^2} = \frac{2\alpha}{\bar{n}}, \tag{5.2}$$

whose solution is denoted by λ_n. Since

$$\begin{aligned} \frac{\sqrt{(\ln\lambda)^2 + \pi^2} - \ln\lambda}{(\ln\lambda)^2 + \pi^2} &= \frac{\pi^2}{\left((\ln\lambda)^2 + \pi^2\right)\left(\sqrt{(\ln\lambda)^2 + \pi^2} + \ln\lambda\right)} \\ &\geqslant \frac{\pi^2}{2\left((\ln\lambda)^2 + \pi^2\right)^{3/2}}, \end{aligned} \tag{5.3}$$

by (5.2) we first observe that there exists a constant c independent of n such that $\left(\ln\lambda_n\right)^3 \geqslant cn$, for n large enough. Writing

$$\ln\lambda = s\sqrt{(\ln\lambda)^2+\pi^2},$$

where

$$s = s(\lambda) = \frac{1}{\sqrt{1+\left(\frac{\pi}{\ln\lambda}\right)^2}}, \tag{5.4}$$

by (5.2) and (5.3) we obtain

$$\frac{\pi^2}{\left((\ln\lambda)^2+\pi^2\right)^{3/2}(1+s)} = \frac{2\alpha}{\overline{n}}.$$

As a consequence

$$\lambda_n = \exp\left(\left(\left(\frac{\overline{n}\pi^2}{2\alpha(1+s(\lambda_n))}\right)^{2/3} - \pi^2\right)^{1/2}\right).$$

Since asymptotically $\left(\ln\lambda_n\right)^2 \geqslant cn^{2/3}$, from (5.4) we have

$$s(\lambda_n) = 1 + \mathcal{O}(n^{-2/3})$$

and therefore

$$\lambda_n = \exp\left(\left(\left(\frac{\overline{n}\pi^2}{4\alpha}\right)^{2/3} - \pi^2 + \mathcal{O}(1)\right)^{1/2}\right). \tag{5.5}$$

Writing

$$\left(\left(\frac{\overline{n}\pi^2}{4\alpha}\right)^{2/3} - \pi^2 + \mathcal{O}(1)\right)^{1/2} = \left(\left(\frac{\overline{n}\pi^2}{4\alpha}\right)^{2/3} - \pi^2\right)^{1/2} + \sigma_n$$

we easily find that

$$\sigma_n = \mathcal{O}(n^{-1/3}).$$

Finally, we obtain the result since

$$\lambda_n = \exp\left(\left(\left(\frac{\overline{n}\pi^2}{4\alpha)}\right)^{2/3} - \pi^2\right)^{1/2}\right)\exp(\sigma_n).$$

□

This approximation is rather good, as can be observed in Fig. 3 where we plot $\ln\lambda_n$ and $\ln\widetilde{\lambda}_n$ for $n = 10, 11, \ldots, 120$. Here the value of λ_n that verifies (5.2) has been numerically computed by using a nonlinear solver.

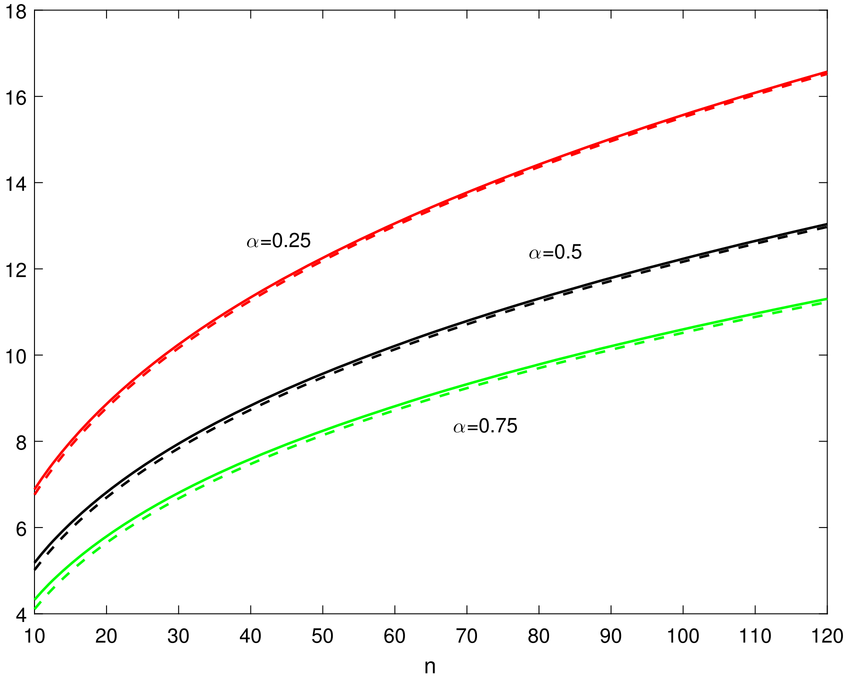

FIG. 3. Comparison between $\ln \lambda_n$ (solid lines) and $\ln \tilde{\lambda}_n$ (dashed lines) for $n = 10, 11, \ldots, 120$.

PROPOSITION 5.2 Let $g_n^{(1)}(\lambda)$ and $g_n^{(2)}(\lambda)$ be the functions defined in (4.7) and (4.8), respectively. Then

$$\max_{\lambda \geqslant 1} g_n^{(1)}(\lambda) = g_n^{(1)}(\lambda_n) = \exp\left(-3\left(n\alpha^2\pi^2\right)^{1/3}\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right), \tag{5.6}$$

$$\max_{\lambda \geqslant 1} g_n^{(2)}(\lambda) = g_n^{(2)}(1) = \exp\left(-\left(8\pi(1-\alpha)n\right)^{1/2}\right)\left(1 + \mathcal{O}\left(n^{-1/2}\right)\right). \tag{5.7}$$

Proof. First of all we need to evaluate $\gamma^-\left(\lambda_n\right)(2\alpha\overline{n})^{1/2}$. Using (4.4) and (5.2) we have

$$\left(\gamma^-\left(\lambda_n\right)\right)^2 = \frac{2\alpha}{\overline{n}}\left(\left(\ln \lambda_n\right)^2 + \pi^2\right).$$

By (5.5) we also have

$$\left(\ln \lambda_n\right)^2 + \pi^2 = \left(\frac{\overline{n}\pi^2}{4\alpha}\right)^{2/3} + \mathcal{O}(1) \tag{5.8}$$

and hence

$$\gamma^{-}(\lambda_n) = \left(\left(\frac{\alpha}{2\overline{n}}\right)^{1/3}\pi^{4/3} + \mathcal{O}(n^{-1})\right)^{1/2}$$
$$= \left(\frac{\alpha}{2\overline{n}}\right)^{1/6}\pi^{2/3}\left(1 + \mathcal{O}(n^{-2/3})\right).$$

Consequently,

$$\gamma^{-}(\lambda_n)(2\alpha\overline{n})^{1/2} = (2\overline{n}\alpha^2\pi^2)^{1/3}\left(1 + \mathcal{O}(n^{-2/3})\right).$$

Using the result obtained in Proposition 5.1 we can write

$$\lambda_n^{-\alpha} = \exp\left(-\alpha\left(\left(\left(\frac{\overline{n}\pi^2}{4\alpha}\right)^{2/3} - \pi^2\right)^{1/2}\right)\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right)$$
$$= \exp\left(-\left(\frac{\overline{n}\alpha^2\pi^2}{4}\right)^{1/3}\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right).$$

Therefore, we have

$$g_n^{(1)}(\lambda_n) = \exp\left(-\left(\frac{\overline{n}\alpha^2\pi^2}{4}\right)^{1/3}\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right)$$
$$\times \exp\left(-(2\overline{n}\alpha^2\pi^2)^{1/3}\left(1 + \mathcal{O}(n^{-2/3})\right)\right)$$
$$= \exp\left(-\left(\overline{n}\alpha^2\pi^2\right)^{1/3}\left(4^{-1/3} + 2^{1/3}\right)\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right).$$

Finally, recalling that $\overline{n} = 4n + 2$ we obtain the result.

As for the function $g_n^{(2)}(\lambda)$, the situation is much simpler. Indeed, since it is monotone decreasing, using (4.4) and (4.8) we have

$$\max_{\lambda \geqslant 1} g_n^{(2)}(\lambda) = g_n^{(2)}(1) = \exp\left(-(2\pi(1-\alpha)\overline{n})^{1/2}\right)$$
$$= \exp\left(-(8\pi(1-\alpha)n)^{1/2}\right)\left(1 + \mathcal{O}\left(n^{-1/2}\right)\right).$$

□

Finally, we can prove the following result.

PROPOSITION 5.3 Let $R_{2n-1,2n}(\mathscr{L})$ be the rational approximation given in (2.5). Then, with respect to the operator norm in $\mathscr{H}$, we have for n large enough

$$\left\|\mathscr{L}^{-\alpha} - R_{2n-1,2n}(\mathscr{L})\right\| \leqslant 4\sin(\alpha\pi)\exp\left(-3(n\alpha^2\pi^2)^{1/3}\right)\left(1+\mathcal{O}(n^{-1/3})\right). \tag{5.9}$$

Proof. First of all, by comparing (5.6) with (5.7) for n large enough we can write

$$\frac{g_n^{(2)}(1)}{g_n^{(1)}(\lambda_n)} \leqslant \frac{1}{n}.$$

Therefore,

$$\begin{aligned}
\max_{\lambda\geqslant 1}\left(g_n^{(1)}(\lambda)+g_n^{(2)}(\lambda)\right) &\leqslant \max_{\lambda\geqslant 1} g_n^{(1)}(\lambda) + \max_{\lambda\geqslant 1} g_n^{(2)}(\lambda) \\
&\leqslant g_n^{(1)}(\lambda_n) + g_n^{(2)}(1) \\
&= g_n^{(1)}(\lambda_n)\left(1+\mathcal{O}\left(n^{-1}\right)\right).
\end{aligned}$$

By Propostion 5.2 we find the result. □

To test the estimate just given in Proposition 5.3 we work with the operator

$$\mathscr{L} = \left[\mathrm{diag}(1,2,\ldots,100)\right]^8 \tag{5.10}$$

so that $\sigma(\mathscr{L}) \subseteq \left[1, 10^{16}\right]$. In Fig. 4 we plot the error and its estimate (5.9) with respect to the number of inversions, that is, $2n$. From now on, for discrete operators the error is plotted with respect to the Euclidean matrix norm.

Notwithstanding the above result, experimentally (see Fig. 5) it can immediately be observed that

$$\max_{\lambda\geqslant 1}\left(g_n^{(1)}(\lambda)+g_n^{(2)}(\lambda)\right) \approx \max\left(g_n^{(1)}(\lambda_n), g_n^{(2)}(1)\right).$$

This is because the contribution of a function in correspondence of the maximum of the other one is negligible. In order to understand whenever $g_n^{(2)}(1)$ may be greater than $g_n^{(1)}(\lambda_n)$ for some values of n and α (as in Fig. 5 for $\alpha = 0.75$) we just need to compare $g_n^{(2)}(1)$ with $g_n^{(1)}(1)$.

Using (4.4), (4.7) and (4.8) the equation $g_n^{(1)}(1) = g_n^{(2)}(1)$ is approximatively equivalent to

$$\exp\left(-\,(2\pi\alpha\overline{n})^{1/2}\right) = \exp\left(-\,(2\pi(1-\alpha)\overline{n})^{1/2}\right),$$

whose solution is $\alpha = 1/2$ independently of n. This means that for $\alpha \leqslant 1/2$,

$$\max_{\lambda\geqslant 1}\left(g_n^{(1)}(\lambda)+g_n^{(2)}(\lambda)\right) \approx g_n^{(1)}(\lambda_n)$$

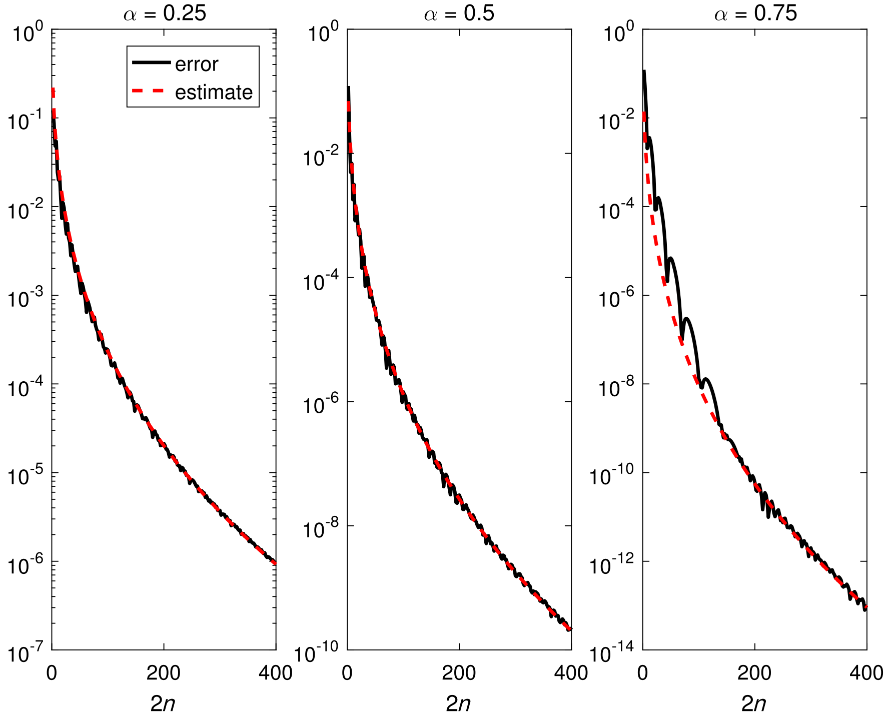

FIG. 4. Error and its estimate given by (5.9) for the operator defined in (5.10).

and therefore the error decays like $\exp\left(-cn^{1/3}\right)$ for some absolute constant c (cf. (5.6)), whereas for $\alpha > 1/2$ the situation is a bit more complicated. By comparing (5.6) with (5.7) we have that asymptotically $g_n^{(2)}(1)$ decay faster than $g_n^{(1)}(\lambda_n)$, so after a certain n^* the decay rate is still of type $\exp\left(-cn^{1/3}\right)$ also for $\alpha > 1/2$. Anyway, for $n \leqslant n^*$ the decay rate is of type $\exp\left(-cn^{1/2}\right)$. The integer n^* comes from the solution with respect to n of

$$g_n^{(1)}(\lambda_n) = g_n^{(2)}(1).$$

Using Proposition 5.2 we can estimate it by solving

$$\exp\left(-3\left(n\alpha^2\pi^2\right)^{1/3}\right) = \exp\left(-\left(8\pi(1-\alpha)n\right)^{1/2}\right).$$

We easily find

$$n^* \approx 4.5\frac{\alpha^4}{(1-\alpha)^3}. \tag{5.11}$$

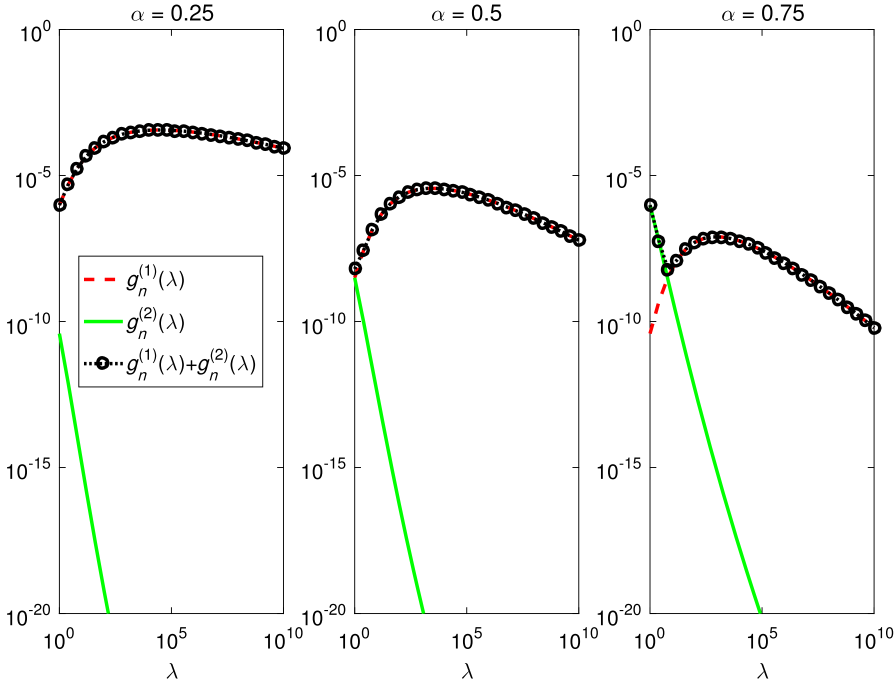

FIG. 5. Behavior of the functions $g_n^{(1)}(\lambda), g_n^{(2)}(\lambda), g_n^{(1)}(\lambda) + g_n^{(2)}(\lambda)$ for $n = 30$.

The previous considerations can be summarized as follows:

$$\left\| \mathscr{L}^{-\alpha} - R_{2n-1,2n}(\mathscr{L}) \right\| \approx 4\sin(\alpha\pi)S(n,\alpha), \tag{5.12}$$

where

$$S(n,\alpha) = \begin{cases} g_n^{(1)}(\lambda_n) & (\forall n \wedge \alpha \leqslant 1/2) \vee (n > n^* \wedge \alpha > 1/2), \\ g_n^{(2)}(1) & (n \leqslant n^* \wedge \alpha > 1/2) \end{cases} \tag{5.13}$$

(see (5.6) and (5.7)).

6. Truncated approaches

The idea of truncating the Gauss–Laguerre rule is clearly not new and is essentially a consequence of the fact that the weights decay exponentially. Among the existing papers on this point we recall Berger (1969), where a truncated approach has been used for the computation of the Laplace transform, and Mastroianni & Monegato (2004), where the authors develop the error analysis of the truncated Gauss–Laguerre rule for a general, absolutely continuous f.

Here we focus on the case where f is an arbitrary continuous function that satisfies $0 \leqslant f(x) \leqslant 1$, since this is the case for the functions that appear in the definition of $I^{(i)}(\lambda), i = 1, 2$. In fact, we clearly have that for $\lambda \geqslant 1$ (see (2.2) and (2.3)),

$$0 \leqslant (1 + e^{-x/\alpha}\lambda)^{-1} \leqslant 1, \qquad 0 \leqslant (e^{-x/(1-\alpha)} + \lambda)^{-1} \leqslant 1.$$

Suppose that a sequence of error approximations $\{\varepsilon_n\}_{n\geqslant 1}$ is available, that is,

$$\left|I(f) - I_n(f)\right| \leqslant \varepsilon_n, \tag{6.1}$$

where now $I_n(f)$ is the n-point Gauss–Laguerre approximation of $I(f)$, with $0 \leqslant f(x) \leqslant 1$. Since

$$\int_0^{+\infty} e^{-x} f(x)\,\mathrm{d}x \leqslant \int_0^{+\infty} e^{-x}\,\mathrm{d}x,$$

let s_n be the solution of

$$\int_{s_n}^{+\infty} e^{-x}\,\mathrm{d}x = \varepsilon_n,$$

that is,

$$s_n = -\ln \varepsilon_n. \tag{6.2}$$

We consider the truncated rule

$$\begin{aligned} I_{k_n}(f) &= \sum_{j=1}^{k_n} w_j^{(n)} f(\vartheta_j^{(n)}) \\ &= I_n(f) - \sum_{j=k_n+1}^{n} w_j^{(n)} f(\vartheta_j^{(n)}), \end{aligned}$$

where $k_n \leqslant n$ is the smallest integer such that $\vartheta_j^{(n)} \geqslant s_n$ for $j \geqslant k_n$. Therefore,

$$\begin{aligned} \left|I(f) - I_{k_n}(f)\right| &= \left|I(f) - I_n(f) + \sum_{j=k_n+1}^{n} w_j^{(n)} f(\vartheta_j^{(n)})\right| \\ &\leqslant \left|I(f) - I_n(f)\right| + \sum_{j=k_n+1}^{n} w_j^{(n)}. \end{aligned}$$

Using the bound (Mastroianni & Occorsio, 2001, Eqs. (2.4) and (2.7))

$$w_j^{(n)} \leqslant C(\vartheta_j^{(n)} - \vartheta_{j-1}^{(n)}) e^{-\vartheta_j^{(n)}}, \quad j = 2, \ldots, n,$$

where C is a constant independent of n, we have (see (6.2))

$$\sum_{j=k_n+1}^{n} w_j^{(n)} \leqslant C \sum_{j=k_n+1}^{n} (\vartheta_j^{(n)} - \vartheta_{j-1}^{(n)}) e^{-\vartheta_j^{(n)}} \leqslant C \int_{\vartheta_{k_n}^{(n)}}^{+\infty} e^{-y}\, \mathrm{d}y = C e^{-\vartheta_{k_n}^{(n)}} \leqslant C e^{-s_n} = C\varepsilon_n,$$

so that finally

$$\left| I(f) - I_{k_n}(f) \right| \leqslant (1 + C)\varepsilon_n.$$

REMARK 6.1 Experimentally, one can easily check that the approximation

$$w_j^{(n)} \approx (\vartheta_j^{(n)} - \vartheta_{j-1}^{(n)}) e^{-\vartheta_j^{(n)}}$$

is very accurate and hence in the numerical experiments we take $C = 1$.

6.1 *A balanced approach*

Let $k_n \leqslant n$ be the smallest integer such that (see (5.9))

$$\vartheta_j^{(n)} \geqslant -\ln\left(4\sin(\alpha\pi)\exp\left(-3\left(n\alpha^2\pi^2\right)^{1/3}\right)\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right), \quad j \geqslant k_n.$$

Using the above theory we have that for n large enough,

$$\begin{aligned}\left\| \mathscr{L}^{-\alpha} - R_{2k_n-1,2k_n}(\mathscr{L}) \right\| \leqslant 4(1 + C)\sin(\alpha\pi)\exp\left(-3\left(n\alpha^2\pi^2\right)^{1/3}\right) \\ \times \left(1 + \mathcal{O}\left(n^{-1/3}\right)\right).\end{aligned} \tag{6.3}$$

In order to derive error estimates with respect to k_n, that is, with respect to the number of inversions, we first need to prove the following result.

PROPOSITION 6.1 For k large enough, the kth root of the Laguerre polynomial of degree n satisfies

$$\vartheta_k^{(n)} = c_k \frac{k^2\pi^2}{4n}(1 + \mathcal{O}(n^{-2})), \quad 1 < c_k \leqslant \left(1 + \frac{1}{k}\right)^2. \tag{6.4}$$

Proof. First of all we need to study the asymptotic behavior of the roots of $J_0(z)$, the Bessel function of the first kind of order 0. By Szegö (1939, Eq. (1.71.7)),

$$J_0(z) = \left(\frac{2}{\pi z}\right)^{1/2} \cos\left(z - \frac{\pi}{4}\right) + \mathcal{O}(z^{-3/2}),$$

we observe that there is a root, say j_k, in $I = [\pi/2 + k\pi, (k+1)\pi]$ since $J_0(z)$ changes sign. Now let

$$z_k = \frac{3}{4}\pi + k\pi \in I \tag{6.5}$$

be the solution of $\cos\left(z - \frac{\pi}{4}\right) = 0$, so that $J_0(z_k) = \mathcal{O}(k^{-3/2})$. Therefore,

$$z_k - j_k = \frac{J_0(z_k)}{J_0'(\xi)}, \quad \xi \in I.$$

Now, since

$$J_0'(\xi) = \left(\frac{2}{\pi}\right)^{1/2}\left[-\frac{1}{2}\xi^{-3/2}\cos\left(\xi - \frac{\pi}{4}\right) - \sin\left(\xi - \frac{\pi}{4}\right)\xi^{-1/2}\right] + \mathcal{O}(\xi^{-5/2})$$

and

$$\left|\sin\left(\xi - \frac{\pi}{4}\right)\right| \geqslant \frac{\sqrt{2}}{2}$$

we deduce that $J_0'(\xi) = \mathcal{O}(k^{-1/2})$. From the above considerations we have

$$z_k - j_k = \mathcal{O}(k^{-1})$$

and then using (6.5) we get

$$j_k^2 = \left(\frac{3}{4}\pi + k\pi + \mathcal{O}(k^{-1})\right)^2 = (k\pi)^2\left(1 + \frac{3}{4k} + \mathcal{O}(k^{-2})\right)^2.$$

By Abramowitz & Stegun (1970, Eq. (22.16.8)),

$$\vartheta_k^{(n)} = \frac{j_k^2}{4n+2}\left[1 + \frac{j_k^2}{4\,(4n+2)^2}\right] + \mathcal{O}\left(n^{-5}\right)$$

we obtain the result. □

Now we want to solve with respect to k,

$$\vartheta_k^{(n)} = -\ln\left(4\sin(\alpha\pi)\exp\left(-3\left(n\alpha^2\pi^2\right)^{1/3}\right)\right)\left(1 + \mathcal{O}\left(n^{-1/3}\right)\right). \tag{6.6}$$

For k large enough, by (6.4) the solution of (6.6) satisfies

$$c_k\frac{k^2\pi^2}{4n}(1 + \mathcal{O}(n^{-2})) = -\ln\left(4\sin(\alpha\pi)\right) + 3\left(n\alpha^2\pi^2\right)^{1/3} + \mathcal{O}(n^{-1/3}), \tag{6.7}$$

that is,

$$c_k \frac{k^2\pi^2}{4n} = 3\left(n\alpha^2\pi^2\right)^{1/3}(1+\mathcal{O}(n^{-1/3})).$$

By the definition of c_k we thus have $k \sim n^{2/3}$ and therefore

$$k^2\left(1+\mathcal{O}(k^{-1/2})\right) = 12\alpha^{2/3}\pi^{-4/3}n^{4/3}$$

which leads to

$$n^{1/3} = \frac{k^{1/2}}{12^{1/4}\alpha^{1/6}\pi^{-1/3}}\left(1+\mathcal{O}(k^{-1/2})\right).$$

Using this value in (6.3) we find

$$\begin{aligned}\left\|\mathscr{L}^{-\alpha} - R_{2k-1,2k}(\mathscr{L})\right\| &\leqslant 4(1+C)\sin(\alpha\pi) \\ &\quad\times \exp\left(-3\frac{\pi}{12^{1/4}}\alpha^{1/2}k^{1/2}\left(1+\mathcal{O}(k^{-1/2})\right)\right)\left(1+\mathcal{O}(k^{-1/2})\right) \\ &\leqslant 4(1+C)\hat{C}\sin(\alpha\pi)\exp\left(-3\frac{\pi}{12^{1/4}}\alpha^{1/2}k^{1/2}\right),\end{aligned}$$

where the constant $\hat{C}$ takes account of the term $\left(1+\mathcal{O}(k^{-1/2})\right)$.

We remark, however, that the above analysis can be simplified by neglecting the terms $\ln(4\sin(\alpha\pi))$ and c_k in (6.7), and solving directly

$$\frac{k^2\pi^2}{4n} = 3\left(n\alpha^2\pi^2\right)^{1/3}.$$

Using the floor function, we denote by

$$k_n^{(1)} = \left\lfloor 2\sqrt{3}\left(\frac{\alpha n^2}{\pi^2}\right)^{1/3}\right\rfloor, \tag{6.8}$$

which experimentally is confirmed to be a value rather close to k_n, in a reasonable range of values of α, say $\alpha \in [0.05, 0.95]$, leading to a method that is almost indistinguishable from the one with k_n. Since

$$n \approx \frac{\pi}{\alpha^{1/2}}\left(\frac{k_n^{(1)}}{2\sqrt{3}}\right)^{3/2} \tag{6.9}$$

using (6.3) we find

$$\left\|\mathscr{L}^{-\alpha} - R_{2k_n^{(1)}-1,2k_n^{(1)}}(\mathscr{L})\right\| \approx 4(1+C)\sin(\alpha\pi)\exp\left(-3.6\alpha^{1/2}\left(2k_n^{(1)}\right)^{1/2}\right). \tag{6.10}$$

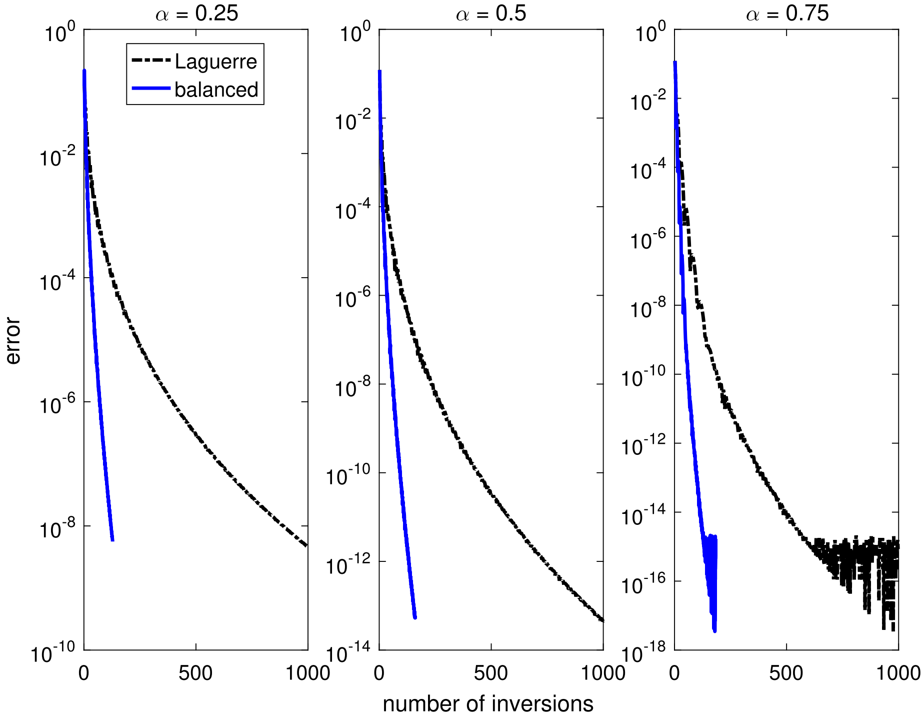

FIG. 6. $\left\|\mathscr{L}^{-\alpha} - R_{2j-1,2j}(\mathscr{L})\right\|$ vs the number of inversions $2j$, for $j = n$ (Laguerre) and $j = k_n^{(1)}$ (balanced).

By using the operator (5.10) again, in Fig. 6 we compare the two errors provided by applying the n-point Gauss–Laguerre rule and the corresponding balanced formula, that is,

$$\left\|\mathscr{L}^{-\alpha} - R_{2j-1,2j}(\mathscr{L})\right\|, \qquad j = n, k_n^{(1)}.$$

We observe great improvement in terms of the computational cost attainable with the truncated approach.

In Fig. 7 we focus attention on the truncated (balanced) approach. We plot the error and its estimate (6.10) with $C = 1$ with respect to the number of inversions, that is, $2k_n^{(1)}$. The results show the accuracy of the estimate.

When $\alpha > 1/2$ the above estimate may be optimistic for $n \leqslant n^*$ (cf. (5.11)). Working with (5.12)–(5.13) with $S(n,\alpha) = g_n^{(2)}(1)$ and following the same analysis that starts from (6.6), by (5.7) we find that $k = 2(1-\alpha)^{1/4}\,(2n/\pi)^{3/4}$ and then the value

$$k_n^{(2)} := 2\left\lfloor (1-\alpha)^{1/4}\left(\frac{2n}{\pi}\right)^{3/4}\right\rfloor \tag{6.11}$$

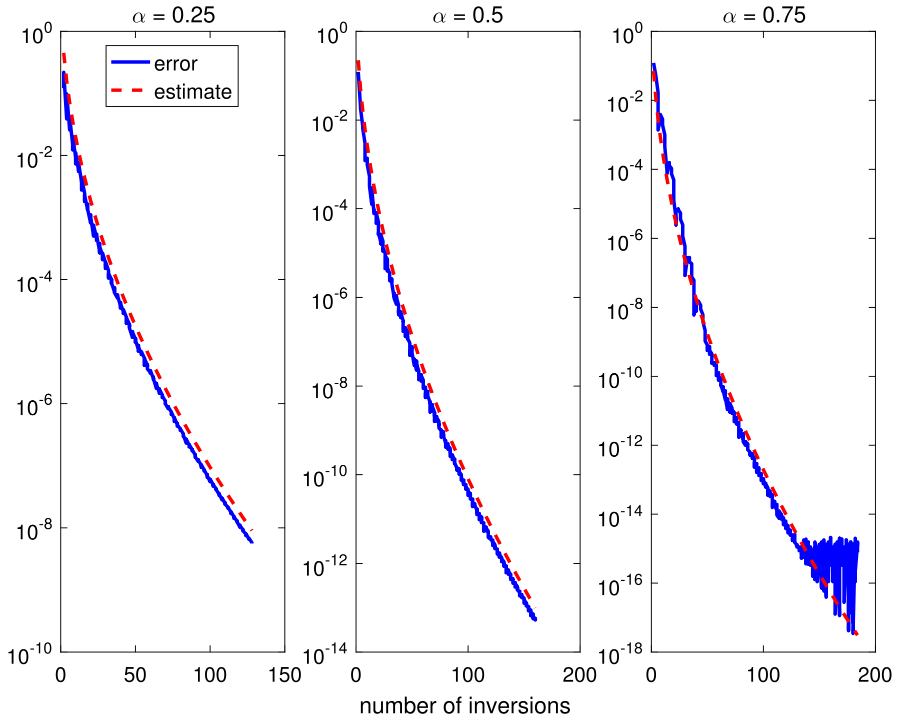

FIG. 7. Error and its estimate given by (6.10) for the operator defined in (5.10).

is very close to k_n. Therefore,

$$\left\| \mathscr{L}^{-\alpha} - R_{2k_n^{(2)}-1,2k_n^{(2)}}(\mathscr{L}) \right\| \approx 4(1+C)\sin(\alpha\pi) \times \exp\left(-2.96(1-\alpha)^{1/3}\left(2k_n^{(2)}\right)^{2/3}\right) \qquad \text{for } (n \leqslant n^*) \wedge (\alpha > 1/2), \tag{6.12}$$

which expresses an initial convergence very fast with respect to the number of inversions. For $\alpha > 1/2$, one should use the first $k_n^{(2)}$ Laguerre points for $n \leqslant n^*$ and then switch to the first $k_n^{(1)}$ for $n > n^*$. Anyway, experimentally it can be observed that the corresponding method does not offer valuable improvement with respect to the choice of the first $k_n^{(1)}$, independently of α and n.

Therefore, the balanced approach that we propose is the one based on (6.8), and reported in the figures, with error estimate given by (6.10) independently of α and n.

6.2 *An equalized approach*

The idea is to work separately on the two integrals and hence to consider approximations of the type

$$\mathscr{L}^{-\alpha} \approx \frac{\sin(\alpha\pi)}{\alpha\pi} R^{(1)}_{k_{n_1}-1,k_{n_1}}(\mathscr{L}) + \frac{\sin(\alpha\pi)}{(1-\alpha)\pi} R^{(2)}_{k_{n_2}-1,k_{n_2}}(\mathscr{L}),$$

in which $R^{(i)}_{k_{n_i}-1,k_{n_i}}(\lambda)$, $i = 1,2$, represents the truncated Gauss–Laguerre rule for $I^{(i)}(\lambda)$ based on the first k_{n_i} roots of the Laguerre polynomials of degree n_i. For $n_1 \neq n_2$ we then use different sets of points, and clearly the total number of inversions is now $k_{n_1} + k_{n_2}$.

We first consider the case where, for a given n, $\varepsilon_n^{(1)}(\lambda)/\alpha \geqslant \varepsilon_n^{(2)}(\lambda)/(1-\alpha)$ (cf. (4.1) and (4.2)) and we define $n_1 = n$. Then we evaluate $k_{n_1} = k^{(1)}_{n_1}$ as in (6.8) and we approximate $I^{(1)}(\mathscr{L})$ with $R^{(1)}_{k_{n_1}-1,k_{n_1}}(\mathscr{L})$. Then we find n_2 ($\leqslant n_1$) such that

$$g^{(1)}_{n_1}(\lambda_{n_1}) = g^{(2)}_{n_2}(1),$$

that is,

$$\exp\left(-3\left(n_1\alpha^2\pi^2\right)^{1/3}\right) = \exp\left(-\left(8\pi(1-\alpha)n_2\right)^{1/2}\right) \tag{6.13}$$

(cf. (5.6) and (5.7)). At this point we compute as in (6.11),

$$k_{n_2} = k^{(2)}_{n_2} = 2\left\lfloor (1-\alpha)^{1/4}\left(\frac{2n_2}{\pi}\right)^{3/4}\right\rfloor, \tag{6.14}$$

and use the Gauss–Laguerre rule $R^{(2)}_{k_{n_2}-1,k_{n_2}}(\mathscr{L})$ for the second integral. Clearly, for each n the error estimate for the equalized approach remains that of the balanced approach given by (6.10), but now we have fewer inversions. In this view, we have to find the relationship between k_{n_1} and k_{n_2}. From (6.13) we get

$$n_2 = \frac{9}{8}\pi^{1/3}\frac{\alpha^{4/3}}{1-\alpha}n_1^{2/3}$$

so that using (6.14) we can express k_{n_2} in terms of n_1. Then by (6.9) we obtain

$$k_{n_2} \approx \frac{3.09}{\alpha^{3/4}(1-\alpha)^{1/2}}k_{n_1}^{3/4},$$

from which we deduce that $(k_{n_1} + k_{n_2}) \leqslant 2k_{n_1}$.

As for the case $\varepsilon_n^{(1)}(\lambda)/\alpha < \varepsilon_n^{(2)}(\lambda)/(1-\alpha)$ the arguments follow the same line. Let $n_2 = n$ and compute the second integral with $R^{(2)}_{k_{n_2}-1,k_{n_2}}(\mathscr{L})$. Then solving (6.13) with respect to n_1 ($\leqslant n_2$) we obtain

$$n_1 = \frac{(8(1-\alpha))^{3/2}}{27\alpha^2\pi^{1/2}}n_2^{3/2}.$$

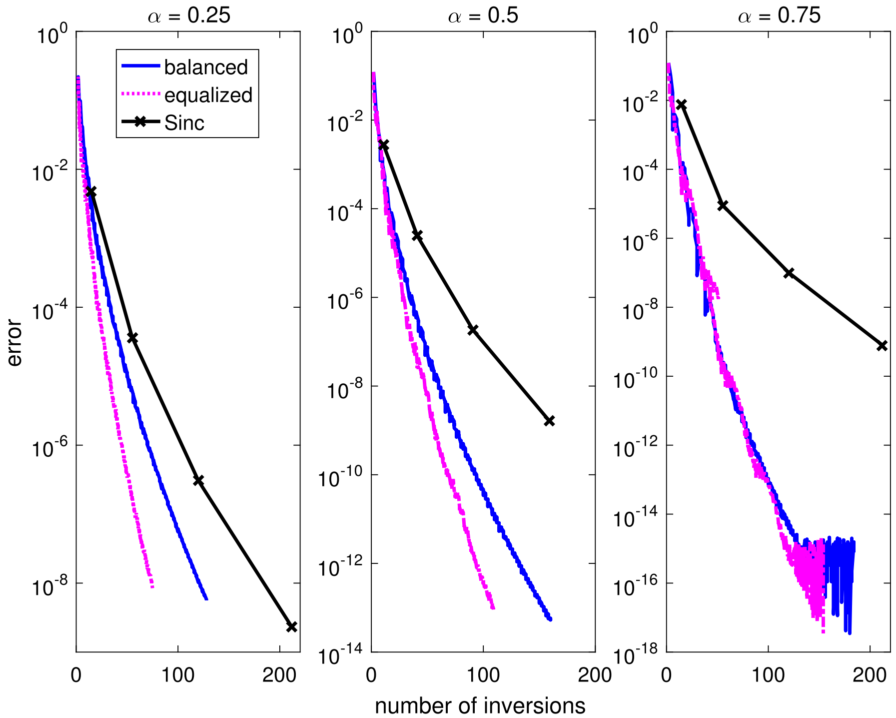

FIG. 8. Comparison between the errors provided by the balanced and equalized approaches with the Sinc quadrature studied in Bonito & Pasciak (2015).

Consequently, as in (6.8),

$$k_{n_1} = k_{n_1}^{(1)} = \left\lfloor 2\sqrt{3}\left(\frac{\alpha n_1^2}{\pi^2}\right)^{1/3} \right\rfloor,$$

and we compute the first integral with $R^{(1)}_{k_{n_1}-1,k_{n_1}}(\mathscr{L})$. Using (6.11) we also have

$$k_{n_2} = 2\left\lfloor (1-\alpha)^{1/4}\left(\frac{2n_2}{\pi}\right)^{3/4} \right\rfloor$$

and therefore, collecting the above expressions, we finally obtain

$$k_{n_1} \approx 0.61\frac{(1-\alpha)^{2/3}}{\alpha}k_{n_2}^{4/3}.$$

As before, the error estimate for the equalized approach is that of the balanced approach given by (6.12) but the number of inversions that we have to consider is now $(k_{n_1}+k_{n_2}) \leqslant 2k_{n_2}$.

In Fig. 8 we consider the comparison between our two truncated approaches together with Sinc rule analyzed in Bonito & Pasciak (2015).

7. Conclusions

In this work we have considered the construction of very fast methods based on the Gauss–Laguerre rule and we have been able to provide accurate error estimates that can be used to *a priori* select the number of points to use. We observe that while all the experiments concern the artificial example (5.10), other tests on finite difference discretizations of the Laplace operator have essentially led to identical results.

Funding

GNCS-INdAM; University of Pisa (PRA_2020_61); FRA-University of Trieste.

Acknowledgements

The authors are members of the INdAM research group GNCS.

References

Abramowitz, M. & Stegun, I. (1970) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Publications.

Aceto, L., Bertaccini, D., Durastante, F. & Novati, P. (2019) Rational Krylov methods for functions of matrices with applications to fractional partial differential equations. *J. Comput. Phys.*, **396**, 470–482.

Aceto, L. & Novati, P. (2019) Rational approximations to fractional powers of self-adjoint positive operators. *Numer. Math.*, **143**, 1–16.

Aceto, L. & Novati, P. (2020) Padé-type approximations to the resolvent of fractional powers of operators. *J. Sci. Comput.*, **83**, Article number: 13.

Barrett, W. (1961) Convergence properties of Gaussian quadrature formulae. *Comput. J.*, **3**, 272–277.

Berger, B. S. (1969) Dynamic response of an infinite cylindrical shell in acoustic medium. *J. Appl. Mech.*, **36**, 342–345.

Bonito, A. & Pasciak, J. E. (2015) Numerical approximation of fractional powers of elliptic operators. *Math. Comp.*, **84**, 2083–2110.

Davis, P. J. & Rabinowitz, P. (1984) *Methods of Numerical Integration*. San Diego: Academic Press.

Elliott, D. (1967) Truncation errors in Padé approximations to certain functions: an alternative approach. *Math. Comp.*, **21**, 398–406.

Harizanov, S., Lazarov, R., Margenov, S., Marinov, P. & Vutov, Y. (2018) Optimal solvers for linear systems with fractional powers of sparse SPD matrices. *Numer. Linear Algebra Appl.*, **25**, e2167.

Harizanov, S., Lazarov, R., Margenov, S., Marinov, P. & Pasciak, J. E. (2020) Analysis of numerical methods for spectral fractional elliptic equations based on the best uniform rational approximation. *J. Comput. Phys.*, **408**, 109285.

Harizanov, S., Lazarov, R, Marinov, P., Margenov, S. & Pasciak, J. E. (2019) Comparison analysis of two numerical methods for fractional diffusion problems based on the best rational approximations of t^{γ} on $[0, 1]$. *Advanced Finite Element Methods with Applications (FEM 2017)* (T. Apel, U. Langer, A. Meyer & O. Steinbach eds), Lecture Notes in Computational Science and Engineering, vol. 128. Cham: Springer.

Harizanov, S. & Margenov, S. (2018) Positive approximations of the inverse of fractional powers of SPD M-matrices. *Control Systems and Mathematical Methods in Economics* (G. Feichtinger, R. Kovacevic & G. Tragler eds). Lecture Notes in Economics and Mathematical Systems, vol. 687. Cham: Springer.

Hofreither, C. (2020) A unified view of some numerical methods for fractional diffusion. *Comput. Math. Appl.*, **80**, 332–350.

Mastroianni, G. & Monegato, G. (2004) Truncated quadrature rules over $(0, \infty)$ and Nyström-type methods. *SIAM J. Numer. Anal.*, **41**, 1870–1892.

Mastroianni, G. & Occorsio, D. (2001) Lagrange interpolation at Laguerre zeros in some weighted uniform spaces. *Acta Math. Hung.*, **91**, 27–52.

Stahl, H. R. (2003) Best uniform rational approximation of x^α on $[0, 1]$. *Acta Math.*, **190**, 241–306.

Szegö, G. (1939) *Orthogonal Polynomials*. Providence, RI: American Mathematical Society.

Vabishchevich, P. N. (2015) Numerically solving an equation for fractional powers of elliptic operators. *J. Comput. Phys.*, **282**, 289–302.

Vabishchevich, P. N. (2018) Numerical solution of time-dependent problems with fractional power elliptic operator. *Comput. Meth. Appl. Math.*, **18**, 111–128.

Vabishchevich, P. N. (2020) Approximation of a fractional power of an elliptic operator. *Numer. Linear Algebra Appl.*, **27**, e2287.

Van Damme, G. (2020) Legendre, Laguerre and Hermite-Gauss quadrature. *MATLAB Central File Exchange*. Available at https://www.mathworks.com/matlabcentral/fileexchange/26737-legendre-laguerre-and-hermite-gauss-quadrature.s