BIT

Sparsity of Runge–Kutta convolution weights for the three-dimensional wave equation

Lehel Banjai · Maryna Kachanovska

Abstract Wave propagation problems in unbounded homogeneous domains can be formulated as time-domain integral equations. An effective way to discretize such equations in time are Runge–Kutta based convolution quadratures. In this paper the behaviour of the weights of such quadratures is investigated. In particular approximate sparseness of their Galerkin discretization is analyzed. Further, it is demonstrated how these results can be used to construct and analyze the complexity of fast algorithms for the assembly of the fully discrete systems.

Keywords Convolution quadrature · Runge–Kutta methods · Time-domain boundary integral equations · Wave equation

Mathematics Subject Classification (2010) 65R20 · 65L06 · 35L05 · 65M38

Communicated by Axel Ruhe.

L. Banjai
School of Mathematical & Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK
e-mail: l.banjai@hw.ac.uk

M. Kachanovska (✉)
Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany
e-mail: mkachanovska@gmail.com

1 Introduction

In many physical applications, e.g. acoustic and electromagnetic scattering, it is necessary to solve an exterior boundary value problem for the three-dimensional wave equation. Such problems can be effectively treated with the use of time-domain boundary integral equations (TDBIE). The well-posedness of such formulations for the wave equation was analyzed in [3,4].

Solution of time domain boundary integral equations is usually performed with Galerkin time-space methods [16], collocation methods [11] or Laplace-domain approaches. A review of these methods can be found in [10]. However, compared to the field of elliptic problems, fast solvers for time domain boundary integral equations are not as extensively developed. A particularly efficient approach to the solution of retarded potential boundary integral equations is offered in [14,15].

One of the methods for the solution of TDBIE, convolution quadrature [22–24], combines Laplace-domain and time-stepping techniques. It is stable, efficient and does not require intricate, highly accurate quadrature in space in contrast to standard Galerkin time-space methods. The applicability of the method to external boundary-value problems for the wave equation was justified in [24]. These results were supported by extensive numerical experiments in [5] where it was also shown that Runge–Kutta convolution quadrature [25] is preferable to multistep convolution quadrature whenever the scattering domain is non-convex. In [6] the theoretical justification of this fact was given.

In this paper we discretize the time-domain single layer boundary operator by an m-stage A-stable Runge–Kutta convolution quadrature [25]. The weights of the quadrature are boundary integral operators. The main part of this paper is devoted to the investigation of the behaviour of the kernels $w_n^h(d)$ of these operators. We prove estimates that show an exponential decay of $\|w_n^h(d)\|$ away from a neighbourhood of $nh \approx d$, where $h > 0$ denotes the timestep. The paper ends with an illustration of how these results can be used to speed up existing algorithms for the computation of convolution weights.

2 Statement of the problem

Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain with boundary Γ and let $\Omega^c = \mathbb{R}^3 \setminus \overline{\Omega}$ be its complement.

We will consider the homogeneous wave equation set in Ω^c,

$$
\begin{aligned}
\frac{\partial^2 u}{\partial t^2} - \Delta u &= 0 \quad \text{in } [0, T] \times \Omega^c, \\
u(0, .) = \frac{\partial u}{\partial t}(0, .) &= 0 \quad \text{in } \Omega^c, \\
u(t, x) = g(t, x) \quad &\text{on } [0, T] \times \Gamma.
\end{aligned}
\tag{2.1}
$$

The solution u of the above system can be represented as the single-layer potential of an unknown density λ

$$u(t, \tilde{x}) = (\mathscr{S}\lambda)(t, \tilde{x}) = \int\limits_0^t \int\limits_\Gamma \frac{\delta(t - \tau - \|\tilde{x} - y\|)}{4\pi \|\tilde{x} - y\|} \lambda(\tau, y) d\Gamma_y d\tau,$$

$$= \int\limits_\Gamma \frac{\lambda(t - \|\tilde{x} - y\|, y)}{4\pi \|\tilde{x} - y\|} d\Gamma_y, \quad (t, \tilde{x}) \in [0, T] \times \Omega^c,$$

where $\delta(\cdot)$ denotes the Dirac delta function. Single layer potential $\mathscr{S}\lambda$ is also known as the *retarded potential*, the name being justified by the second expression above. For any density λ, the function $u = \mathscr{S}\lambda$ satisfies the first two equations in (2.1), therefore to solve (2.1) it remains to choose λ so that the boundary condition is also satisfied. The single layer potential $\mathscr{S}\lambda$ is continuous across Γ, so letting in the above equation $\tilde{x} \to x \in \Gamma$ and using the boundary condition from (2.1), we obtain a *boundary integral equation* for the unknown density λ

$$g(t, x) = (\mathscr{V}\lambda)(t, x) = \int\limits_0^t \int\limits_\Gamma \frac{\delta(t - \tau - \|x - y\|)}{4\pi \|x - y\|} \lambda(\tau, y) d\Gamma_y d\tau,$$

$$\forall (t, x) \in [0, T] \times \Gamma. \tag{2.2}$$

Here, the operator $\mathscr{V}$ is called the single layer boundary integral operator. For the existence and uniqueness of solutions of this equation see [3].

For further discussion we will require the Laplace transforms of $\mathscr{S}$ and $\mathscr{V}$. With the Laplace transform defined by

$$\mathscr{L}f(s) = \int\limits_0^\infty e^{-st} f(t) dt, \quad \operatorname{Re} s > 0,$$

and causal f, i.e. $f(t) = 0$, for $t \leq 0$, it holds that

$$\mathscr{L}(f(\cdot - r))(s) = \int\limits_0^\infty e^{-st} f(t - r) dt = e^{-sr} \mathscr{L}f(s), \quad r \geq 0. \tag{2.3}$$

Hence the Laplace transforms of $\mathscr{S}$ and $\mathscr{V}$ are given respectively by

$$S(s)\varphi(\tilde{x}) = \int\limits_\Gamma \frac{e^{-s\|\tilde{x} - y\|}}{4\pi \|\tilde{x} - y\|} \varphi(y) d\Gamma_y, \quad \tilde{x} \in \Omega^c,$$

and

$$V(s)\varphi(x) = \int\limits_\Gamma \frac{e^{-s\|x - y\|}}{4\pi \|x - y\|} \varphi(y) d\Gamma_y, \quad x \in \Gamma.$$

Next, we address the time-discretization of retarded potentials.

2.1 Convolution quadrature based on backward differences

Let $h > 0$ denote the timestep and $t_j = jh$ the equally spaced time-points. Then the derivative of a causal function f can be approximated by

$$f'(t) \approx \tfrac{1}{h}(f(t) - f(t - h)).\tag{2.4}$$

Taking the Laplace transform of the above approximation and applying (2.3) gives

$$s\mathscr{L}f(s) \approx \frac{1 - e^{-sh}}{h}\mathscr{L}f(s).\tag{2.5}$$

We can also reverse this procedure: approximate the differentiation symbol s in the Laplace domain by $\frac{1-e^{-sh}}{h} = s + s\,O((sh))$, see (2.5), and compute the inverse Laplace transform of the approximation to obtain the backward difference approximation of the derivative (2.4).

Convolution quadrature of $\mathscr{S}\lambda$ and $\mathscr{V}\lambda$ proceeds in a similar way. First the approximation in the Laplace domain is made

$$V(s)\mathscr{L}\lambda \approx V\left(\tfrac{1-e^{-sh}}{h}\right)\mathscr{L}\lambda\tag{2.6}$$

and then the inverse Laplace transform of the approximation is computed and used as an approximation of $\mathscr{V}\lambda$. Let the following expansion hold

$$V\left(\tfrac{1-e^{-sh}}{h}\right) = \sum_{j=0}^{\infty}\omega_j(V)e^{-shj} = \sum_{j=0}^{\infty}\omega_j(V)e^{-st_j}, \quad \operatorname{Re} s > 0.$$

Remark 2.1 Note that $V(s)$ is an analytic and bounded function of s for $\operatorname{Re} s > 0$,

$$\|V(s)\|_{H^{1/2}(\Gamma)\leftarrow H^{-1/2}(\Gamma)} \le C(\sigma)|s|, \quad \operatorname{Re} s \ge \sigma > 0,$$

see [3]. Also $1 - e^{-sh}$ is an analytic function of e^{-sh} and $\operatorname{Re}(1 - e^{-sh}) > 0$ for $\operatorname{Re} s > 0$. Hence, the above expansion is well defined and the linear operators $\omega_j(V) : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ are bounded.

Using (2.3) again, we see that the inverse Laplace transform of the approximation in (2.6) is given by

$$\mathscr{V}\lambda(t) \approx \sum_{j=0}^{\infty}\omega_j(V)\lambda(t - t_j).$$

Assuming causality of λ we obtain the convolution quadrature approximation at $t = t_n$:

$$\sum_{j=0}^{n}\omega_j(V)\lambda(t_n - t_j) = \sum_{j=0}^{n}\omega_{n-j}(V)\lambda(t_j).$$

Error and stability analysis of this first order time-discretization method and of convolution quadrature based on other A-stable linear multistep methods can be found in [24].

2.2 Runge–Kutta based convolution quadrature

Since A-stable linear multistep methods have order restricted to $p \leq 2$, Runge–Kutta based methods need to be considered. For the importance of high order methods in wave propagation problems see [5,6] and for the analysis of the resulting discrete systems see [25] and [8,9].

Let an m-stage Runge–Kutta method be given by its Butcher tableau

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}.$$

In terms of A, b, and c, an m-stage Runge–Kutta discretization of the initial value problem $y' = f(t, y)$, $y(0) = y_0$, is given by the recurrence

$$Y_{ni} = y_n + h \sum_{j=1}^{m} a_{ij} f(t_n + c_j h, Y_{nj}), \quad i = 1, \ldots, m,$$

$$y_{n+1} = y_n + h \sum_{j=1}^{m} b_j f(t_n + c_j h, Y_{nj});$$

here, h is the time-step and $t_j = jh$. The values Y_{ni} and y_n are approximations to $y(t_n + c_i h)$ and $y(t_n)$, respectively. This Runge–Kutta method is said to be of (classical) order $p \geq 1$ and stage order q if for sufficiently smooth right-hand sides f,

$$Y_{0i} - y(c_i h) = O(h^{q+1}), \text{ for } i = 1, \ldots, m, \quad \text{and} \quad y_1 - y(t_1) = O(h^{p+1}),$$

as $h \to 0$.

The corresponding stability function is defined by $R(z) = 1 + z b^T (I - Az)^{-1} \mathbb{1}$, where $\mathbb{1} = (1 \ \ldots \ 1)^T$ and the following approximation property holds

$$R(z) = e^z + O(z^{p+1}). \tag{2.7}$$

We will make the following assumptions on the stability function of the Runge–Kutta method and will further assume that the coefficient matrix A is invertible. These assumptions are required by the theory of Runge–Kutta convolution quadrature for hyperbolic problems as described in [9].

Assumption 2.1 (a) The Runge–Kutta method is A-stable, namely $|R(z)| \leq 1$ for all z, s.t. $\operatorname{Re} z \leq 0$.

(b) $|R(\infty)| = 1 - b^T A^{-1} \mathbb{1} = 0$.

(c) $|R(iy)| < 1$, for all $y \in \mathbb{R} \setminus \{0\}$.

Note that the second assumption above implies $c_m = 1$. Examples of Runge–Kutta methods satisfying these assumptions are the Radau IIA and Lobatto IIIC methods.

As in the previous section we will want to approximate the derivative of a function f, but this time at the vector of stages:

$$f(t + ch) = \begin{pmatrix} f(t + c_1 h) \\ \vdots \\ f(t + c_m h) \end{pmatrix}.$$

Using (2.3) and proceeding as in the previous section we see that an approximation in the Laplace domain of the form

$$s e^{csh} \mathscr{L} f(s) \approx \frac{\Delta(e^{-sh})}{h} e^{csh} \mathscr{L} f(s)$$

is required, where

$$e^{csh} = \begin{pmatrix} e^{c_1 sh} \\ \vdots \\ e^{c_m sh} \end{pmatrix}$$

and $\Delta(\zeta) : \mathbb{C} \to \mathbb{C}^{m \times m}$ is a matrix valued function with the property

$$\Delta(e^{-z}) e^{cz} \approx z e^{cz}.$$

The next lemma defines such a function and proves some of its properties.

Lemma 2.1 *Let*

$$\Delta(\zeta) = \left(A + \frac{\zeta}{1 - \zeta} \mathbb{1} b^T \right)^{-1} = A^{-1} - \zeta A^{-1} \mathbb{1} b^T A^{-1}. \tag{2.8}$$

Then the following hold:

(a) *For $z \to 0$ it holds*

$$\Delta(e^{-z}) e^{cz} = z e^{cz} + O(z^{q+1}).$$

(b) *If $\mu \notin \sigma(A^{-1})$, then $R(\mu) = \zeta^{-1}$ if and only if $\mu \in \sigma(\Delta(\zeta))$.*

Proof Note that the equality in (2.8) is readily proved using the Sherman–Morrison formula and $b^T A^{-1} \mathbb{1} = 1$, where the latter is implied by Assumption 2.1(b). Result (b) follows from the expression

$$(zI - \Delta(\zeta))^{-1} = A(zA - I)^{-1} - \frac{\zeta}{1 - R(z)\zeta}(zA - I)^{-1}\mathbb{1}b^T(zA - I)^{-1}$$

proved in [25, Lemma 2.4].

Proof of (a) requires a few more steps. In [9, Lemma 2.5], it has been shown that

$$zb^T e^{cz} = e^z - 1 + O(z^{p+1}) \quad \text{and} \quad zAe^{cz} = e^{cz} - \mathbb{1} + O(z^{q+1}).$$

Hence,

$$\begin{aligned}
\Delta(e^{-z})^{-1}ze^{cz} &= e^{cz} - \mathbb{1} + \frac{1}{1 - e^{-z}}\mathbb{1} - \frac{e^{-z}}{1 - e^{-z}}\mathbb{1} + O(z^{q+1}) \\
&= e^{cz} + O(z^{q+1}).
\end{aligned}$$

$\square$

Therefore we are in a similar position as in the previous section. The last step that we need to do is construct the expansion

$$V\left(\frac{\Delta(\zeta)}{h}\right) = \sum_{j=0}^{\infty} W_n^h(V)\zeta^j.$$

Remark 2.2 Note that A-stability and Lemma 2.1(b) imply that the eigenvalues of $\Delta(\zeta)$ for $|\zeta| < 1$ all lie in the right-half complex plane. Therefore the same arguments as in Remark 2.1 tell us that the above expansion is well defined and that the convolution weights $W_n^h(V)$ are $m \times m$ matrices of bounded linear operators mapping from $H^{-1/2}(\Gamma)$ to $H^{1/2}(\Gamma)$.

Recalling the definition of $V(s)$ we see that $W_n^h \equiv W_n^h(V)$ are integral operators defined by

$$W_n^h\lambda(x) = \int_\Gamma w_n^h(\|x - y\|)\lambda(y)d\Gamma_y, \quad x \in \Gamma,$$

where the kernels $w_n^h(d)$ are defined by the corresponding expansion

$$\frac{\exp\left(-\frac{\Delta(\zeta)}{h}d\right)}{4\pi d} = \sum_{n=0}^{\infty} w_n^h(d)\zeta^n. \tag{2.9}$$

Let us denote by g_n and $\mathbf{g}_n$ the following functions

$$g_n(x) = g(nh, x), \quad \mathbf{g}_n(x) = \begin{pmatrix} g(nh + c_1h, x) \\ \vdots \\ g(nh + c_mh, x) \end{pmatrix}.$$

With this notation the Runge–Kutta convolution quadrature of (2.2) is given by

$$g_n(x) = \sum_{i=0}^{n} \left(W_{n-i}^h \lambda_i \right)(x),$$

where λ_n denotes

$$\lambda_n(x) = \begin{pmatrix} \lambda(nh + c_1 h, x) \\ \vdots \\ \lambda(nh + c_m h, x) \end{pmatrix}.$$

In the remainder of the paper we will require the scaled convolution kernels $w_n(d) := 4\pi d w_n^h(hd)$. Notice that $w_n(d)$ are the coefficients of the following expansion:

$$\exp\left(-\Delta(\zeta)d\right) = \sum_{n=0}^{\infty} w_n(d)\zeta^n. \tag{2.10}$$

3 Sparsity of Runge–Kutta convolution weights

Our task in this section is to find estimates for convolution weights $w_n^h(d)$ in terms of d and n. To do so, we first derive bounds for the scaled convolution weights $w_n(d)$ and then use these results to show that similar bounds hold also for $w_n^h(d)$.

The scaled convolution weight $w_n(d)$ for $d > 0$ can also be expressed as

$$w_n(d) = \frac{1}{2\pi i} \oint_\gamma R(z)^{n-1}(I - Az)^{-1}\mathbb{1}b^T(I - Az)^{-1}e^{-zd}dz, \tag{3.1}$$

see [25]. Here, γ represents a contour that encloses all the eigenvalues of A^{-1}.

To prove the main estimates, we need to choose the contour γ carefully. First, we consider the domain $\Upsilon_r, r > 0$:

$$\Upsilon_r = \{z \in \mathbb{C} \ : \ |R(z)| > r\}.$$

We denote by γ_r the boundary of this domain, i.e., $\gamma_r := \partial \Upsilon_r$. Hence, $|R(z)| = r$ holds for all $z \in \gamma_r$. Next, we prove some properties of domains Υ_r.

Let

$$A_+ = \left\{z \in \mathbb{C} \ : \ |R(z)| > |e^z|, \ \text{Re}\, z > 0\right\}$$

denote the order star of R restricted to the right-half complex plane, see [20]. In fact A_+ denotes just the m bounded fingers containing the m, counting multiplicities,

singularities of R. Since $|R(iy)| < 1$ for $y \neq 0$, the origin is the only point of the intersection of the closure of the order star with the imaginary axis and hence

$$\overline{A}_+ \subset \Upsilon_1 \cup \{0\}.$$

Recall that the stability function of the Runge–Kutta method is a rational function

$$R(z) = \frac{P(z)}{Q(z)},$$

where $P(z)$ and $Q(z)$ are polynomials, see [20]. Without loss of generality we can assume that $Q(0) = 1$. The so-called E-polynomial is defined by

$$E(y) = |Q(iy)|^2 - |P(iy)|^2 = e_0 y^{2s} + O(y^{2s+2}). \tag{3.2}$$

It can be shown that for Runge–Kutta methods satisfying Assumption 2.1 the constant e_0 is positive.

   We will need the following lemma. We believe this result to be known, however were not able to find an exact reference. Therefore, we provide its proof in the appendix.

Lemma 3.1 *For an A-stable Runge–Kutta method of order p there exist $q, v > 0$, such that the domain*

$$\{(x, y) \in \mathbb{R}^2 \; : \; |y| < vx^{\frac{1}{\ell}}, \; 0 < x < q\}$$

belongs to Υ_1 (and intersects all the order star fingers). Here

$$\ell = \begin{cases} p + 1, & \text{if } p \text{ is odd,} \\ 2s, & \text{if } p \text{ is even,} \end{cases}$$

where s is defined by (3.2) with $e_0 > 0$.

Lemma 3.2 *Under Assumption 2.1, the domain Υ_1 is located in the open right-half plane and is bounded and connected (possibly multiply).*

Proof The boundedness follows directly from the assumption $R(\infty) = 0$. A-stability and the bound $|R(iy)| < 1$, $y \in \mathbb{R} \setminus \{0\}$ imply that Υ_1 is located in the open right-half plane.

   Let $\widetilde{\Upsilon}_1$ be a connected (possibly multiply) component of Υ_1. Then, by the maximum principle, $\widetilde{\Upsilon}_1$ must contain a singularity of $R(z)$ and the closure of the corresponding finger (minus the origin). According to Lemma 3.1, the intersection of $\widetilde{\Upsilon}_1$ with all the other fingers is nonempty. Since $\widetilde{\Upsilon}_1$ contains all the singularities of Υ_1, by the maximum modulus principle applied to $R(z)$, it coincides with Υ_1. $\qquad\square$

Remark 3.1 The domain Υ_1 is not necessarily simply connected and can contain a hole Υ'. Namely, a bounded domain Υ' can exist, s.t. $R(z)$ vanishes in one of its interior points, $|R(z)| < 1$ inside Υ' and $\partial\Upsilon' \subset \partial\Upsilon_1$. This is the case for the diagonally
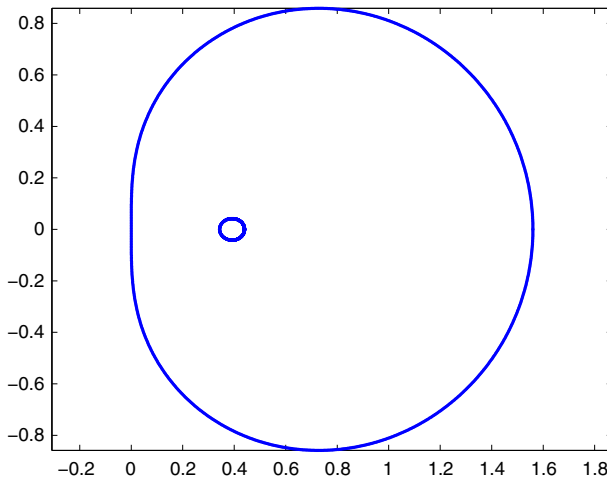
Fig. 1 The boundary of Υ_1 for an *A*-stable DIRK method defined by (3.3)

implicit Runge–Kutta (DIRK) method of order 2, see [1, Theorem 5], defined by the Butcher tableau

$$\begin{array}{c|cc} a & a & 0 \\ 1 & 1-a & a \\ \hline & 1-a & a \end{array}, \quad a = 1 + \frac{1}{\sqrt{2}}. \tag{3.3}$$

Its stability function is given by $R(z) = \frac{1+(1-2a)z}{(1-az)^2}$ and satisfies Assumption 2.1, see also [20, Section 4.6, p.98]. The boundary of the domain Υ_1 for this method is shown in Fig. 1.

Remark 3.2 Note that by continuity of $R(z)$, in a sufficiently small vicinity of $r = 1$, Υ_r stays bounded and connected.

Corollary 3.1 *If the stability function of a Runge–Kutta method that satisfies Assumption 2.1 coincides with a Padé approximant for the exponential, the domain Υ_1 is simply connected. Further, for small enough $\varepsilon > 0$, the perturbed domain $\Upsilon_{1\pm\varepsilon}$ is also simply connected.*

Proof For the proof we need two ingredients:

1. Ehle's Conjecture [26, Theorem 7]. Any Padé approximation $R(z) = \frac{P(z)}{Q(z)}$, $\deg P = k$, $\deg Q = m$ is *A*-stable iff $m - 2 \le k \le m$.
2. All zeros of such Padé approximants lie in the open left-half plane, see [12,13].

Hence, the existence of a bounded domain Υ', s.t. $|R(z)| < 1$ inside Υ' and $\partial \Upsilon' \subset \partial \Upsilon_1$ (i.e. a hole in Υ_1), contradicts the maximum modulus principle applied to the analytic function $\frac{1}{R(z)}, z \in \mathbb{C}_+$. The proof for the perturbed domain is similar. The main thing to

notice is that if ε is small enough, then there exists $\varepsilon' > 0$ such that the zeros of $R(z)$ are contained in the half-plane $\mathrm{Re}\, z < -\varepsilon'$ and that $\Upsilon_{1\pm\varepsilon}$ is contained in $\mathrm{Re}\, z > -\varepsilon'$. □

Again, the stability functions of the Radau IIA and Lobatto IIIC methods satisfy the assumptions of the above corollary.

Lemma 2.1(b) provides us with an easy way to draw the curves $\gamma_r = \partial\Upsilon_r$, i.e., by plotting the eigenvalues of $\Delta(\zeta)$ for all $|\zeta| = 1/r$. In [5] the multiplicity of the eigenvalues of $\Delta(\zeta)$ for the 2- and 3-stage Radau IIA Runge–Kutta methods was discussed. In both cases $\Delta(\zeta)$ has only simple eigenvalues for $|\zeta| = 1$, as explained also by Corollary 3.1. For the 2-stage version, eigenvalues of multiplicity greater than 1 occur only for $\zeta = -5 \pm 3\sqrt{3}$. For a plot of these curves at the critical values and $r = 1$ see Fig. 2.

Since the domain Υ_1 is connected, by Remark 3.2, Υ_r is also connected (not necessarily simply) for sufficiently small $|r - 1|$. Let δ_* be such that

$$\text{the domain } \Upsilon_r \text{ is bounded and connected for } |r - 1| < \delta_*. \tag{3.4}$$

From now on we will make use of γ chosen as the positively oriented boundary of a domain Υ_r, i.e. $\gamma = \gamma_r$, with $|r - 1| < \delta_*$.

Remark 3.3 Note that the total length of γ_r is bounded, see [2, Lemma 3], by

$$|\gamma_r| \le 4md(\gamma_r),$$
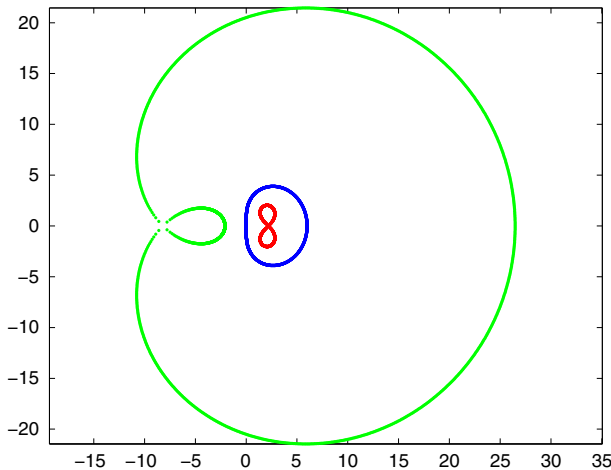
where $d(\gamma_r)$ is the diameter of γ_r.

Fig. 2 Curves γ_r, for the 2-stage Radau IIA method, are plotted for $r = 1$ (the *middle* curve in blue) and the critical values $r = r_1 = 5 + 3\sqrt{3}$ (the *outer* curve in green) and $r = r_2 = 3\sqrt{3} - 5$ (the *inner* curve in red). For $r > r_1$ and $r < r_2$ the curve splits into two disjoint curves

From (3.1) it follows that the Euclidean norm of $w_n(d)$ is bounded as

$$
\|w_n(d)\| \leq \frac{1}{2\pi} \left\| \int_{\gamma_r} R(z)^{n-1} e^{-zd} (I - Az)^{-1} \mathbb{1} b^T (I - Az)^{-1} dz \right\|
$$

$$
\leq \frac{1}{2\pi} |\gamma_r| r^{n-1} \max_{z \in \gamma_r} \|e^{-zd} (I - Az)^{-1} \mathbb{1} b^T (I - Az)^{-1}\|.
$$

Denoting by $U(z) = (I - Az)^{-1}$, we can deduce the bound

$$
\max_{z \in \gamma_r} \|(I - Az)^{-1} \mathbb{1} b^T (I - Az)^{-1}\| \leq \max_{z \in \gamma_r} \|(I - Az)^{-1}\|^2 \|\mathbb{1} b^T\|
$$

$$
\leq \max_{z \in \gamma_r} \|U(z)\|^2 \|b\| \sqrt{m},
$$

which implies that

$$
\|w_n(d)\| \leq \frac{1}{2\pi} r^{n-1} |\gamma_r| \|b\| \sqrt{m} \max_{z \in \gamma_r} |e^{-zd}| \max_{z \in \gamma_r} \|U(z)\|^2. \tag{3.5}
$$

To understand the behaviour of a scaled convolution weight $w_n(d)$ we need to find a bound on $\max_{z \in \gamma_r} |e^{-zd}|$. To do so, we use the fact that the stability function $R(z)$ approximates e^z, see (2.7), and thus $\max_{z \in \gamma_r} |e^{-zd}|$ can be expressed via the value of $|R(z)|$ on γ_r.

For a Runge–Kutta method of order p we can write

$$
R(z) = e^z + f(z),
$$

where $f(z) = O(z^{p+1})$.

Let us consider $z \in \gamma_r$ and $d \in \mathbb{R}_{>0}$. Multiplying the last equation by $e^{-z} R(z)^{-1}$, we get the following identity:

$$
e^{-z} = R(z)^{-1} \left(1 + f(z) e^{-z}\right). \tag{3.6}
$$

Clearly,

$$
\max_{z \in \gamma_r} |e^{-z}| = e^{-x_0},
$$

where $x_0 = \min_{z \in \gamma_r} \operatorname{Re} z$. Let $z_0 = x_0 + i y_0 \in \gamma_r$ be a point such that

$$
\operatorname{Re} z' \geq \operatorname{Re} z_0, \quad \text{for all } z' \in \gamma_r. \tag{3.7}
$$

Note that such a point is not necessarily unique. Taking the modulus and raising both sides of (3.6) to the dth power, we obtain

$$
\max_{z \in \gamma_r} |e^{-zd}| = |e^{-z_0}|^d = |R(z_0)^{-1} \left(1 + f(z_0) e^{-z_0}\right)|^d.
$$

Hence

$$\max_{z \in \gamma_r} |e^{-zd}| = r^{-d}|1 + f(z_0)e^{-z_0}|^d, \tag{3.8}$$

where $z_0 = x_0 + iy_0$ is a point satisfying (3.7).

Remark 3.4 The points $z_0 \in \gamma_r$ with the smallest real part can alternatively be characterized by the following two properties:

– For all z such that $\text{Re } z < \text{Re } z_0$, it holds that

$$|R(z_0)| > |R(z)|. \tag{3.9}$$

This is due to the analyticity of $R(z)$, as well as the definition of γ_r and points z_0.
– Let us fix $x_0 = \text{Re } z_0$. Note that the continuous function $|R(x_0 + iy)| \to 0$ as $y \to \infty$. Then y_0, see (3.7), are the points in which $|R(x_0 + iy)|$ achieves its local maximum. Let us show this. We assume the contrary, namely, that there exists $y' \in \mathbb{R}$ s.t.

$$|R(x_0 + iy')| > |R(x_0 + iy_0)|.$$

Since $R(z)$ is analytic in $z' = x_0 + iy'$, there exists an ϵ-neighborhood of z' s.t.

$$|R(z)| > |R(x_0 + iy_0)|, \quad |z - z'| \le \epsilon.$$

Taking $z = z' - \epsilon = (x_0 - \epsilon) + iy'$, we arrive at the contradiction to (3.9).

In order to bound this product we need to understand how x_0 and $x_0 + iy_0$ behave. This question has been studied in [19] examining the behaviour of $R(z)$ in the order star [26]. Namely, when $r \to 1$, all such points z_0 satisfy $|z_0| \le C|r - 1|^a$ for some $C, a > 0$. This and the fact that $f(z) = O(z^{p+1})$ will allow us to obtain bounds on the scaled convolution weights. Here we will employ the results from [19].

Definition 3.1 [19] Given a rational function $R(z)$ we define the error growth function as the real-valued function $\phi(x) := \sup_{\text{Re } z < x} |R(z)|$.

Theorem 3.1 (Theorem 7 in [19]) *Let $R(z) = \frac{P(z)}{Q(z)}$ be an A-stable approximation to e^z of exact order $p \ge 1$, i.e.,*

$$R(z) = \frac{P(z)}{Q(z)} = e^z + C_{p+1}z^{p+1} + O(z^{p+2}), \quad \text{for } z \to 0, \ C_{p+1} \ne 0. \tag{3.10}$$

Furthermore, assume $|R(iy)| < 1$ for $y \ne 0$ and $|R(\infty)| < 1$. Then we have for $x \to 0$:

– *if p is odd,*

$$\phi(x) = e^x + O(x^{p+1}).$$

– if p is even and $(-1)^{p/2} C_{p+1} x < 0$,

$$\phi(x) = e^x + O(x^{p+1}).$$

– if p is even and $(-1)^{p/2} C_{p+1} x > 0$,

$$\phi(x) = e^x + O(x^{1+p/(2s-p)}),$$

where s is defined by

$$E(y) = |P(iy)|^2 - |Q(iy)|^2 = e_0 y^{2s} + O(y^{2s+2}), \ e_0 > 0. \qquad (3.11)$$

Remark 3.5 [19] For $x < \mathrm{Re}\, \lambda_{min}$, with λ_{min} being an eigenvalue of A^{-1} with the smallest real part, $\phi(x)$ is a strictly monotonically increasing continuous function.

The following proposition shows that for $r \to 1$, $x_0 = \min_{z \in \gamma_r} \mathrm{Re}\, z$ is close to $r - 1$.

Proposition 3.1 *Let $R(z)$ be the stability function of the Runge–Kutta method satisfying Assumption* 2.1, *let* (3.10) *hold and let $x_0 = \min_{z \in \gamma_r} \mathrm{Re}\, z$. Then for $r \to 1$:*

$$x_0 = r - 1 + o(|r - 1|).$$

Proof By definition, $|R(z)| = r$ on γ_r. Since the error growth function $\phi(x)$ is a strictly monotonically increasing continuous function, see Remark 3.5, $\phi(x_0) = r$. The statement of the proposition follows from the application of the implicit function theorem to the 3 cases of Theorem 3.1 and the fact that $\phi(0) = 1$, $\frac{d\phi}{dx}(0) = 1$. □

Next proposition shows that when $r \approx 1$, the point z_0 defined by (3.7) lies in a small circle centered at the origin.

Proposition 3.2 *Let $R(z)$ be the stability function of the Runge–Kutta method satisfying Assumption* 2.1 *and* (3.10). *Then there exist $\delta_0 > 0$ and $K > 0$, s.t., for all r, with $|r - 1| < \delta_0$, the point $z_0 \in \gamma_r$ defined by* (3.7) *lies inside one of the circles specified below:*

1. *for p odd:*

$$|z_0| \leq K |r - 1|.$$

2. *for p even:*
   (a) *if $r > 1$ and $(-1)^{\frac{p}{2}} C_{p+1} < 0$ or $r < 1$ and $(-1)^{\frac{p}{2}} C_{p+1} > 0$,*

$$|z_0| \leq K |r - 1|.$$

(b) *if $r > 1$ and $(-1)^{\frac{p}{2}}C_{p+1} > 0$ or $r < 1$ and $(-1)^{\frac{p}{2}}C_{p+1} < 0$,*

$$|z_0| \le K|r-1|^{\frac{1}{2s-p}},$$

where s is defined by (3.11).

Proof The proof of this statement follows closely the proof of Theorem 7 in [19]. Recall that $z_0 = x_0 + iy_0$ is a (not necessarily unique) point in which $|R(x_0 + iy)|$, $y \in \mathbb{R}$, achieves its local maximum, see Remark 3.4. As argued in [19], for $x_0 \to 0$ the maximum of $|R(x_0 + iy)|$, $y \in \mathbb{R}$, lies inside the order star close to the origin. We consider the following cases for $r \to 1$ (and, consequently, $x_0 \to 0$, see Proposition 3.1):

1. p is odd. As shown in the proof of Theorem 7 in [19], the local extrema of $|R(x_0 + iy)|$ lie asymptotically (as $x_0 \to 0$) on the lines $y = x_0 \tan(k\pi/p)$, $k = 0, 1, \ldots, p-1$. Since $|R(z)|$ achieves an extremum at z_0, it follows that $|z_0| \le C|x_0|$, where $C > 0$ and depends on the Runge–Kutta method. Proposition 3.1 gives an expression for x_0 (using $\phi(x_0) = r$):

$$x_0 = r - 1 + o(|r-1|).$$

Hence,

$$|z_0| \le K|r-1|,$$

for some $K > 0$.

2. p is even.
   (a) As proved in Theorem 7 in [19], for $(-1)^{p/2}C_{p+1}x_0 < 0$ and $x_0 \to 0$, the point z_0 satisfies $|z_0| \le C|x_0|$, $C > 0$. The statement of the proposition is then obtained with similar arguments as in the previous case and the fact that $\operatorname{sgn}x_0 = \operatorname{sgn}(r-1)$.
   (b) For the last case, namely $(-1)^{p/2}C_{p+1}x_0 > 0$, in the proof of Theorem 7 in [19] it was shown that the local extrema of $|R(x_0 + iy)|$, $y \in \mathbb{R}$, for $x_0 \to 0$, are achieved near the imaginary axis and lie on a curve $y^{2s-p} = Dx_0$, for some $D \in \mathbb{R}$ and with s defined in (3.11).
   Then there exists $C > 0$, such that for all sufficiently small x_0,

$$|z_0| \le C\left(|x_0| + |x_0|^{1/(2s-p)}\right) = C|x_0|^{1/(2s-p)}\left(|x_0|^{1-1/(2s-p)} + 1\right).$$

According to Proposition 3.4 in [20] $2s \ge p+1$, therefore, for even p, $2s \ge p+2$. This implies that $|x_0|^{1-1/(2s-p)} = o(|x_0|)$, hence

$$|z_0| \le K|r-1|^{\frac{1}{2s-p}},$$

for some $K > 0$. $\qquad\square$

Now we have all the estimates necessary to prove the next proposition on the decay of the scaled convolution weights.

Proposition 3.3 *Let $R(z)$ be the stability function of an m-stage Runge–Kutta method of order p satisfying Assumption* 2.1 *and* (3.10).

Let s be defined by (3.11). *Then there exist positive constants G, G', C, C' and $\bar{\delta} \in (0, 1)$, such that for $n \geq 1$ and $0 < \delta < \bar{\delta}$ the following estimates hold:*

1. *p is odd*

$$
\begin{aligned}
\|w_n(d)\| &\leq G(1-\delta)^{n-d}(1+C\delta^{p+1})^d \quad \text{for } d \leq n, \\
\|w_n(d)\| &\leq G'(1+\delta)^{n-d}(1+C'\delta^{p+1})^d \quad \text{for } d > n;
\end{aligned}
\tag{3.12}
$$

2. *p is even*
   (a) $C_{p+1}(-1)^{\frac{p}{2}} > 0$

$$
\begin{aligned}
\|w_n(d)\| &\leq G(1-\delta)^{n-d}(1+C\delta^{p+1})^d \quad \text{for } d \leq n, \\
\|w_n(d)\| &\leq G'(1+\delta)^{n-d}(1+C'\delta^{\frac{p+1}{2s-p}})^d \quad \text{for } d > n;
\end{aligned}
\tag{3.13}
$$

   (b) $C_{p+1}(-1)^{\frac{p}{2}} < 0$

$$
\begin{aligned}
\|w_n(d)\| &\leq G(1-\delta)^{n-d}(1+C\delta^{\frac{p+1}{2s-p}})^d \quad \text{for } d \leq n, \\
\|w_n(d)\| &\leq G'(1+\delta)^{n-d}(1+C'\delta^{p+1})^d \quad \text{for } d > n.
\end{aligned}
\tag{3.14}
$$

The scaled convolution weight $w_0(d)$ satisfies:

$$
\|w_0(d)\| \leq \exp(-\mu d),
\tag{3.15}
$$

for some $\mu > 0$.

Constants $G, G', C, C', \bar{\delta}, \mu$ depend only on the Runge–Kutta method and are independent of n and d.

Proof Let us start with $w_0(d)$. From the definition of the scaled convolution weights

$$
\begin{aligned}
\exp(-\Delta(\zeta)d) &= \sum_{n=0}^{\infty} w_n(d)\,\zeta^n, \\
\Delta(\zeta) &= A^{-1} - \zeta A^{-1}\mathbb{1}b^T A^{-1},
\end{aligned}
$$

it follows that $w_0(d) = \exp(-A^{-1}d)$. All the eigenvalues of A lie on the right of the imaginary axis (due to the A-stability of the Runge–Kutta method) and hence the same holds for the eigenvalues of A^{-1}. The bound on $w_0(d)$ can then be obtained from the definition of the matrix exponential.

For the general case $w_n(d)$, $n \geq 1$, we use the bounds derived before, inserting (3.8) into (3.5):

$$
\begin{aligned}
\|w_n(d)\| &\leq \frac{1}{2\pi} r^{n-1} \|b\| \sqrt{m} |\gamma_r| \max_{z \in \gamma_r} |e^{-zd}| \max_{z \in \gamma_r} \|U(z)\|^2 \\
&= \frac{1}{2\pi} r^{n-d-1} |\gamma_r| \max_{z \in \gamma_r} \|U(z)\|^2 \|b\| \sqrt{m} |1 + f(z_0) e^{-z_0}|^d, \quad (3.16)
\end{aligned}
$$

where $f(z) = R(z) - e^z$ and z_0 is such that $\mathrm{Re}\, z' \geq \mathrm{Re}\, z_0$ for all $z' \in \gamma_r$.

Let us first derive a bound for $|1 + f(z_0) e^{-z_0}|$. For $|z| < \frac{1}{\lambda_0}$, where λ_0 is the spectral radius of A, we can expand $R(z) = 1 + z b^T (I - Az)^{-1} \mathbb{1}$ with the help of Neumann series to obtain an explicit expression for $f(z)$:

$$
f(z) = R(z) - e^z = z \sum_{l=p}^{\infty} b^T A^l \mathbb{1} z^l - \sum_{l=p+1}^{\infty} \frac{z^l}{l!}.
$$

For $|z| < \frac{1}{\|A\|}$, we can trivially bound

$$
|1 + f(z) e^{-z}|^d \leq \left(1 + C|z|^{p+1}\right)^d, \quad (3.17)
$$

where C depends on the Runge–Kutta method, but not on z or d.

Now let $n > d$. We choose $r < 1$, $r = 1 - \delta$, $0 < \delta < \min\left\{\delta_*, \frac{1}{\|A\|}, 1\right\}$. Here δ_* is the constant from (3.4).

Then the bound (3.16), using (3.17), can be written as

$$
\|w_n(d)\| \leq \frac{1}{2\pi} (1 - \delta)^{n-d-1} |\gamma_{1-\delta}| \max_{z \in \gamma_{1-\delta}} \|U(z)\|^2 \|b\| \sqrt{m} \left(1 + C|z_0|^{p+1}\right)^d,
$$

where z_0 is such that $\mathrm{Re}\, z_0 < \mathrm{Re}\, z$ for all $z \in \gamma_{1-\delta}$.

The total length of $\gamma_{1-\delta}$ as well as $\max_{z \in \gamma_{1-\delta}} \|U(z)\|$ can be bounded by constants that depend on the Runge–Kutta method, see also Lemma 3.2 and Remarks 3.2 and 3.3. Applying Proposition 3.2 to estimate $(1 + C|z_0|^{p+1})^d$, we obtain the required expressions for the case $n > d$.

The bound for $n < d$ can be obtained similarly by setting $r = 1 + \delta$, with $0 < \delta < \min\left\{\delta_*, \frac{1}{\|A\|}, 1\right\}$. □

Remark 3.6 Note that for even p the above bounds imply that when $2s - p < p + 1$ scaled convolution weights decay exponentially. However, since $2s \leq 2m$ (m is the number of stages and the degree of the denominator in $R(z) = \frac{P(z)}{Q(z)}$), for exponential decay it suffices that $p \geq m$.

Remark 3.7 From the proof it can be seen that the increase of widths of (approximate) supports of convolution weights is due to the term $|1 + f(z_0) e^{-z_0}|$, which we bounded by a constant greater than 1. We have not observed in numerical experiments a case

where this term is noticeably smaller than 1. Note that if this term were smaller than 1, the convolution weights $w_n(d)$ would decay exponentially with increasing $n \approx d$.

We have shown that scaled convolution weights $w_n(d)$ exhibit exponential decay outside of a neighborhood of $n \approx d$, which is an expression of the strong Huygens principle. Additionally, the above estimates suggest that the size of the approximate support of a convolution weight $w_n(d)$ increases with n. Let us examine this in more detail.

We consider Runge–Kutta methods satisfying the following assumption.

Assumption 3.1 Let the order of the Runge–Kutta method be $p \geq 1$. If p is even, then let $2s - p < p + 1$, where s is as in (3.11), see also Remark 3.6.

In particular, this assumption holds true for the Radau IIA and Lobatto IIIC methods. For $n \geq 1$, $w_n(d)$ is bounded by $\epsilon > 0$ outside the interval $\left[d_1^{(n,\epsilon)}, d_2^{(n,\epsilon)}\right]$, where

$$
\begin{aligned}
d_1^{(n,\epsilon)} &= \sup \left\{ d \geq 0 : \ \|w_n(d')\| < \epsilon, \ \text{ for all } 0 \leq d' < d \right\}, \\
d_2^{(n,\epsilon)} &= \inf \left\{ d \geq 0 : \ \|w_n(d')\| < \epsilon, \ \text{ for all } d' > d \right\}.
\end{aligned}
$$

The set

$$
\left\{ d : \ \|w_n(d')\| < \epsilon, \ 0 \leq d' < d \right\}
$$

is non-empty for $n \geq 1$, since $w_n(0) = 0$ (as we show in the proof of Proposition 3.4) and $w_n(d)$ is smooth as a function of $d \geq 0$. The set

$$
\left\{ d : \ \|w_n(d')\| < \epsilon, \ d' > d \right\}
$$

is non-empty for all $n \geq 0$ by Proposition 3.3 for the Runge–Kutta methods under consideration. Hence, the values $d_1^{(n,\epsilon)}, d_2^{(n,\epsilon)}$ are defined for all $n \geq 1$.

To find estimates on $d_1^{(n,\epsilon)}$ and $d_2^{(n,\epsilon)}$, $n > 0$, we make use of the bounds of Proposition 3.3 that can be written in a more general form:

$$
\begin{aligned}
\|w_n(d)\| &\leq G(1 - \delta)^{n-d}(1 + C\delta^\alpha)^d \quad \text{for all } d \leq n, \\
\|w_n(d)\| &\leq G'(1 + \delta)^{n-d}(1 + C'\delta^{\alpha'})^d \quad \text{for all } d > n,
\end{aligned}
$$

for all $0 < \delta < \bar{\delta} < 1$ and constants $C, C', G, G', \alpha, \alpha', \bar{\delta} > 0$ depending only on the Runge–Kutta method. We can estimate $d_1^{(n,\epsilon)}$ and $d_2^{(n,\epsilon)}$ from

$$
G(1 - \delta)^{n-d}(1 + C\delta^\alpha)^d \leq \epsilon, \tag{3.18}
$$

$$
G'(1 + \delta)^{n-d}(1 + C'\delta^{\alpha'})^d \leq \epsilon. \tag{3.19}
$$

The first inequality is satisfied if

$$
\begin{aligned}
d &\leq n \log \frac{1}{1-\delta} \left(\log \frac{1}{1-\delta} + \log(1 + C\delta^\alpha) \right)^{-1} \\
&\quad - \log \frac{G}{\epsilon} \left(\log \frac{1}{1-\delta} + \log(1 + C\delta^\alpha) \right)^{-1} \\
&= n - \left(\log \frac{1}{1-\delta} + \log(1 + C\delta^\alpha) \right)^{-1} \left(n \log \left(1 + C\delta^\alpha \right) + \log \frac{G}{\epsilon} \right).
\end{aligned}
$$

Hence, for small enough δ, there exist constants $c_1, c_2 > 0$ s.t. (recall that $\alpha > 1$)

$$
d \leq n - \frac{c_1 n \delta^\alpha + \log \frac{G}{\epsilon}}{\delta(1 - c_2 \delta)}
$$

implies (3.18). This in turn implies that

$$
d_1^{(n,\epsilon)} \geq n - \frac{c_1 n \delta^{\alpha-1} + \delta^{-1} \log \frac{G}{\epsilon}}{1 - c_2 \delta}.
$$

Let us choose δ to minimize the expression $c_1 n \delta^{\alpha-1} + \delta^{-1} \log \frac{G}{\epsilon}$ as $n \to \infty$. In particular, setting

$$
\delta = c n^{-\frac{1}{\alpha}} \log^{\frac{1}{\alpha}} \frac{G}{\epsilon}, \tag{3.20}
$$

with c being a small constant independent of n, ϵ (chosen so that $\delta < \bar{\delta}$, where $\bar{\delta}$ is defined in Proposition 3.3), ensures that

$$
d_1^{(n,\epsilon)} \geq n - C_1 n^{\frac{1}{\alpha}} \log^{1-\frac{1}{\alpha}} \frac{G}{\epsilon}, \tag{3.21}
$$

for some C_1 depending on the Runge–Kutta method. Similarly,

$$
d_2^{(n,\epsilon)} \leq n + C_1' n^{\frac{1}{\alpha'}} \log^{1-\frac{1}{\alpha'}} \frac{G'}{\epsilon}, \tag{3.22}
$$

for $C_1' > 0$ that depends on the Runge–Kutta method. To check this estimate, we plot the dependence of numerically determined values $\Delta_1^{n,\epsilon} = n - d_1^{n,\epsilon}$ and $\Delta_2^{n,\epsilon} = d_2^{n,\epsilon} - n$ on n in Fig. 3 for different Runge–Kutta methods.

Our estimates predict that for methods with odd order p, namely BDF1 ($p = 1$), 2-stage Radau IIA ($p = 3$) and 3-stage Radau IIA ($p = 5$), $\Delta_1^{n,\epsilon}$, $\Delta_2^{n,\epsilon}$ increase as $O\left(n^{\frac{1}{p+1}}\right)$. This is in close agreement with the numerical results shown in Fig. 3.

For Runge–Kutta methods of even orders obtained estimates predict that for larger n and d the width of a convolution weight gets larger in a non-symmetric manner: $\Delta_1^{n,\epsilon}$ can get larger with increasing n faster than $\Delta_2^{n,\epsilon}$ or vice versa. For larger n
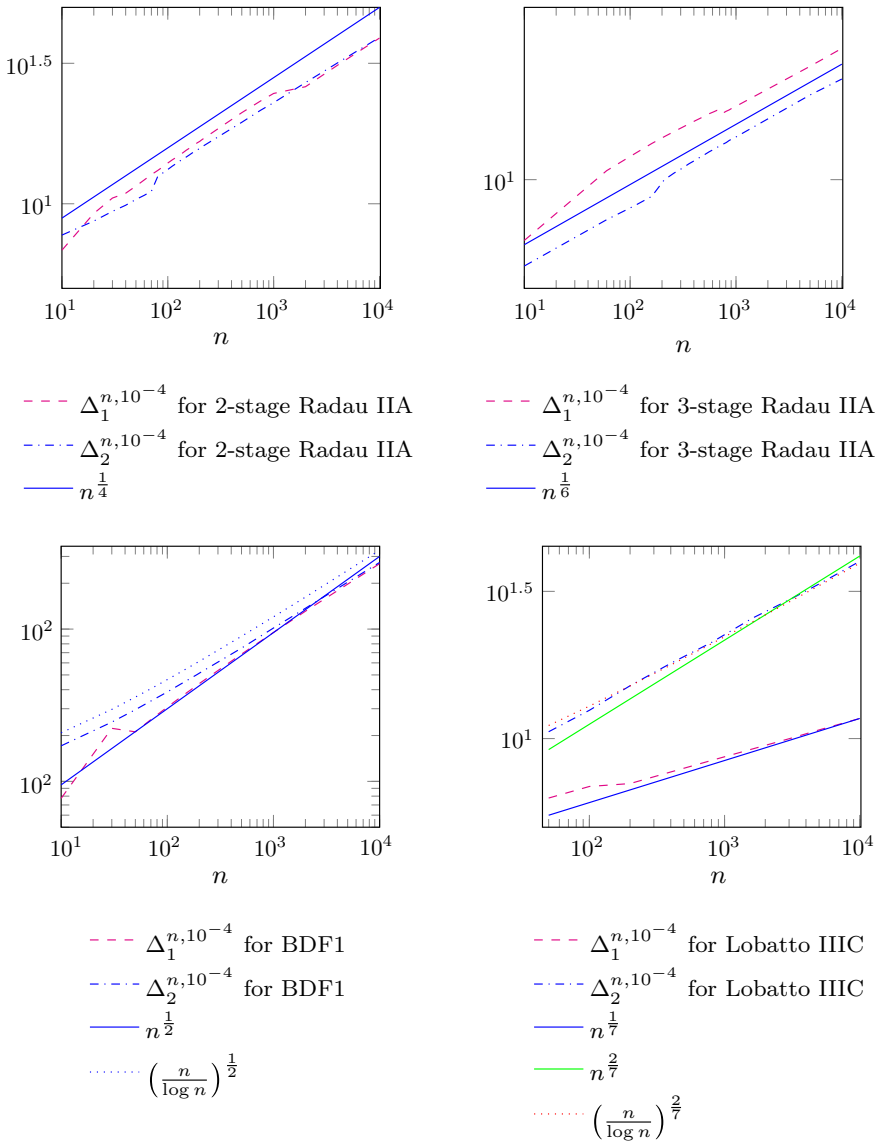
Fig. 3 Dependence of $\Delta_1^{n,10^{-4}}$, $\Delta_2^{n,10^{-4}}$ on n for different Runge–Kutta methods

and d the asymmetry becomes more visible. This can be illustrated by the Lobatto IIIC method of 6th order. Numerical experiments indicate that with increasing n the part of the approximate support of the convolution weight $w_n(d)$ of the Lobatto IIIC method corresponding to $d < n$ increases slower than the part of the approximate support corresponding to $d > n$. This effect can be explained by estimates (3.13) as follows. The stability function of the 4-stage Lobatto IIIC method is the $(2, 4)$-Padé approximation to e^z. For such approximants the sign of C_{p+1} is negative (see, for
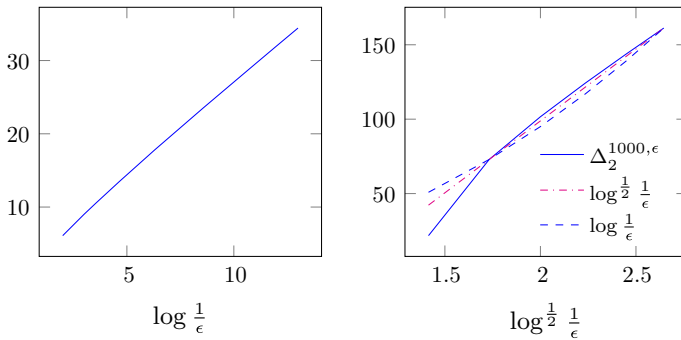
Fig. 4 We examine the dependence of $\Delta_2^{n,\epsilon}$ for the BDF1 method on ϵ. In the *left plot* we show the dependence of $\Delta_2^{10,\epsilon}$ on $\log\frac{1}{\epsilon}$, and in the *right plot* the dependence of $\Delta_2^{10^3,\epsilon}$ on $\log^{\frac{1}{2}}\frac{1}{\epsilon}$ is demonstrated. We compare the rate of growth of $\Delta_2^{10^3,\epsilon}$ to $O(\log\frac{1}{\epsilon})$ and $O(\log^{\frac{1}{2}}\frac{1}{\epsilon})$: this rate is closer to $O(\log^{\frac{1}{2}}\frac{1}{\epsilon})$ rather than $O(\log\frac{1}{\epsilon})$

example, [20, Theorem 3.11]); then the sign of $C_{p+1}(-1)^{\frac{p}{2}}$ is positive. According to the estimates (3.13), $\Delta_1^{n,\epsilon} = O\left(n^{\frac{1}{p+1}}\right)$, with $p = 6$, while $\Delta_2^{n,\epsilon} = O\left(n^{\frac{2s-p}{p+1}}\right)$, with $s = 4$ for Lobatto IIIC (this value can be obtained by examining $E(y) = O(y^{2s})$ for small $y \in \mathbb{R}$). Again, the results in Fig. 3 are in close agreement with these estimates.

Remark 3.8 The choice of δ provided by (3.20) is valid only if

$$\left(\frac{\log\frac{G}{\epsilon}}{n}\right)^{\frac{1}{\alpha}}$$

can be bounded by a constant (otherwise it is impossible to choose $c > 0$ independent of n, $\epsilon > 0$ in (3.20)). Hence, the suggested dependence of $\Delta_1^{n,\epsilon}$, $\Delta_2^{n,\epsilon}$ on the accuracy $\epsilon > 0$ is not necessarily valid in the regime when n is fixed and $\epsilon \to 0$. In this case an optimal choice of δ is $\delta = c > 0$, and $\Delta_1^{n,\epsilon}$, $\Delta_2^{n,\epsilon}$ increase as $O\left(\log\frac{1}{\epsilon}\right)$ rather than $O\left(\log^{\frac{1}{\alpha}}\frac{1}{\epsilon}\right)$.

This can be seen in Fig. 4, where we plot $\Delta_2^{10,\epsilon}$ and $\Delta_2^{10^3,\epsilon}$ for BDF1 convolution weights for a range of $\epsilon > 0$.

The next proposition is a corollary of Proposition 3.3 and shows that (unscaled) convolution weights $w_n^h(d)$ also experience exponential decay away from $\frac{d}{h} \approx n$.

Proposition 3.4 *Let $R(z)$ be the stability function of an m-stage Runge–Kutta method of order p satisfying Assumptions 2.1 and (3.10).*

Let s be defined by (3.11). Then there exist positive constants G, G', C, C' and $\bar\delta \in (0, 1)$, such that for $n \geq 1$ and $0 < \delta < \bar\delta$ the following estimates hold:

1. *p is odd*

$$\|w_n^h(d)\| \leq \frac{G}{h}(1-\delta)^{n-\frac{d}{h}}(1+C\delta^{p+1})^{\frac{d}{h}} \quad for \ \frac{d}{h} \leq n,$$
$$\|w_n^h(d)\| \leq \frac{G'}{d}(1+\delta)^{n-\frac{d}{h}}(1+C'\delta^{p+1})^{\frac{d}{h}} \quad for \ \frac{d}{h} > n; \tag{3.23}$$

2. *p is even*
   (a) $C_{p+1}(-1)^{\frac{p}{2}} > 0$

$$\|w_n^h(d)\| \leq \frac{G}{h}(1-\delta)^{n-\frac{d}{h}}(1+C\delta^{p+1})^{\frac{d}{h}} \quad for \ \frac{d}{h} \leq n,$$
$$\|w_n^h(d)\| \leq \frac{G'}{d}(1+\delta)^{n-\frac{d}{h}}(1+C'\delta^{\frac{p+1}{2s-p}})^{\frac{d}{h}} \quad for \ \frac{d}{h} > n; \tag{3.24}$$

   (b) $C_{p+1}(-1)^{\frac{p}{2}} < 0$

$$\|w_n^h(d)\| \leq \frac{G}{h}(1-\delta)^{n-\frac{d}{h}}(1+C\delta^{\frac{p+1}{2s-p}})^{\frac{d}{h}} \quad for \ \frac{d}{h} \leq n,$$
$$\|w_n^h(d)\| \leq \frac{G'}{d}(1+\delta)^{n-\frac{d}{h}}(1+C'\delta^{p+1})^{\frac{d}{h}} \quad for \ \frac{d}{h} > n. \tag{3.25}$$

The convolution weight $w_0^h(d)$ satisfies:

$$\|w_0^h(d)\| \leq \frac{\exp(-\mu\frac{d}{h})}{4\pi d},$$

for some $\mu > 0$.

Constants $G, G', C, C', \bar{\delta}, \mu$ depend only on the Runge–Kutta method and do not depend on n, d and h.

Proof Let us again start with $w_0^h(d)$. From the definition of the scaled convolution weights it follows that $w_0^h(d) = \frac{w_0(\frac{d}{h})}{4\pi d}$, and the required bound can be obtained from (3.15). Note, however, that the convolution weight $w_0^h(d)$ has a singularity at $d = 0$.

Bounds for the case $\frac{d}{h} > n$ can be obtained straightforwardly from expressions (3.12, 3.13, 3.14) applied to $w_n^h(d) = \frac{w_n(\frac{d}{h})}{4\pi d}$.

The case $\frac{d}{h} \leq n$ has to be treated separately: we cannot directly apply Proposition 3.3 for bounding $w_n^h(d) = \frac{w_n(\frac{d}{h})}{4\pi d}$, since for small d this bound would be far from optimal. We will proceed as follows. First, we will show that a scaled convolution weight $w_n(d)$ has a zero at $d = 0$ of multiplicity at least n. Next, this fact and ideas from the proof of Proposition 3.3 will be used to demonstrate that away from a neighbourhood of $d \approx nh$ convolution weights $w_n^h(d)$ decay exponentially.

Let us first expand the generating function of the scaled convolution weights $e^{-\Delta(\zeta)d}$, see (2.10), into a Taylor series in ζ and then into a series in d:

$$\exp(-\Delta(\zeta)d) = \sum_{n=0}^{\infty} w_n(d)\zeta^n,$$

$$\exp(-\Delta(\zeta)d) = \sum_{n=0}^{\infty} \frac{(-\Delta(\zeta))^n}{n!} d^n.$$

Matching the powers of ζ we obtain the following expansion for $w_n(d)$, $n \geq 0$:

$$w_n(d) = \sum_{m=n}^{\infty} d^m f_m^n(A, b, h),$$

where $f_m^n(A, b, h)$ are matrix-valued functions of A, b and h.

Therefore, $w_n(0) = 0$, $n \geq 1$. The above also implies that convolution weights $w_n^h(d)$, $n \geq 1$, have a zero at $d = 0$ of order at least $n - 1$.

Let us recall the representation of the scaled convolution weights (3.1):

$$w_n(d) = \frac{1}{2\pi i} \oint_\gamma R(z)^{n-1}(I - Az)^{-1}\mathbb{1}b^T(I - Az)^{-1}e^{-zd}dz, \qquad (3.26)$$

where γ is a contour that encloses all the eigenvalues of A^{-1}, $n \geq 1$. From this it follows that

$$w_n(0) = \frac{1}{2\pi i} \oint_\gamma R(z)^{n-1}(I - Az)^{-1}\mathbb{1}b^T(I - Az)^{-1}dz = 0. \qquad (3.27)$$

Now let us prove the bounds (3.23, 3.24, 3.25) for $\frac{d}{h} < n$. Let $d \neq 0$. We express $e^{-z\frac{d}{h}}$ in terms of an integral of a parameter $0 \leq \rho \leq 1$:

$$e^{-z\frac{d}{h}} = 1 - \frac{zd}{h} \int_0^1 e^{-z\frac{d}{h}\rho}d\rho.$$

The definition (3.26) can then be rewritten:

$$w_n^h(d) = \frac{1}{2\pi i d} \oint_\gamma R(z)^{n-1}(I - Az)^{-1}\mathbb{1}b^T(I - Az)^{-1}e^{-z\frac{d}{h}}dz$$

$$= \frac{1}{2\pi i d} \oint_\gamma R(z)^{n-1}(I - Az)^{-1}\mathbb{1}b^T(I - Az)^{-1}dz$$

$$-\frac{1}{2\pi i h}\int\limits_0^1\oint\limits_\gamma R(z)^{n-1}(I-Az)^{-1}\mathbb{1}b^T(I-Az)^{-1}ze^{-z\frac{d}{h}\rho}dzd\rho.$$

The first term in the above sum equals 0, due to (3.27). The modulus of the second term, namely,

$$\frac{1}{2\pi i h}\int\limits_0^1\oint\limits_\gamma R(z)^{n-1}(I-Az)^{-1}\mathbb{1}b^T(I-Az)^{-1}ze^{-z\frac{d}{h}\rho}dzd\rho,$$

can be estimated using the mean value theorem. We first bound the value of the integral

$$I(\rho,d)=\frac{1}{2\pi i h}\oint\limits_\gamma R(z)^{n-1}(I-Az)^{-1}\mathbb{1}b^T(I-Az)^{-1}ze^{-z\frac{d}{h}\rho}dz$$

repeating the arguments of the proof of Proposition 3.3. Note that two changes have to be made. First, d has to be replaced with $\frac{d}{h}\rho$. Further, instead of bounding $\|(I-Az)^{-1}\mathbb{1}b^T(I-Az)^{-1}\|$ we now bound $\|(I-Az)^{-1}\mathbb{1}b^T(I-Az)^{-1}z\|$, for z lying on $\gamma=\gamma_r$, by a constant that does not depend on d, h or n, but does depend on the Runge–Kutta method.

It is not difficult to see that there exist positive constants G, C, $\bar\delta\in(0,\ 1)$, $q>0$ such that for $n\geq 1$ and $0<\delta<\bar\delta$ the following estimate holds:

$$|I(\rho,d)|\leq\frac{1}{h}G(1-\delta)^{n-\frac{d}{h}\rho}(1+C\delta^q)^{\frac{d}{h}\rho},$$

for $\frac{d}{h}\rho\leq n$. In the above expression q is either $p+1$ or $\frac{p+1}{2s-p}$, as in Proposition 3.3. Clearly this estimate is valid for all d such that $\frac{d}{h}<n$ and $\rho\in[0,1]$.

Next, we bound $\int\limits_0^1 I(\rho,d)d\rho$ as:

$$\left|\int\limits_0^1 I(\rho,d)d\rho\right|\leq\max_{\rho\in[0,\ 1]}|I(\rho,d)|$$

$$\leq\frac{1}{h}G(1-\delta)^n\max_{\rho\in[0,\ 1]}\left(\frac{1+C\delta^q}{1-\delta}\right)^{\frac{d}{h}\rho}$$

$$\leq\frac{1}{h}G(1-\delta)^{n-\frac{d}{h}}\left(1+C\delta^q\right)^{\frac{d}{h}}.$$

This finishes the proof of the statement. □

4 Computation of convolution weights

Let us write

$$\mathcal{K}_d(\zeta) = \frac{\exp\left(-\Delta(\zeta)\frac{d}{h}\right)}{4\pi d}.$$

The expansion (2.9) shows that $w_n^h(d)$ is the nth Taylor coefficient of $\mathcal{K}_d(\zeta)$. Therefore, the Cauchy integral formula gives another representation of $w_n^h(d)$,

$$w_n^h(d) = \frac{1}{2\pi i} \oint_{\mathscr{C}} \mathcal{K}_d(\zeta)\zeta^{-n-1}d\zeta.$$

Let us choose the contour $\mathscr{C}$ to be the circle centered at the origin with radius $\varrho < 1$. Discretizing this integral with the composite Trapezoid rule gives the approximation

$$w_n^h(d) = \varrho^{-n} \sum_{j=0}^{N} \mathcal{K}_d\left(\varrho e^{ij\frac{2\pi}{N+1}}\right) e^{-ijn\frac{2\pi}{N+1}} + O(\varrho^{N+1}), \quad n = 0, 1, \ldots, N. \tag{4.1}$$

In practice, the parameter $\varrho > 0$ cannot be chosen arbitrarily small in finite precision arithmetic. If eps denotes the machine precision the best accuracy that can be achieved is $\sqrt{\text{eps}}$ with the choice $\varrho = \text{eps}^{\frac{1}{2N}}$, see [23]. Using FFT, $w_n^h(d)$ can be computed in $O(N \log N)$ time for all $n = 0, 1, \ldots, N$. However, if d is bounded by Kh, for a constant $K > 0$, it is possible to avoid the scaling parameter ϱ as described in the next proposition.

Proposition 4.1 *Let w_n^h, $n \geq 0$, be the Runge–Kutta convolution weights (2.9), and let the Runge–Kutta method satisfy Assumption 2.1.*

Let $K > 0$ be fixed. There exist $\mu_1, \mu_2, \mu_3 > 0$, s.t. for all $\epsilon > 0, h > 0, L \in \mathbb{N}$ satisfying

$$L \geq \mu_1 \log \frac{1}{\epsilon h} + \mu_2 K + \mu_3,$$

and all $0 < d \leq Kh$ the following holds true.

1. *There exists an L-term approximation to the convolution kernel $\mathcal{K}_d(\xi) = \frac{\exp\left(-\Delta(\xi)\frac{d}{h}\right)}{4\pi d}$:*

$$\left| \mathcal{K}_d(\xi) - \sum_{\ell=0}^{L-1} w_\ell^h(d)\xi^\ell \right| \leq \epsilon$$

for all $\xi \in \mathbb{C} : |\xi| \leq 1$.

2. *Convolution weights can be approximated with an accuracy ϵ by an L-term discrete Fourier transform of the convolution kernel*

$$\left| w_n^h(d) - \frac{1}{L} \sum_{\ell=0}^{L-1} \mathcal{K}_d(e^{i\ell\frac{2\pi}{L}}) e^{-i\ell n \frac{2\pi}{L}} \right| \leq \epsilon$$

for all $n < L$.

Proof Let us prove the first statement using the bounds on convolution weights derived in Proposition 3.4. The second statement then follows directly from the first statement by the application of the aliasing formula.

By definition, for all $|\xi| < 1$

$$\mathcal{K}_d(\xi) = \sum_{\ell=0}^{\infty} w_\ell^h(d)\xi^\ell = \sum_{\ell=0}^{L-1} w_\ell^h(d)\xi^\ell + \sum_{\ell=L}^{\infty} w_\ell^h(d)\xi^\ell.$$

Let us show that given $\epsilon > 0$, there exists L s.t.

$$E_L(\xi, d) = \left\| \sum_{\ell=L}^{\infty} w_\ell^h(d)\xi^\ell \right\| \leq \epsilon. \tag{4.2}$$

First, let $L > K$. In a generalized form, the bounds on the scaled convolution weights $w_\ell^h(d)$ for $\ell \geq L$ and $d \leq Kh < \ell h$ can be stated as

$$\left\| w_\ell^h(d) \right\| \leq \frac{G}{h} (1-\delta)^{\ell - \frac{d}{h}} (1 + A\delta^\alpha)^{\frac{d}{h}},$$

for some $0 < \delta < \bar{\delta} < 1$ and $A, G, \alpha, \bar{\delta} > 0$ being constants. Then, after inserting this bound into the expression (4.2) for $E_L(\xi, d)$,

$$E_L(\xi, d) \leq \frac{G}{h} \left(\frac{1 + A\delta^\alpha}{1-\delta} \right)^{\frac{d}{h}} \sum_{\ell=L}^{\infty} (1-\delta)^\ell$$

$$\leq \frac{G}{h} \left(\frac{1 + A\delta^\alpha}{1-\delta} \right)^{\frac{d}{h}} (1-\delta)^L \delta^{-1} \leq \frac{G}{h\delta} \left(\frac{1 + A\delta^\alpha}{1-\delta} \right)^K (1-\delta)^L, \tag{4.3}$$

where we used that $d \leq Kh$. From the above it can be seen that L has to be chosen larger than K and so that

$$\frac{G}{h\delta} \left(\frac{1 + A\delta^\alpha}{1-\delta} \right)^K (1-\delta)^L < \epsilon.$$

Namely,

$$L \geq K \left(\log(1 + A\delta^\alpha) \log^{-1} \frac{1}{1-\delta} + 1 \right)$$
$$+ \log \frac{1}{h\epsilon} \log^{-1} \frac{1}{1-\delta} + \log \frac{G}{\delta} \log^{-1} \frac{1}{1-\delta}.$$

This proves the statement of the proposition for $|\xi| < 1$. For $|\xi| = 1$ the correctness of the statement can be seen from the bound (4.3) which is valid for $|\xi| = 1$.   □

Remark 4.1 Note that in the above statement the dependence of L on $\log \frac{1}{h}$ cannot be removed. To illustrate this fact, in Fig. 5 we plot $n_* = \sup\{n \in \mathbb{N} \mid \|w_\ell^h(d)\| < \epsilon$, for all $\ell \geq n$, $d \leq D\}$ for different values of h and ϵ and fixed $\frac{D}{h} = 10$, for the 3-stage Radau IIA method of order five.

A similar statement holds for the scaled convolution weights.

Proposition 4.2 *Let* w_n, $n \geq 0$, *be the scaled Runge–Kutta convolution weights* (2.10), *and let the Runge–Kutta method satisfy Assumption* 2.1.

Let $K > 0$ *be fixed. There exist* $\mu_1, \mu_2, \mu_3 > 0$, *s.t. for all* $\epsilon > 0$ *and* $L \in \mathbb{N}$ *satisfying*

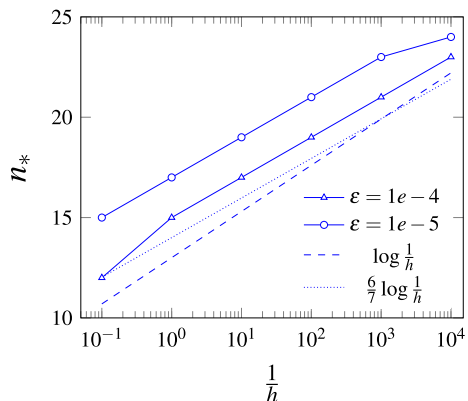$$L \geq \mu_1 \log \frac{1}{\epsilon} + \mu_2 K + \mu_3,$$

the following holds true for $0 \leq d \leq K$.

1. *There exists an* L-*term approximation to the convolution kernel* $K_d(\xi) = \exp(-\Delta(\xi)d)$:

$$\left| K_d(\xi) - \sum_{\ell=0}^{L-1} w_\ell(d)\xi^\ell \right| \leq \epsilon$$

for all $\xi \in \mathbb{C} : |\xi| \leq 1$.

Fig. 5 Dependence of $n_* = \sup\{n \in \mathbb{N} \mid \|w_\ell^h(d)\| < \epsilon$, for all $\ell \geq n$, $d \leq D\}$ on $\frac{1}{h}$, for $\epsilon = 1e - 4$, $1e - 5$ and fixed $\frac{D}{h} = 10$, for the 3-stage Radau IIA method

2. *Scaled convolution weights can be approximated with the accuracy ϵ by an L-term discrete Fourier transform of the convolution kernel $K_d(\xi)$*

$$\left| \mathrm{w}_n(d) - \frac{1}{L} \sum_{\ell=0}^{L-1} K_d(\xi)(\mathrm{e}^{i\ell\frac{2\pi}{L}})\mathrm{e}^{-i\ell n\frac{2\pi}{L}} \right| \le \epsilon$$

for all $n < L$.

Proof The proof of the proposition mimics the proof of Proposition 4.1.   □

5 Applications

The idea of using the sparsity of convolution weights to speed up calculations is not new. The most straightforward way of using the sparsity is to notice that for large enough n, the weights W_j^h, $j > n$, need not be computed at all but can be approximated by zero. In order to describe more advanced algorithms we need to briefly introduce the Galerkin boundary element discretization of convolution weights.

Let the boundary Γ be split into M disjoint panels $\tau_1, \ldots, \tau_M$ so that $\Gamma = \cup_i \bar\tau_i$. The span of piecewise constant basis functions

$$b_i(x) = \begin{cases} 1 & x \in \tau_i \\ 0 & x \notin \tau_i \end{cases}, \quad i = 1, 2, \ldots, M,$$

defines a finite dimensional subspace $X \subset H^{-1/2}(\Gamma)$.

The corresponding Galerkin discretization of a convolution weight W_n^h is given by

$$\begin{aligned}
\left(\mathbf{W}_n^h\right)_{ij} &= \int_\Gamma \left(W_n^h b_i\right)(x)b_j(x)d\Gamma_x \\
&= \int_\Gamma \int_\Gamma w_n^h(\|x - y\|)b_i(y)b_j(x)d\Gamma_y d\Gamma_x \\
&= \int_{\tau_i} \int_{\tau_j} w_n^h(\|x - y\|)d\Gamma_y d\Gamma_x.
\end{aligned}$$

Remark 5.1 Convolution weights W_n^h are $m \times m$-matrices of operators bounded as mappings from $H^{-1/2}(\Gamma)$ to $H^{1/2}(\Gamma)$. Therefore if

$$W_n^h = \begin{pmatrix} W_n^{11} & \cdots & W_n^{1m} \\ \vdots & \cdots & \vdots \\ W_n^{m1} & \cdots & W_n^{mm} \end{pmatrix},$$

then $\mathbf{W}_n^h$ should be understood as a matrix of matrices (a tensor) and

$$\left(\mathbf{W}_n^h\right)_{ij} = \begin{pmatrix} \int_\Gamma \left(W_n^{11} b_i\right)(x) b_j(x) d\Gamma_x & \cdots & \int_\Gamma \left(W_n^{1m} b_i\right)(x) b_j(x) d\Gamma_x \\ \vdots & \cdots & \vdots \\ \int_\Gamma \left(W_n^{m1} b_i\right)(x) b_j(x) d\Gamma_x & \cdots & \int_\Gamma \left(W_n^{mm} b_i\right)(x) b_j(x) d\Gamma_x \end{pmatrix}.$$

The sparsity of convolution weights shows that many entries in the Galerkin matrices need not be computed. This fact, though only for linear multistep based methods, has already been used in [17,18,21] to construct and analyse an efficient algorithm. In the next section we analyze the complexity of such an algorithm for Runge–Kutta convolution quadrature.

5.1 Sparse Runge–Kutta convolution quadrature

In this section, using the new estimates proved in this paper, we analyse the complexity of sparse Runge–Kutta convolution quadrature. We will throughout consider Runge–Kutta methods satisfying Assumptions 2.1 and 3.1.

The main idea behind the a priori cut-off strategy of [17,18,21] is to approximate $\mathbf{W}_n^h$, $n = 0, \ldots, N$ (for a fixed time interval $[0, T]$, $N = \lceil T/h \rceil$) by sparse matrices. For Runge–Kutta based methods we can use the results of Proposition 3.3 to decide what proportion of the entries in the matrices can be set to zero. More precisely, let i, j, n be s.t. for all $x \in \tau_i$, $y \in \tau_j$

$$\left\| w_n \left(\frac{\|x - y\|}{h} \right) \right\| < \epsilon,$$

where $\epsilon > 0$ is a given accuracy. Then, using $w_n^h(d) = \frac{w_n\left(\frac{d}{h}\right)}{4\pi d}$, we obtain

$$\left\| \left(\mathbf{W}_n^h\right)_{ij} \right\| = \left\| \int_{\tau_i} \int_{\tau_j} \frac{w_n\left(\frac{\|x-y\|}{h}\right)}{4\pi \|x - y\|} b_i(x) b_j(y) d\Gamma_x d\Gamma_y \right\|$$

$$\leq \int_{\tau_i} \int_{\tau_j} \frac{\epsilon}{4\pi \|x - y\|} \left| b_i(x) b_j(y) \right| d\Gamma_x d\Gamma_y \leq C_\Gamma \epsilon \|b_i\|_{L_2(\Gamma)} \|b_j\|_{L_2(\Gamma)}, \quad (5.1)$$

for some constant $C_\Gamma > 0$ that depends only on Γ. The last inequality was obtained from the L_2-continuity of the Laplace single layer boundary operator, see also [21].

Let us denote the spatial meshwidth by Δx and assume that

$$c\Delta x < \operatorname{diam} \tau_i < \Delta x, \quad i = 1, \ldots, M,$$

for some $c > 0$ independent of Δx. Therefore, there exists a positive $C' > 0$, s.t. $\left\| \left(\mathbf{W}_n^h\right)_{ij} \right\| \leq C' (\Delta x)^2 \epsilon$.

In [18] it is shown (for the BDF2 method) that the stability and convergence of the method based on the a priori cut-off strategy for BDF2 convolution quadrature is ensured if ϵ is chosen s.t.

$$\log \frac{1}{\epsilon} = O(\log M). \tag{5.2}$$

Similar analysis as in [18] shows that we can use the same lower bound for ϵ.

Additionally, let us assume that

$$\Delta x = O(h^\nu), \quad \nu > 0. \tag{5.3}$$

The case $\nu = 1$ is of particular interest in applications, e.g., scattering of waves of high frequency content where the time step and meshwidth are restricted more by the highest frequency present in the system than the required accuracy. Namely, given the highest frequency in the system f, the time step h is chosen as $h \approx Cf^{-1}$, for $C > 0$, and the meshwidth $\Delta x \approx Kh$, $K > 0$, see [14]. Under the assumption $M = O\left((\Delta x)^{-2}\right)$, this implies the following condition on N (the number of time steps) and M (the size of the spatial discretization):

$$M = O\left(N^{2\nu}\right). \tag{5.4}$$

To estimate the number of non-zero elements in $\left(\mathbf{W}_n^h\right)^{kl}$, $k, l = 1, \ldots, m$, $0 \le n \le N$, recall that for some $\alpha > 1$, $w_n\left(\frac{\|x-y\|}{h}\right)$ is negligible if

$$\left| \frac{\|x - y\|}{h} - n \right| \ge \max(\Delta_1^{n,\epsilon}, \Delta_2^{n,\epsilon}) \sim n^{1/\alpha} \log^{1-1/\alpha} \frac{1}{\epsilon}.$$

Hence the support of the kernel is widest for $n = N$. Therefore, for each of the M^2 pairs (i, j), $i, j = 1, \ldots, M$, the corresponding entry in the Galerkin matrix $\left(\mathbf{W}_n^h\right)_{ij}$ is non-zero for at most $O(N^{1/\alpha} \log^{1-1/\alpha} \frac{1}{\epsilon})$ convolution weights. Summing all this together we obtain

$$O(M^2 N^{1/\alpha} \log^{1-1/\alpha} \tfrac{1}{\epsilon}) = O(M^2 N^{1/\alpha} \log^{1-1/\alpha} M)$$

as an estimate of the number of non-zero entries in the sparse approximation of the N convolution weights.

For the 3-stage Radau IIA method $\alpha = p+1 = 6$. Taking $\nu = 1$, i.e., $M = O(N^2)$, we obtain as the total storage cost for N convolution weights

$$O(M^2 N^{1/6} \log^{5/6} M) = O(M^{2+1/12} \log^{5/6} M).$$

Hence, a straightforward application of sparsity does not give an algorithm of linear complexity. Nonetheless, the memory requirements are very close to those of many MOT and Galerkin methods, which scale as $O(M^2)$.

As a remedy, in [21] it was suggested to combine the a priori cutoff strategy with panel clustering. Although the extension of these ideas to Runge–Kutta CQ may lead to a method with improved memory requirements, the total complexity of such an algorithm would not scale better than $O(MN^2)$, due to the use of back substitution to solve the convolution system. This difficulty needs to be circumvented in order to allow for efficient large scale Runge–Kutta convolution quadrature algorithms. Improved complexity can be recovered if FFT methods are used. In the next section we describe briefly how FFT can be combined with the sparsity of the convolution weights to a good effect.

5.2 FFT and sparsity

The use of FFT, as in (4.1), usually destroys any sparsity, but in [7] it was shown that also in this case advantage can be made of sparsity. Here we will just briefly explain the main idea.

In (4.1) we have seen that the convolution kernels can be approximated by a discrete Fourier transform and hence the same is true of the convolution weights

$$\mathbf{W}_n^h \approx \varrho^{-n} \frac{1}{N+1} \sum_{\ell=0}^{N} \mathbf{V}(\varrho e^{i\ell \frac{2\pi}{N+1}}) e^{-i\ell n \frac{2\pi}{N+1}},$$

for $n = 0, 1, \ldots, N$. This shows that N convolution weights can be computed in $O(N \log N)$ time. Recall that for a fixed time interval $[0, T]$, $N = \lceil T/h \rceil$.

From Proposition 3.3 we know that for a fixed $n_0 > 0$

$$\|w_n^h(d)\| = \frac{1}{4\pi d} \left\| w_n\left(\frac{d}{h}\right) \right\| \leq \frac{C}{4\pi d}(1-\delta)^{n-\frac{d}{h}} \leq \frac{C}{4\pi d}(1-\delta)^{n-n_0},$$

for all $d \in (0, \ n_0 h)$ and $n > n_0$ with constant C. Hence there exists $n_1 \propto \log \varepsilon^{-1}$ such that

$$\|w_n^h(d)\| \leq \frac{\varepsilon}{4\pi d}, \quad \text{for all } 0 < d < n_0 h \text{ and } n > n_1.$$

Therefore for $n > n_1$, the "near-field" of the matrices $\mathbf{W}_n^h$ can be approximated by zero [c.f. (5.1)]:

$$\left(\mathbf{W}_n^h\right)_{ij} = \int_{\tau_i} \int_{\tau_j} w_n^h(\|x-y\|) d\Gamma_y d\Gamma_x \approx 0, \quad \text{if } \text{dist}(\tau_i, \tau_j) < n_0 h, \ n > n_1.$$

Unfortunately this is not true for the $N+1$ matrices $\mathbf{V}(\varrho e^{i\ell \frac{2\pi}{N+1}})$, further problem being that the near-field forms the part of the matrix which fast methods such as hierarchical matrices and the fast multipole method cannot be applied to. To considerably reduce the computational costs, we show now how to reduce the number of matrices for which

the near-field needs to be computed from $O(N)$ to $O\left(\log\frac{1}{\epsilon}\right)$ and still make effective use of the FFT. First we compute the n_1 weights $\mathbf{W}_n^h$, $n = 0, 1, \ldots, n_1$, using

$$\mathbf{W}_n^h \approx \frac{\varrho^{-n}}{n_1 + 1} \sum_{\ell=0}^{n_1} \mathbf{V}(\varrho e^{i\ell\frac{2\pi}{n_1+1}}) e^{-i\ell n_1 \frac{2\pi}{N+1}}, \quad n = 0, 1, \ldots, n_1.$$

This is not an expensive operation since n_1 depends only logarithmically on the desired accuracy ϵ, see Proposition 4.2 and (5.1). The remaining weights are computed using a similar formula as before in $O(N \log N)$ time

$$\mathbf{W}_n^h \approx \frac{\varrho^{-n}}{N + 1} \sum_{\ell=0}^{N} \widetilde{\mathbf{V}}(\varrho e^{i\ell\frac{2\pi}{N+1}}) e^{-i\ell n \frac{2\pi}{N+1}}, \quad n = n_1 + 1, \ldots, N,$$

where

$$(\widetilde{\mathbf{V}}(s))_{ij} = \begin{cases} 0 & \text{if } \operatorname{dist}(\tau_i, \tau_j) < n_0 h \\ (\widetilde{\mathbf{V}}(s))_{ij} & \text{otherwise.} \end{cases}$$

Thereby N convolution weights can still be computed in $O(N \log N)$ time, but only for a $O\left(\log\frac{1}{\epsilon}\right)$ number of evaluations of $\mathbf{V}(s)$ is it necessary to compute the near-field.

For a more thorough explanation of how to use such ideas in a full algorithm for *solving* the discretized equations see [7].

6 Conclusions

We have analysed the behaviour of convolution weights of Runge–Kutta convolution quadrature. We have proved that convolution weights $w_n^h(d)$ decay exponentially away from $nh \approx d$. The obtained estimates explain the dependence of the size of the approximate support of a convolution weight on the order of the underlying Runge–Kutta method. The results of this work can be used for the design and complexity analysis of fast algorithms based on sparsity for the solution of TDBIE for the three-dimensional wave equation.

Appendix: Proof of Lemma 3.1

Lemma 7.1 (Lemma 3.1) *For an A-stable Runge–Kutta method of order p there exist* $q, \nu > 0$, *such that the domain*

$$\{(x, y) \in \mathbb{R}^2 \ : \ |y| < \nu x^{\frac{1}{\ell}}, \ 0 < x < q\}$$

belongs to Υ_1 *(and intersects all the order star fingers). Here*

$$\ell = \begin{cases} p + 1, & \text{if } p \text{ is odd,} \\ 2s, & \text{if } p \text{ is even,} \end{cases}$$

where s is defined by (3.2) with $e_0 > 0$.

Proof Let us first remark that this proof is similar to the proof of Theorem 7 in [19].

We rewrite the stability function as

$$R(z) = e^z + C_{p+1}z^{p+1} + r(z), \quad C_{p+1} \neq 0, \tag{7.1}$$

where $r(z) = O(z^{p+2})$.

Let us consider (x, y) satisfying

$$|R(x + iy)| = 1.$$

Clearly, $R(0) = 1$ and $\frac{\partial |R(x+iy)|}{\partial x}\big|_{(0,0)} = 1$. By the implicit function theorem, there exists $\epsilon > 0$ and a unique continuously differentiable function $f : B_\epsilon(0) \to \mathbb{R}$, s.t. $f(0) = 0$, and $|R(f(y), y)| = 1$. Then, for $y \to 0$,

$$x = f(0) + f'(0)y + O(y^2) = O(y).$$

To prove the statement of the lemma, we explicitly study the behavior of the function $f(y)$ in the vicinity of 0. Let us consider the following cases.

1. p is odd.

$$|R(x + iy)|^2 = e^{2x} + 2C_{p+1} \operatorname{Re}\left(e^{x+iy}(x - iy)^{p+1}\right)$$
$$+ 2\operatorname{Re}\left(e^{x+iy}r(x - iy)\right) + C_{p+1}^2(x^2 + y^2)^{p+1} + |r(x + iy)|^2$$
$$+ 2C_{p+1} \operatorname{Re}\left(r(x + iy)(x - iy)^{p+1}\right). \tag{7.2}$$

Next we expand this expression into the Taylor series in x and y, singling out higher order terms while making use of $x = O(y)$:

$$e^{2x} = 1 + x + O(x^2),$$

$$\operatorname{Re}\left(e^{x+iy}(x - iy)^{p+1}\right) = \operatorname{Re}\sum_{k=0}^{p+1}\binom{p+1}{k}x^k i^{p+1-k}y^{p+1-k}(1 + r_1^{(2)}(x))$$
$$\times (1 + r_2^{(2)}(y)), \tag{7.3}$$

where $r_1^{(2)}(x) = e^x - 1 = O(x)$, $r_2^{(2)}(y) = O(y)$.

Since $p + 1$ is even, $\operatorname{Re}\left(e^{x+iy}(x - iy)^{p+1}\right) = (-1)^{\frac{p+1}{2}}y^{p+1} + O(y^{p+2})$. Finally, $\operatorname{Re}\left(e^{x+iy}r(x - iy)\right) = O(y^{p+2})$, which follows from the definition of $r(z) = O(z^{p+2})$, and

$$(x^2 + y^2)^{p+1} = O(y^{2p+2}),$$
$$|r(x + iy)|^2 = O(y^{2p+2}),$$
$$\operatorname{Re}\left(r(x + iy)(x - iy)^{p+1}\right) = O(y^{2p+2}).$$

Summarizing,

$$|R(x + iy)|^2 = 1 + 2x + 2C_{p+1}(-1)^{\frac{p+1}{2}} y^{p+1} + O(y^{p+2} + x^2).$$

Hence,

$$|R(x + iy)|^2 = 1 + y^{p+1}\left(2C_{p+1}(-1)^{\frac{p+1}{2}} + 2xy^{-p-1} + O(y) + O(x^2 y^{-p-1})\right). \quad (7.4)$$

We look for x, y lying in an ϵ-neighborhood of 0 and satisfying $|R(x + iy)|^2 = 1$. From the expression (7.4) we deduce that $x = f(y)$ with

$$x = -C_{p+1}(-1)^{\frac{p+1}{2}} y^{p+1} + O(y^{p+2}).$$

Note as well that (3.2) with $e_0 > 0$ implies that for sufficiently small y $|R(iy)|^2 < 1$ (recall as well the convention $Q(0) = 1$), hence from (7.1)

$$C_{p+1}(-1)^{\frac{p+1}{2}} < 0.$$

Hence there exist $\tilde{a}, \nu > 0$ s.t. $|R(x + iy)|^2 > 1$ for $\left\{(x, y) \in \mathbb{R}^2 : 0 < x < \tilde{a}, \; |y| < \nu x^{\frac{1}{p+1}}\right\}$.

2. *p is even.*

In this case we will make use of the properties of the E-polynomial (3.11), similarly to the approach taken in the proof of Theorem 7 in [19].

Let us define

$$\psi_y(x) = |R(x + iy)|^2.$$

For a fixed y we can expand the above expression into a Taylor series in x:

$$\psi_y(x) = |R(iy)|^2 + x\frac{d\psi_y}{dx}(0) + O(x^2). \quad (7.5)$$

Using (3.11), we can rewrite the first term:

$$|R(iy)|^2 = \frac{|P(iy)|^2}{|Q(iy)|^2} = 1 - \frac{E(y)}{|Q(iy)|^2},$$

which, after expansion into Taylor series (and using the convention $Q(0) = 1$), gives

$$|R(iy)|^2 = 1 - e_0 y^{2s} + O(y^{2s+2}).$$

Next, we need an expression for $\frac{d}{dx}\psi_y(0)$. From (7.2, 7.3), for odd p (using the same arguments as previously),

$$|R(x + iy)|^2 = 1 + 2x + 2C_{p+1}(p + 1)xy^p(-1)^{\frac{p}{2}} + O(xy^{p+1} + y^{p+2} + x^2)$$

Hence

$$\frac{d\psi_y}{dx}(0) = 2 + 2C_{p+1}(p+1)y^p(-1)^{\frac{p}{2}} + O(y^{p+1}).$$

Substituting the above into (7.5), we obtain

$$\begin{aligned}
\psi_y(x) &= 1 - e_0 y^{2s} + x(2 + 2C_{p+1}(p+1)y^p(-1)^{\frac{p}{2}} + O(y^{p+1})) \\
&\quad + O(y^{2s+2}) + O(x^2) \\
&= 1 + y^{2s}\left(-e_0 + 2xy^{-2s} + 2C_{p+1}(p+1)xy^{p-2s}(-1)^{\frac{p}{2}}\right. \\
&\quad \left. + O\left(xy^{p+1-2s} + x^2 y^{-2s}\right) + O(y^2)\right).
\end{aligned}$$

Recall that $e_0 > 0$. From the above expression we can see that $|R(x+iy)|^2 = 1$ in the vicinity of zero for (x, y):

$$x = e_0 y^{2s} + O(y^{2s+1}).$$

Hence there exist $\tilde{q}, \nu > 0$ s.t. $|R(x+iy)|^2 > 1$ for all

$$\left\{(x, \ y) \mid 0 < x < \tilde{q}, \ |y| < \nu x^{\frac{1}{2s}}\right\}.$$

Since the bounds derived are asymptotically optimal, this domain indeed intersects all the order star fingers. □

References

1. Alexander, R.: Diagonally Implicit Runge–Kutta Methods for Stiff O.D.E'.s. SIAM J. Numer. Anal. **14**(6), 1006–1021 (1977)
2. Ambroladze, A., Wallin, H.: Rational interpolants with preassigned poles, theoretical aspects. Stud. Math. **132**(1), 1–14 (1999)
3. Bamberger, A., Ha-Duong, T.: Formulation variationnelle espace-temps pour le calcul par potentiel retardé de la diffraction d'une onde acoustique. Math. Methods Appl. Sci. **8**(1), 405–435 (1986)
4. Bamberger, A., Ha-Duong, T.: Formulation variationnelle pour le calcul de la diffraction d'une onde acoustique par une surface rigide. Math. Methods Appl. Sci. **8**(1), 598–608 (1986)
5. Banjai, L.: Multistep and multistage convolution quadrature for the wave equation: algorithms and experiments. SIAM J. Sci. Comput. **32**(5), 2964–2994 (2010)
6. Banjai, L.: Time-domain Dirichlet-to-Neumann map and its discretization. IMA J. Numer. Anal. (2013). doi:10.1093/imanum/drt032
7. Banjai, L., Kachanovska, M.: Fast convolution quadrature for the wave equation in three dimensions (submitted)
8. Banjai, L., Lubich, C.: An error analysis of Runge–Kutta convolution quadrature. BIT Numer. Math. **51**, 483–496 (2011)
9. Banjai, L., Lubich, C., Melenk, J.: Runge–Kutta convolution quadrature for operators arising in wave propagation. Numer. Math. **119**(1), 1–20 (2011)
10. Costabel, M.: Time-dependent problems with the boundary integral equation method. In: Stein, E., de Borst, R., Hughes, T.J. (eds.) Encyclopedia of Computational Mechanics. Wiley, New York (2004)
11. Davies, P.J., Duncan, D.B.: Stability and convergence of collocation schemes for retarded potential integral equations. SIAM J. Numer. Anal. **42**(3), 1167–1188 (2004)

12. Ehle, B.L.: On Padé approximations to the exponential function and A-stable methods for the numerical solution of initial value problems. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)-University of Waterloo (Canada) (1969)

13. Ehle, B.L.: A-stable methods and Padé approximations to the exponential. SIAM J. Math. Anal. **4**, 671–680 (1973)

14. Ergin, A., Shanker, B., Michielssen, E.: The plane-wave time-domain algorithm for the fast analysis of transient wave phenomena. IEEE Antennas Propag. Mag. **41**(4), 39–52 (1999)

15. Ergin, A.A., Shanker, B., Michielssen, E.: Fast evaluation of three-dimensional transient wave fields using diagonal translation operators. J. Comput. Phys. **146**(1), 157–180 (1998)

16. Ha-Duong, T.: On retarded potential boundary integral equations and their discretization. In: Davies, P., Duncan, D., Martin, P., Rynne, B. (eds.) Topics in Computational Wave Propagation. Direct and Inverse Problems, pp. 301–336. Springer, Berlin (2003)

17. Hackbusch, W., Kress, W., Sauter, S.A.: Sparse convolution quadrature for time domain boundary integral formulations of the wave equation by cutoff and panel-clustering. In: Schanz, M., Steinbach, O. (eds.) Boundary Element Analysis, Lecture notes in Applied and Computational Mechanics, vol. 29, pp. 113–134. Springer, Berlin (2007)

18. Hackbusch, W., Kress, W., Sauter, S.A.: Sparse convolution quadrature for time domain boundary integral formulations of the wave equation. IMA J. Numer. Anal. **29**(1), 158–179 (2009)

19. Hairer, E., Bader, G., Lubich, C.: On the stability of semi-implicit methods for ordinary differential equations. BIT Numer. Math. **22**(2), 211–232 (1982)

20. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II. Springer Series in Computational Mathematics, vol. 14. Springer, Berlin (2010)

21. Kress, W., Sauter, S.: Numerical treatment of retarded boundary integral equations by sparse panel clustering. IMA J. Numer. Anal. **28**(1), 162–185 (2008)

22. Lubich, C.: Convolution quadrature and discretized operational calculus I. Numerische Mathematik **52**(2), 129–145 (1988)

23. Lubich, C.: Convolution quadrature and discretized operational calculus II. Numerische Mathematik **52**(4), 413–425 (1988)

24. Lubich, C.: On the multistep time discretization of linear initial-boundary value problems and their boundary integral equations. Numerische Mathematik **67**, 365–389 (1994)

25. Lubich, C., Ostermann, A.: Runge–Kutta methods for parabolic equations and convolution quadrature. Math. Comput. **60**(201), 105–131 (1993)

26. Wanner, G., Hairer, E., Nørsett, S.P.: Order stars and stability theorems. BIT Numer. Math. **18**, 475–489 (1978)