INTERNAL ERROR PROPAGATION IN EXPLICIT RUNGE–KUTTA METHODS[*]

DAVID I. KETCHESON[†], LAJOS LÓCZI[†], AND MATTEO PARSANI[‡]

Abstract. In practical computation with Runge–Kutta methods, the stage equations are not satisfied exactly, due to roundoff errors, algebraic solver errors, and so forth. We show by example that propagation of such errors within a single step can have catastrophic effects for otherwise practical and well-known methods. We perform a general analysis of internal error propagation, emphasizing that it depends significantly on how the method is implemented. We show that for a fixed method, essentially any set of internal stability polynomials can be obtained by modifying the implementation details. We provide bounds on the internal error amplification constants for some classes of methods with many stages, including strong stability preserving methods and extrapolation methods. These results are used to prove error bounds in the presence of roundoff or other internal errors.

Key words. Runge–Kutta methods, internal stability, roundoff error, strong stability preservation, extrapolation, ordinary differential equations

AMS subject classifications. 65L05, 65L06, 65L70, 65G50

DOI. 10.1137/130936245

1. Error propagation in Runge–Kutta methods. Runge–Kutta (RK) methods are used to approximate the solution of initial value ODEs:

$$(1.1) \qquad \begin{cases} U'(t) = F(t, U(t)), \\ U(t_0) = U_0, \end{cases}$$

often resulting from the semidiscretization of partial differential equations (PDEs). An s-stage RK method approximates the solution of (1.1) as follows:

$$(1.2a) \qquad Y_i = U_n + \tau \sum_{j=1}^{s} a_{ij} F(t_n + c_j \tau, Y_j) \qquad (1 \le i \le s),$$

$$(1.2b) \qquad U_{n+1} = U_n + \tau \sum_{j=1}^{s} b_j F(t_n + c_j \tau, Y_j).$$

Here U_n is a numerical approximation to $U(t_n)$, $\tau = t_{n+1} - t_n$ is the step size, and the stage values Y_j are approximations to the solution at times $t_n + c_j \tau$.

Most analysis of RK methods assumes that the stage equations (1.2a) are solved exactly. In practice, perturbed solution and stage values are computed:

$$(1.3a) \qquad \widetilde{Y}_i = \widetilde{U}_n + \tau \sum_{j=1}^{s} a_{ij} F(t_n + c_j\tau, \widetilde{Y}_j) + \tilde{r}_i \qquad (1 \le i \le s)$$

$$(1.3b) \qquad \widetilde{U}_{n+1} = \widetilde{U}_n + \tau \sum_{j=1}^{s} b_j F(t_n + c_j\tau, \widetilde{Y}_j) + \tilde{r}_{s+1}.$$

The *internal errors* (or stage residuals) $\tilde{r}_j$ include errors due to
- roundoff,
- finite accuracy of an iterative algebraic solver (for implicit methods).

The perturbed equations (1.3) are also used to study accuracy by taking $\widetilde{Y}$, $\widetilde{U}$ to be exact solution values to the ODE or PDE system, in which case the stage residuals include
- temporal truncation errors,
- spatial truncation errors,
- perturbations due to imposition of boundary conditions.

Such analysis is useful for explaining the phenomenon of *order reduction* due to stiffness [8] or imposition of boundary conditions [4, 1]. The theory of BSI-stability and B-convergence has been developed to understand these phenomena, and the relevant method property is usually the *stage order* [8].

The study of both kinds of residuals (due to roundoff or truncation errors) is referred to as *internal stability* [26, 24, 23, 22, 29]. We focus on the issue of amplification of roundoff errors in explicit RK (ERK) schemes, although we will see that some of our results and techniques are applicable to other internal stability issues. Since roundoff errors are generally much smaller than truncation errors, their propagation within a single step is not usually important. However for ERK methods with a large number of stages, the constants appearing in the propagation of internal errors can be so large that amplification of roundoff becomes an issue [27, 26, 28]. Amplification of roundoff errors in methods with many stages is increasingly important because there now exist several classes of practical RK methods that use many stages, including Runge–Kutta–Chebyshev (RKC) methods [29], extrapolation methods [10], deferred correction methods [5], some strong stability preserving (SSP) methods [9], and other stabilized ERK methods [19, 20]. Furthermore, these methods are naturally implemented not in the Butcher form (1.2) but in a *modified Shu–Osher form* [7, 12, 9]:

$$(1.4) \qquad Y_i = v_i U_n + \sum_{j=1}^{s} \left(\alpha_{ij} Y_j + \tau \beta_{ij} F(t_n + c_j\tau, Y_j) \right) \quad (1 \le i \le s+1),$$

$$U_{n+1} = Y_{s+1}.$$

As we will see, propagation of roundoff errors in these schemes should be based on the perturbed equations

$$(1.5) \qquad \widetilde{Y}_i = v_i \widetilde{U}_n + \sum_{j=1}^{s} \left(\alpha_{ij} \widetilde{Y}_j + \tau \beta_{ij} F(t_n + c_j\tau, \widetilde{Y}_j) \right) + \tilde{r}_i \qquad (1 \le i \le s+1),$$

$$\widetilde{U}_{n+1} = \widetilde{Y}_{s+1}$$

rather than on (1.3), because internal error propagation (in contrast to traditional error propagation) depends on the form used to implement the method. Through an

example in section 1.2 we will see that even when methods (1.2) and (1.4) are equivalent, the corresponding perturbed methods (1.3) and (1.5) may propagate internal errors in drastically different ways. Thus the residuals in (1.5) and in (1.3) will in general be different. In section 2, we elaborate on this difference and derive, for the first time, completely general expressions for the internal stability polynomials.

We emphasize here that the difference between (1.4) and (1.2) is distinct from the reordering of step sizes that was used to improve internal stability in [26]. Methods (1.4) and (1.2) can have different internal stability properties even when they are algebraically equivalent stage for stage.

In section 2.2, we introduce the *maximum internal amplification factor*, a simple characterization of how a method propagates internal errors. Although we follow tradition and use the term *internal stability*, it should be emphasized that this topic does not relate to stability in the usual sense, as there is no danger of unbounded blow-up of errors, only their substantial amplification. In this sense, the maximum internal amplification factor is similar to a condition number in that it is an upper bound on the factor by which errors may be amplified. In section 2.4 we show that for a fixed ERK method, essentially any set of internal stability polynomials can be obtained by modifying the implementation.

In sections 3 and 4, we analyze internal error propagation for SSP and extrapolation methods, respectively. Theorem 3.1 shows that SSP methods exhibit no internal error amplification when applied under the usual assumption of forward Euler contractivity. Additional results in these sections provide bounds on the internal amplification factor for general initial value problems. Much of our analysis follows along the lines of what was done in [29] for RKC methods. First we determine closed-form expressions for the stability polynomials and internal stability polynomials of these methods. Then we derive bounds and estimates for the maximum internal amplification factor. Using these bounds, we prove error bounds in the presence of roundoff error for whole families of methods where the number of stages may be arbitrarily large.

1.1. Preliminaries. In this subsection we define the basic setting and notation for our work. We consider the initial value problem (1.1), where $U : [t_0, T] \to \mathbb{R}^m$ and $F : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$. To shorten the notation, we will sometimes omit the first argument of F, writing $F(U)$ when there is no danger of confusion.

The RK method (1.2) and its properties are fully determined by the matrix $A = [a_{ij}] \in \mathbb{R}^{s \times s}$ and column vector $b = [b_j] \in \mathbb{R}^s$, which are referred to as the Butcher coefficients [3].

Let us define

(1.6)

$$
\alpha = \begin{pmatrix} \alpha_{1:s} & 0 \\ \alpha_{s+1} & 0 \end{pmatrix}, \qquad \alpha_{1:s} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1s} \\ \vdots & & \vdots \\ \alpha_{s1} & \cdots & \alpha_{ss} \end{pmatrix} \quad \alpha_{s+1} = (\alpha_{s+1,1}, \ldots, \alpha_{s+1,s}),
$$

$$
\beta = \begin{pmatrix} \beta_{1:s} & 0 \\ \beta_{s+1} & 0 \end{pmatrix}, \qquad \beta_{1:s} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1s} \\ \vdots & & \vdots \\ \beta_{s1} & \cdots & \beta_{ss} \end{pmatrix} \quad \beta_{s+1} = (\beta_{s+1,1}, \ldots, \beta_{s+1,s}),
$$

$$
v = \begin{pmatrix} v_{1:s} \\ v_{s+1} \end{pmatrix} = (I_{s+1} - \alpha)\mathbb{1}, \quad v_{1:s} = (v_1, \ldots, v_s)^\top,
$$

where $\mathbb{1}$ is the column vector of length $s+1$ with all entries equal to unity, and I_k is the $k \times k$ identity matrix. We always assume that $(I_s - \alpha_{1:s})^{-1}$ exists; methods without this property are not well defined [9]. The methods (1.2) and (1.4) are equivalent under the conditions

$$(1.7) \qquad A = (I_s - \alpha_{1:s})^{-1}\beta_{1:s},$$
$$b = \beta_{s+1} + \alpha_{s+1}(I_s - \alpha_{1:s})^{-1}\beta_{1:s}.$$

We assume that all methods satisfy the conditions for stage consistency of order one, i.e.,

$$(1.8a) \qquad v_i = 1 - \sum_j \alpha_{ij},$$

$$(1.8b) \qquad c = A\mathbb{1} = (I_s - \alpha_{1:s})^{-1}\beta_{1:s}\mathbb{1}.$$

Finally, define

$$(1.9) \qquad \mathbf{Y} = (Y_1, \ldots, Y_{s+1})^\top, \qquad \mathbf{F(Y)} = (F(Y_1), \ldots, F(Y_s), 0)^\top,$$
$$(1.10) \qquad \boldsymbol{\alpha} = \alpha \otimes I_m, \qquad \boldsymbol{\beta} = \beta \otimes I_m, \qquad \mathbf{v} = v \otimes I_m,$$

where $\otimes$ denotes the Kronecker product. Method (1.4) can also then be written

$$(1.11) \qquad \mathbf{Y} = \mathbf{v}U_n + \boldsymbol{\alpha}\mathbf{Y} + \tau\boldsymbol{\beta}\mathbf{F(Y)},$$
$$U_{n+1} = Y_{s+1}.$$

Recall that m denotes the dimension of U and s denotes the number of stages; boldface symbols are used for vectors and matrices with dimension(s) of size $m(s+1)$ whenever $m \geq 2$. When considering scalar problems ($m = 1$), we use nonbold symbols for simplicity.

When studying internal error amplification over a single step, we will sometimes omit the tilde over U_n to emphasize that we do not consider propagation of errors from previous steps.

Remark 1.1. The Butcher representation of an RK method is the particular Shu–Osher representation obtained by setting α_{ij} to zero for all i, j and setting

$$\beta = \beta_0 := \begin{pmatrix} A & 0 \\ b^\top & 0 \end{pmatrix}.$$

1.2. An example. Here we present an example demonstrating the effect of internal error amplification. We consider the following initial value problem (problem D2 of the nonstiff DETEST suite [13]), whose solution traces an ellipse with eccentricity 0.3:

$$(1.12a) \qquad x''(t) = -x/r^3, \qquad x(0) = 0.7, \qquad x'(0) = 0,$$
$$(1.12b) \qquad y''(t) = -y/r^3, \qquad y(0) = 0, \qquad y'(0) = \sqrt{13/7},$$
$$(1.12c) \qquad r^2 = x^2 + y^2.$$

We note that very similar results would be obtained with many other initial value problems. We first compute the solution at $t = 20$ using Fehlberg's 5(4) RK pair,
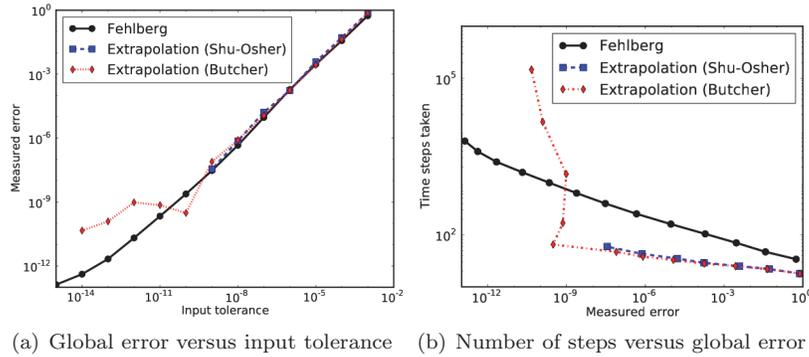
(a) Global error versus input tolerance   (b) Number of steps versus global error

FIG. 1. *Internal instability of a* 12th-*order extrapolation method applied to problem* (1.12).

which is not afflicted by any significant internal amplification of error. Results, shown in Figure 1, are typical and familiar to any student of numerical analysis. As the tolerance is decreased, the step size controller uses smaller steps and achieves smaller local—and global—errors, at the cost of an increased amount of work. Eventually, the truncation errors become so small that the accumulation of roundoff errors is dominant and the overall error cannot be decreased further.

Next we perform the same computation using a 12th-order extrapolation method based on the first-order explicit Euler method [15]; this method has 67 stages. The (embedded) error estimator for the extrapolation method is based on the 11th-order diagonal extrapolation entry. This pair is naturally implemented in a certain Shu–Osher form (see Algorithm 1 in section 4). The results, also shown in Figure 1, are similar to those of Fehlberg's method for large tolerances, although the number of steps required is much smaller for this 12th-order method. However, for tolerances less than 10^{-9}, the extrapolation method fails completely. The step size controller rejects every step and continually reduces the step size; the integration cannot be completed to the desired tolerance in a finite number of steps—even though that tolerance is six orders of magnitude larger than roundoff of 32 bit calculation!

Finally, we perform the same computation using an alternative implementation of the 12th-order extrapolation method. The Butcher form (1.2) is used for this implementation; it seems probable that no extrapolation method has ever previously been implemented in this (unnatural) way. The results are again shown in Figure 1; for large tolerances they are identical to the Shu–Osher implementation. For tolerances below 10^{-9}, the Butcher implementation is able to complete the integration, albeit using an excessively large number of steps, and with errors much larger than those achieved by Fehlberg's method at the same tolerance.

What is the cause for the surprising behavior of the extrapolation method? We will return to and explain this example after describing the relevant theory.

2. Internal error amplification.

2.1. Internal stability functions. Internal stability can be understood by considering the linear, scalar ($m = 1$) initial value problem

(2.1)
$$\begin{cases} U'(t) = \lambda U(t), & \lambda \in \mathbb{C}, \\ U(t_0) = U_0. \end{cases}$$

We will consider the application of an RK method in Shu–Osher form (1.11); recall that the Butcher form (1.2) is included as a special case. Application of the RK method (1.11) to the test problem (2.1) yields

$$(2.2a) \qquad\qquad Y = v_{1:s}U_n + \alpha_{1:s}Y + z\beta_{1:s}Y,$$
$$(2.2b) \qquad\qquad U_{n+1} = v_{s+1}U_n + \alpha_{s+1}Y + z\beta_{s+1}Y,$$

where $z = \tau\lambda$. By solving (2.2a) for Y and substituting into (2.2b) we obtain

$$(2.3) \qquad\qquad U_{n+1} = P(z)U_n,$$

where $P(z)$ is the *stability function*, which in Shu–Osher variables takes the form

$$(2.4) \qquad P(z) := v_{s+1} + (\alpha_{s+1} + z\beta_{s+1})\left(I_s - \alpha_{1:s} - z\beta_{1:s}\right)^{-1} v_{1:s}.$$

If the stage equations are satisfied exactly, then $P(z)$ completely determines the behavior of the numerical scheme for linear problems. However, it is known that RK methods with many stages may exhibit loss of accuracy even when $|P(z)| \leq 1$, due to propagation of errors within a single time step [27, 26, 28]. In order to investigate the internal stability of an RK method we apply the perturbed scheme (1.5), which for problem (2.1) yields

$$(2.5a) \qquad\qquad \widetilde{Y} = v_{1:s}\widetilde{U}_n + \alpha_{1:s}\widetilde{Y} + z\beta_{1:s}\widetilde{Y} + \tilde{r}_{1:s},$$
$$(2.5b) \qquad\qquad \widetilde{U}_{n+1} = v_{s+1}\widetilde{U}_n + \alpha_{s+1}\widetilde{Y} + z\beta_{s+1}\widetilde{Y} + \tilde{r}_{s+1}.$$

Let $\epsilon_n = \widetilde{U}_n - U_n$ and $d = [d_j] = [\widetilde{Y}_j - Y_j]$. Subtracting (2.2) from (2.5) gives

$$(2.6) \qquad\qquad d = v_{1:s}\epsilon_n + \alpha_{1:s}d + z\beta_{1:s}d + \tilde{r}_{1:s},$$
$$(2.7) \qquad\qquad \epsilon_{n+1} = v_{s+1}\epsilon_n + \alpha_{s+1}d + z\beta_{s+1}d + \tilde{r}_{s+1}.$$

By solving (2.6) for d and substituting the resulting expression in (2.7), one arrives at the error formula

$$(2.8) \qquad\qquad \epsilon_{n+1} = P(z)\epsilon_n + Q(z;\alpha,\beta)\tilde{r}_{1:s} + \tilde{r}_{s+1}.$$

The stability function $P(z)$ has already been defined in (2.4), and the *internal stability functions* $Q(z;\alpha,\beta)$ are

$$(2.9)$$
$$Q(z;\alpha,\beta) \equiv (Q_1(z), Q_2(z), \ldots, Q_s(z)) := (\alpha_{s+1} + z\beta_{s+1})\left(I_s - \alpha_{1:s} - z\beta_{1:s}\right)^{-1}.$$

Note that for convenience we have omitted the last component, $Q_{s+1}(z)$, which is always equal to 1. We will often suppress the explicit dependence of Q on α, β to keep the notation simple. Using (1.7), we can obtain the expression

$$(2.10) \qquad Q(z) = zb^\top(I - zA)^{-1}(I - \alpha_{1:s})^{-1} + \alpha_{s+1}(I - \alpha_{1:s})^{-1}.$$

We will refer to ϵ as "error," though its exact interpretation depends on what U and $\widetilde{U}$ refer to. If $\tilde{r}$ represents roundoff error, then (2.8) indicates the effect of roundoff on the overall solution. The one-step error is given by the sum of two terms: one governed by $P(z)$, accounting for propagation of errors committed in previous steps,

and one governed by $Q(z)$, accounting for propagation of the internal errors within the current step. In particular, $Q_j(z)$ governs the propagation of the perturbation $\tilde{r}_j$, appearing in stage j. Clearly we must have $|P(z)| \leq 1$ for stable propagation of errors, but if the magnitude of $|Q_j(z)|\tilde{r}_j$ is larger than the magnitude of the desired tolerance, then the second term is also important.

Note that $Q_j(z)$ herein is denoted by $Q_{sj}(z)$ in [29]. We are mostly interested in ERK methods, for which $Q_j(z)$ is a polynomial of degree at most $s - j$ and the first component of $\tilde{r}$ is zero, since no error is made in setting $Y_1 = U_n$.

Remark 2.1. For a method in Butcher form (i.e., with $\alpha = 0$), (2.10) reads

$$(2.11) \qquad Q^{\mathrm{B}}(z) := Q(z; 0, \beta_0) = zb^\top (I - zA)^{-1}.$$

Formulas (2.10) and (2.11) differ in an important way: we have $Q^{\mathrm{B}}(z) \to 0$ as $z \to 0$, so that the effects of internal errors vanish in the limit of infinitesimal step size. On the other hand, (2.10) does not have this property, so internal errors may still be amplified by a finite amount, no matter how small the step size is. As we will see, this explains the different behavior of the two extrapolation implementations in section 1.2.

2.1.1. Local defects. Equation (2.8) can also be used to study the discretization error. If we take $\widetilde{U}_n = U(t_n)$, $\widetilde{Y}_i = U(t_n + c_i\tau)$, $\widetilde{U}_{n+1} = U(t_{n+1})$, then ϵ_{n+1} is the global error and (2.8) describes how the stage errors contribute to it. We have

$$\tilde{r}_{1:s} = \sum_{k=0}^{p} \frac{\tau^k}{k!} \theta_k u_h^{(k)}(t_n) + \mathcal{O}(\tau^{p+1}),$$

where

$$(2.12) \qquad \theta_0 = \mathbb{1} - v - \alpha\mathbb{1},$$
$$\theta_k = \frac{1}{k!} (I_s - \alpha_{1:s}) (c^k - kAc^{k-1}) \qquad (1 \leq k \leq p).$$

Note that here $c \in \mathbb{R}^{s+1}$ with $c_{s+1} = 1$ and c^k denotes the vector with jth entry equal to c_j^k. Substituting the above into (2.8), we obtain

$$(2.13) \qquad \epsilon_{n+1} = P(z)\epsilon_n + \sum_{k=2}^{p} \frac{\tau^k}{k!} u_h^{(k)}(t_n) Q(z)\theta_k + \mathcal{O}(\tau^{p+1}).$$

For a method of order p, it can be shown that $\tau^k Q(z)\theta_k = \mathcal{O}(z^{p+1})$, where p is the classical order of the method, so the expected rate of convergence will be observed in the limit $z \to 0$. On the other hand, in problems arising from semidiscretization of a PDE, it often does not make sense to consider the limit $z \to 0$ but only the limit $\tau \to 0$. In that case, it can be shown only that $\tau^k Q(z)\theta_k = \mathcal{O}(\tau^{\tilde{p}+1})$, where $\tilde{p}$ is the *stage order* of the method; for all explicit methods, $\tilde{p} = 1$. This difference is responsible for the phenomenon of order reduction [23].

If the stage equations are solved exactly, then the stage and solution values computed using the Shu–Osher form (1.5) are identically equal to those computed using the Butcher form (1.3). By comparing the Butcher and Shu–Osher forms, one finds that in general the residuals are related by

$$\tilde{r}_{1:s}^{\mathrm{B}} = (I_s - \alpha_{1:s})^{-1} \tilde{r}_{1:s}^{\mathrm{SO}},$$
$$\tilde{r}_{s+1}^{\mathrm{B}} = \alpha_{s+1}(I_s - \alpha_{1:s})^{-1} \tilde{r}_{1:s}^{\mathrm{SO}} + \tilde{r}_{s+1}^{\mathrm{SO}}.$$

Here $\tilde{r}^{\mathrm{B}}$, $\tilde{r}^{\mathrm{SO}}$ denote the residuals in (1.3) and (1.5), respectively. Thus if the stage equations are solved exactly, the product of $Q(z)$ with the residuals is independent of the form used for implementation.

   If we wish to study the overall error in the presence of roundoff (i.e., the combined effect of discretization error and roundoff error), we may take U_n to be the solution given by the RK method in the presence of roundoff and $\widetilde{U}_n = U(t_n)$. This leads to

$$(2.14a) \qquad \epsilon_{n+1} = P(z)\epsilon_n + Q(z)\tilde{r} + \sum_{k=2}^{p} \frac{\tau^k}{k!} u_h^{(k)}(t_n) Q(z)\theta_k + \mathcal{O}(\tau^{p+1})$$

$$(2.14b) \qquad = P(z)\epsilon_n + Q(z)\tilde{r} + \mathcal{O}(\tau^{p+1})$$

for a method of order p, where now $\tilde{r}$ denotes only the roundoff errors. The effect of roundoff becomes significant when the last two terms in (2.14b) have similar magnitude.

   2.1.2. Internal stability polynomials and implementation: An example. A given RK method can be rewritten in infinitely many different Shu–Osher forms; this rewriting amounts to algebraic manipulation of the stage equations and has no effect on the method if one assumes the stage equations are solved exactly. However, the internal stability of a method depends on the particular Shu–Osher form used to implement it.

   For example, consider the two-stage, second-order optimal SSP method [9]. It is often written and implemented in the following modified Shu–Osher form:

$$(2.15a) \qquad\qquad\qquad Y_1 = U_n,$$

$$(2.15b) \qquad\qquad\qquad Y_2 = Y_1 + \tau F(Y_1),$$

$$(2.15c) \qquad\qquad U_{n+1} = Y_3 = \frac{1}{2}U_n + \frac{1}{2}(Y_2 + \tau F(Y_2)).$$

Applying (2.15) to the test problem (2.1) and introducing perturbations in the stages we have

$$\widetilde{Y}_1 = U_n,$$
$$\widetilde{Y}_2 = (1+z)U_n + \tilde{r}_2,$$
$$\widetilde{U}_{n+1} = \frac{1}{2}U_n + \frac{1}{2}(1+z)\widetilde{Y}_2 + \tilde{r}_3.$$

Substituting the equation for $\widetilde{Y}_2$ into that for $\widetilde{U}_{n+1}$ yields

$$\widetilde{U}_{n+1} = \left(1 + z + \frac{1}{2}z^2\right)U_n + \frac{1+z}{2}\tilde{r}_2 + \tilde{r}_3,$$

from which we can read off the stability polynomial $P(z) = 1 + z + z^2/2$ and the second-stage internal stability polynomial

$$Q_2(z) = (1+z)/2.$$

However, in the Butcher form, the equation for U_{n+1} is written as

$$U_{n+1} = U_n + \frac{1}{2}\tau(F(Y_1) + F(Y_2)),$$

which leads by a similar analysis to

$$Q_2(z) = z/2.$$

More generally, the equation for U_{n+1} can be written

$$U_{n+1} = \left(\frac{1}{2} + \beta_{31}\right) U_n + \left(\frac{1}{2} - \beta_{31}\right) Y_2 + \beta_{31}\tau F(Y_1) + \frac{1}{2}\tau F(Y_2)$$

with an arbitrary parameter $\beta_{31} \in \mathbb{R}$. By choosing a large value of β_{31}, the internal stability of the implementation can be made arbitrarily poor. For this simple method it is easy to see what a reasonable choice of implementation is, but for methods with very many stages it is far from obvious. In section 2.4 we study this further.

Note that the stability polynomial $P(z)$ is independent of the choice of Shu–Osher form.

2.2. Bounds on the amplification of internal errors. We are interested in bounding the amount by which the residuals $\tilde{r}$ may be amplified within one step, under the assumption that the overall error propagation is stable. In the remainder of this section, we introduce some basic definitions and straightforward results that are useful in obtaining such bounds. It is typical to perform such analysis in the context of the autonomous linear *system* of ODEs [29]:

$$(2.16) \qquad \begin{cases} U'(t) = LU(t) + g(t), \\ U(t_0) = U_0, \end{cases}$$

where $L \in \mathbb{R}^{m \times m}$, $g : \mathbb{R} \to \mathbb{R}^m$. Results based on such analysis are typically useful in the context of more complicated problems, whereas analyzing nonlinear problems directly does not usually yield further insight [26, 22, 29, 4].

Application of the perturbed RK method (1.5) to problem (2.16) leads to (2.14b) but with $z = \tau L$. Taking norms of both sides one may obtain

$$(2.17) \qquad \|\epsilon_{n+1}\| \leq \|P(z)\|\|\epsilon_n\| + \sum_{j=1}^{s+1} \|Q_j(z)\|\|\tilde{r}_j\| + \mathcal{O}(\tau^{p+1}).$$

It is thus natural to introduce the following.

DEFINITION 2.2 (maximum internal amplification factor). *The* maximum internal amplification factor *of an s-stage RK method* (1.4) *with respect to a set* $\mathcal{S} \subset \mathbb{C}$ *is*

$$(2.18) \qquad \mathcal{M} \equiv \mathcal{M}(\alpha, \beta, \mathcal{S}) := \max_{j=1,2,\dots,s} \sup_{z \in \mathcal{S}} |Q_j(z)|,$$

where $Q_j(z)$ is defined in (2.10). *When the set $\mathcal{S}$ is not specified, it is taken to be the absolute stability region of the method $\mathcal{S} = \{z \in \mathbb{C} : |P(z)| \leq 1\}$ with $P(z)$ given by* (2.4).

In order to control numerical errors, the usual strategy is to reduce the step size. To understand the behavior of the error for very small step sizes, it is therefore useful to consider the value

$$\mathcal{M}_0 := \mathcal{M}(\alpha, \beta, \{0\}) = \max_{j=1,2,\dots,s} |Q_j(0)|.$$

To go further, we need to make an assumption about $\sigma(L)$, the spectrum of L. The next theorem provides bounds on the error in the presence of roundoff or

iterative solution errors, when L is normal. Similar results could be obtained when L is nonnormal by considering pseudospectra.

THEOREM 2.3. *Let an RK method of order p with coefficients α, β be applied to (2.16), where L is a normal matrix and $\tau\sigma(L) \in \mathcal{S}$. Let $\epsilon_n = \widetilde{U}_n - U(t_n)$, where $\widetilde{U}_n$ satisfies the perturbed RK equations (1.5). Then*

$$(2.19)\qquad \|\epsilon_{n+1}\| \leq \|\epsilon_n\| + (s+1)\mathcal{M}(\alpha,\beta,\mathcal{S}) \max_{1\leq j\leq s+1} \|\tilde{r}_j\| + \mathcal{O}(\tau^{p+1}).$$

The factor $s+1$ can be replaced by s for explicit methods, since then $\tilde{r}_1 = 0$.

Proof. Use (2.17) and the fact that since L is normal, $\|Q_j(\tau L)\| = \max_{\lambda\in\sigma(L)} |Q_j(\tau\lambda)|$. ◻

Remark 2.4. Combining the above with analysis along the lines of [29], one obtains error estimates for application to PDE semidiscretizations. The amplification of spatial truncation errors does not depend on the implementation, as can be seen by working out the product $Q(Z)\beta$.

As an example, in Table 1 we list approximate maximum internal amplification factors for some common RK methods. All these methods, like most RK methods, have relatively small factors so that their internal stability is generally not a concern.

2.3. Understanding the example from section 1.2. Using the theory of the last section, we can fully explain the results of section 1.2. First, in Table 2, we give the values of $\mathcal{M}$ and $\mathcal{M}_0$ for the three methods. Observe that Fehlberg's method, like most methods with few stages, has very small amplification constants. Meanwhile, the Euler extrapolation method, with 67 stages, has a very large $\mathcal{M}$. However, when implemented in Butcher form, it necessarily has $\mathcal{M}_0 = 0$.

For the extrapolation method in Shu–Osher form, the local error will generally be at least $\mathcal{M}_0 \cdot \epsilon_{\text{machine}} \approx 10^{-10}$ for double precision calculations. Therefore, the pair will fail when the requested tolerance is below this level, which is just what we observe.

TABLE 1
Approximate maximum internal amplification factors for some RK methods. For RK pairs, the amplification factor of the higher-order method is listed.

Method	$\mathcal{M}$	$\mathcal{M}_0$
Three-stage, third-order SSP [25]	1.7	0
Three-stage, third-order Heun [11]	3.2	0
Classic four-stage, fourth-order method [17]	1.7	0
Merson 4(3) pair [18]	5.6	0
Fehlberg 5(4) pair [6]	5.4	0
Bogacki–Shampine 5(4) pair [2]	7.0	0
Prince–Dormand order 8 [21]	138.8	0
10-stage, fourth-order SSP [14]	2.4	0.6
10-stage, first-order RKC [29]	10.0	10.0
18-stage, second-order RKC [29]	27.8	22.6

TABLE 2
Maximum internal amplification factors for methods used in the example in section 1.2.

Method	$\mathcal{M}$	$\mathcal{M}_0$
Fehlberg 5(4)	5.4	0
Euler extrapolation 12(11) (Shu–Osher)	$3.4 \cdot 10^5$	$1.3 \cdot 10^5$
Euler extrapolation 12(11) (Butcher)	$1.7 \cdot 10^5$	0

The extrapolation method in Butcher form also begins to be afflicted by amplification of roundoff errors at this point (since it has a similar value of $\mathcal{M}$). Reducing the step size has the effect of reducing the amplification constant, but only at a linear rate (i.e., in order to reduce the local error by a factor of 10, roughly 10 times as many steps are required). This is observed in Figure 1(b). Meanwhile, the global error does not decrease at all, since the number of steps taken (hence the number of roundoff errors committed) is increasing just as fast as the local error is decreasing. This is observed in Figure 1(a).

2.4. Effect of implementation on internal stability polynomials. In section 2.1.2, we gave an example showing how the internal stability functions of a method could be modified by the choice of Shu–Osher implementation. It is natural to ask just how much control over these functions is possible.

Given an s-stage ERK method, let $Q_1^{\mathrm{B}}, Q_2^{\mathrm{B}}, \ldots, Q_s^{\mathrm{B}}$ denote the internal stability polynomials corresponding to the Butcher form implementation. Let d_j denote the degree of Q_j^{B} and let w_j denote the coefficient of z^{d_j} in Q_j^{B}. Let $Q_1^{\mathrm{SO}}, Q_2^{\mathrm{SO}}, \ldots, Q_s^{\mathrm{SO}}$ be any set of polynomials with the same degrees d_j and leading coefficients w_j, but otherwise arbitrary. Then it is typically possible to find a Shu–Osher implementation of the given method such that the internal stability polynomials are $Q_1^{\mathrm{SO}}, Q_2^{\mathrm{SO}}, \ldots, Q_s^{\mathrm{SO}}$.

How can this be done? Comparing (2.10) and (2.11), it is clear that, except for the constant terms, the internal stability polynomials corresponding to a given Shu–Osher implementation are linear combinations of the internal stability polynomials of the Butcher implementation:

$$(2.20) \qquad Q^{\mathrm{SO}}(z) = Q^{\mathrm{B}}(z)\gamma + \alpha_{s+1}\gamma.$$

Here $\gamma = (I - \alpha_{1:s})^{-1}$ can be any lower-triangular matrix with unit entries on the main diagonal. Given $Q^{\mathrm{B}}(z)$, one can choose γ to obtain the desired polynomials $Q^{\mathrm{SO}}(z)$ except for the constant terms and then choose α_{s+1} to obtain the desired constant terms. We have added the qualifier *typically* above because it is necessary that successive subsets of the Q_j^{B} span the appropriate polynomial spaces.

3. Internal amplification factors for explicit strong stability preserving methods. In this section we prove bounds on and estimates of the internal amplification factors for optimal explicit SSP RK methods of orders 2 and 3. These methods have extraordinarily good internal stability properties.

We begin with a result showing that no error amplification occurs when an SSP method is applied to a contractive problem. Recall that an SSP RK method with SSP coefficient $\mathcal{C}$ can be written in the *canonical Shu–Osher form*, where $\alpha = \mathcal{C}\beta$, $\alpha_{ij} \geq 0$ for all i, j, and the row sums of α are no greater than 1 [9].

THEOREM 3.1. *Suppose that F is contractive with respect to some seminorm $\|\cdot\|$:*

$$(3.1) \quad \|U + \tau F(U) - (\widetilde{U} + \tau F(\widetilde{U}))\| \leq \|U - \widetilde{U}\| \quad \text{for all } U, \widetilde{U} \text{ and for } 0 \leq \tau \leq \tau_0.$$

Let an ERK method with SSP coefficient $\mathcal{C} > 0$ be given in canonical Shu–Osher form α, β and applied with step size $\tau \leq \mathcal{C}\tau_0$. Then

$$(3.2) \qquad \|\epsilon_{n+1}\| \leq \|\epsilon_n\| + \sum_{j=1}^{s+1} \|\tilde{r}_j\|.$$

Proof. The proof proceeds along the usual lines of results for SSP methods. From (1.4) and (1.5), we have

$$Y_i = v_i U_n + \sum_{j=1}^{i-1} \alpha_{ij} \left(Y_j + \frac{\tau}{\mathcal{C}} F(Y_j) \right),$$

$$\widetilde{Y}_i = v_i \widetilde{U}_n + \sum_{j=1}^{i-1} \alpha_{ij} \left(\widetilde{Y}_j + \frac{\tau}{\mathcal{C}} F(\widetilde{Y}_j) \right) + \tilde{r}_i.$$

Taking the difference of these two equations, we have

$$d_i = v_i \epsilon_n + \sum_{j=1}^{i-1} \alpha_{ij} \left(d_j + \frac{\tau}{\mathcal{C}} \left(F(\widetilde{Y}_j) - F(Y_j) \right) \right) + \tilde{r}_i,$$

where d and ϵ are defined in section 2.1. Applying $\| \cdot \|$ and using convexity, (3.1), and (1.8a), we find

$$\|d_i\| \leq v_i \|\epsilon_n\| + \sum_{j=1}^{i-1} \alpha_{ij} \left\| d_j + \frac{\tau}{\mathcal{C}} \left(F(\widetilde{Y}_j) - F(Y_j) \right) \right\| + \|\tilde{r}_i\|$$

$$\leq v_i \|\epsilon_n\| + \sum_{j=1}^{i-1} \alpha_{ij} \|d_j\| + \|\tilde{r}_i\|$$

$$\leq \max \left(\|\epsilon_n\|, \max_{1 \leq j \leq i-1} \|d_j\| \right) + \|\tilde{r}_i\|.$$

Applying the last inequality successively for $i = 1, 2, \ldots, s + 1$ shows that $\|d_i\| \leq \|\epsilon_n\| + \sum_{j=1}^{i} \|\tilde{r}_j\|$. In particular, for $i = s + 1$ this gives (3.2). □

The above result is the only one in this work that applies directly to nonlinear problems. It is useful since it can be applied under the same assumptions that make SSP methods useful in general. On the other hand, SSP methods are very often applied under circumstances in which the contractivity assumption is not justified— for instance, in the time integration of WENO semidiscretizations. In the remainder of this section, we develop bounds on the amplification factor for some common SSP methods. Such bounds can be applied even when the contractivity condition is not fulfilled.

3.1. Optimal second-order SSP methods. Here we study the family of optimal second-order SSP RK methods [9], corresponding to the most natural implementation. For any number of stages $s \geq 2$, the optimal method is

(3.3) $$Y_1 = U_n,$$

$$Y_j = Y_{j-1} + \frac{\tau}{s-1} F(Y_{j-1}), \qquad\qquad 2 \leq j \leq s,$$

$$U_{n+1} = \frac{1}{s} U_n + \frac{s-1}{s} \left(Y_s + \frac{\tau}{s-1} F(Y_s) \right).$$

THEOREM 3.2. *The maximum internal amplification factor for the method* (3.3) *satisfies*

$$\mathcal{M}_s^{\text{SSP2}} \leq \frac{s+1}{s}.$$

Proof. It is convenient to define

$$(3.4) \qquad \nu(z) = 1 + \frac{z}{s-1}.$$

Then the stability and internal stability functions are

$$P(z) = \frac{1}{s} + \frac{s-1}{s}\nu(z)^s, \qquad Q_j(z) = \frac{s-1}{s}\nu(z)^{s-j+1} \quad (2 \le j \le s).$$

(For brevity, we omit the details of the derivation here, but we will give a detailed proof of the analogous SSP3 case in Lemma 3.3.) We have

$$\mathcal{S}_s = \left\{ z \in \mathbb{C} : \left| \nu(z)^s + \frac{1}{s-1} \right| \le \frac{s}{s-1} \right\}.$$

Let $z \in \mathcal{S}_s$ be given; we will show that $|Q_j(z)| \le \frac{s+1}{s}$ for each j. We see that $\nu(z)^s$ lies in a disk of radius $s/(s-1)$ centered at $-1/(s-1)$. Hence

$$\mathcal{S}_s \subset \mathcal{S}'_s := \left\{ z \in \mathbb{C} : |\nu(z)^s| \le \frac{s+1}{s-1} \right\}.$$

For $|\nu| \le 1$ the desired result is immediate. For $|\nu| > 1$, we have

$$\max_{j=2,\ldots,s} \sup_{z \in \mathcal{S}_s} |Q_j(z)| = \max_{j=2,\ldots,s} \sup_{z \in \mathcal{S}_s} \frac{s-1}{s} |\nu(z)|^{s-j+1}$$

$$\le \max_{j=1,\ldots,s} \sup_{z \in \mathcal{S}'_s} \frac{s-1}{s} |\nu(z)|^{s-j+1}$$

$$\le \sup_{z \in \mathcal{S}'_s} \frac{s-1}{s} |\nu(z)|^s \le \frac{s-1}{s} \cdot \frac{s+1}{s-1} = \frac{s+1}{s}. \qquad \square$$

Some scaled and shifted stability regions are depicted in Figure 2.

Combining the above result with Theorem 2.3, one obtains full error estimates for application to linear systems and linear PDEs.
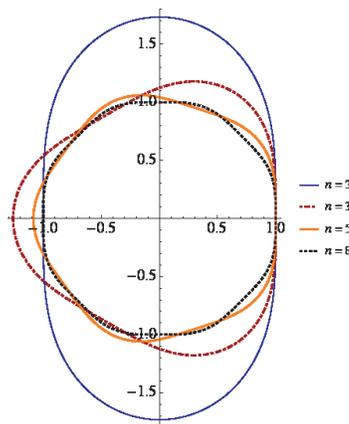

FIG. 2. *Scaled and shifted stability regions (in the sense of* (3.4)*) of the optimal second-order SSP RK methods for $s \in \{2,3,5,8\}$. The boundary curves on the plot are given by $\{\nu \in \mathbb{C} : \left| \frac{s-1}{s}\nu^s + \frac{1}{s} \right| = 1\}$. For any s, the corresponding region has s-fold rotational symmetry.*

3.2. Optimal third-order SSP methods. In this section we give various upper and lower bounds on the maximum internal amplification factor for the optimal third-order SSP RK methods with $s \equiv s_n := n^2$ ($n \geq 2$) stages that were proposed in [14]. The optimal third-order SSP RK method with $s = n^2$ stages can be written as follows:

$$(3.5) \qquad Y_1 = U_n,$$

$$Y_j = Y_{j-1} + \frac{\tau}{n^2 - n} F(Y_{j-1}), \qquad 2 \leq j \leq s, \, j \neq k_n,$$

$$Y_{k_n} = \frac{n-1}{2n-1} Y_{k_n-1} + \frac{n}{2n-1} Y_{m_n} + \frac{\tau}{n(2n-1)} F(Y_{k_n-1}),$$

$$U_{n+1} = Y_{s+1} = Y_s + \frac{\tau}{n^2 - n} F(Y_s),$$

where $n \geq 2$, $k_n := \frac{n(n+1)}{2} + 1$ and $m_n := \frac{(n-1)(n-2)}{2} + 1$.

Just as in the proof of Theorem 3.2, it is convenient to apply some normalization and introduce the scaling and shift

$$(3.6) \qquad \nu_n(z) := 1 + \frac{z}{n^2 - n} \qquad (n \geq 2, \, z \in \mathbb{C}).$$

The following lemma (with a proof in [16]) shows that the stability polynomial and the internal stability polynomials of this method can simply be expressed in terms of ν_n.

LEMMA 3.3. *For any $n \geq 2$, the stability function of the optimal third-order SSP RK method with $s = n^2$ stages is*

$$P(z) = \frac{n-1}{2n-1} \nu_n(z)^{n^2} + \frac{n}{2n-1} \nu_n(z)^{(n-1)^2},$$

while the internal stability functions are

$$Q_j(z) = \begin{cases} \dfrac{n-1}{2n-1} \nu_n(z)^{n^2-j+1} + \dfrac{n}{2n-1} \nu_n(z)^{(n-1)^2-j+1}, & 2 \leq j \leq m_n, \\[2mm] \dfrac{n-1}{2n-1} \nu_n(z)^{n^2-j+1}, & m_n + 1 \leq j \leq k_n - 1, \\[2mm] \nu_n(z)^{n^2-j+1}, & k_n \leq j \leq n^2. \end{cases}$$

This means that $\mathcal{M}_s^{\mathrm{SSP3}} \equiv \mathcal{M}_s^{\mathrm{SSP3}}(\mathcal{S}_s) = \max_{j=2,\dots,s} \sup_{z \in \mathcal{S}_s} |Q_j(z)|$ with $s = n^2$ ($n \geq 2$) and

$$\mathcal{S}_s = \left\{ z \in \mathbb{C} : \left| \frac{n-1}{2n-1} \nu_n(z)^{n^2} + \frac{n}{2n-1} \nu_n(z)^{(n-1)^2} \right| \leq 1 \right\}.$$

First, the triangle inequality shows that $\{z \in \mathbb{C} : |\nu_n(z)| \leq 1\} \subset \mathcal{S}_s$; the set $\{z \in \mathbb{C} : |\nu_n(z)| \leq 1\}$ appears as the unit disk in Figure 3. By taking into account the explicit forms of the Q_j polynomials provided by Lemma 3.3, we see, again by the triangle inequality, that $|Q_j(z)| \leq 1$ for all $z \in \mathbb{C}$ with $|\nu_n(z)| \leq 1$. The following trivial corollary is established.
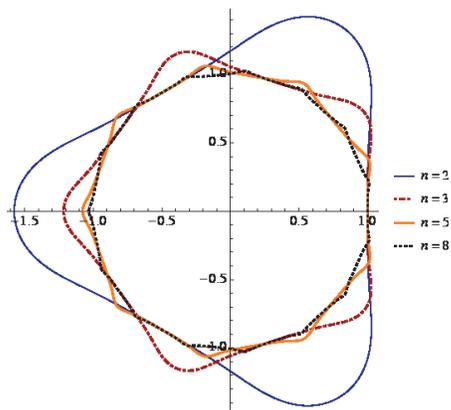
FIG. 3. *Scaled and shifted stability regions (see (3.6)) of the optimal third-order SSP RK methods with n^2 stages for $n \in \{2, 3, 5, 8\}$. The boundary curves on the plot are given by $\{\nu \in \mathbb{C} : \left| \frac{n-1}{2n-1} \nu^{n^2} + \frac{n}{2n-1} \nu^{(n-1)^2} \right| = 1\}$. For any n, the corresponding region has $(2n-1)$-fold rotational symmetry.*

TABLE 3
Maximum internal amplification factors for the first few optimal third-order SSP RK methods with $s = n^2$ stages. The values in the table have been rounded up.

n	$\mathcal{M}_s^{\mathrm{SSP3}}$	n	$\mathcal{M}_s^{\mathrm{SSP3}}$
2	1.575	7	2.314
3	1.794	8	2.411
4	1.956	9	2.501
5	2.091	10	2.585
6	2.209		

COROLLARY 3.4. *We have*

$$\mathcal{M}_s^{\mathrm{SSP3}}(\alpha, \beta, D_{\mathcal{C}}) \leq 1,$$

where $\mathcal{C} = n^2 - n$ and $D_{\mathcal{C}}$ is the disk $\{z \in \mathbb{C} : |z + \mathcal{C}| \leq \mathcal{C}\}$.

As a consequence, in order to obtain estimates of $\mathcal{M}_s^{\mathrm{SSP3}}(\mathcal{S}_s)$, it is enough to bound $|Q_j|$ only on the "petals" in Figure 3, that is, for $z \in \mathcal{S}_s$ with $|\nu_n(z)| \geq 1$. In the rest of this section, we just summarize our results: most of the proofs are technical; hence, due to space limitations, they are given in [16].

For any $\varrho \geq 1$ and $n \geq 2$ we define

$$\mu_n^-(\varrho) := -1 - \frac{n \varrho^{(n-1)^2} \left(1 - \left(1 - \frac{1}{n} \right) \varrho^{2n-1} \right)}{2n - 1}$$

and denote by ν_n^* the unique root of μ_n^- in the interval $[1, +\infty)$. Then we prove that

$$\mathcal{M}_s^{\mathrm{SSP3}} = (\nu_n^*)^{(n^2 - n)/2} \qquad (s = n^2, \ n \geq 2).$$

By computing ν_n^* numerically, $\mathcal{M}_s^{\mathrm{SSP3}}$ can easily be determined for small values of n; see Table 3.

For larger n values the following inequalities hold.

THEOREM 3.5. *For any $n \geq 9$ and with $s = n^2$, the maximum internal amplification factor for the optimal s-stage, third-order SSP RK method satisfies*

$$\frac{9}{10}\sqrt{\frac{n}{\ln(n)}} < \left(1 + \frac{\ln(n)}{n^2} - \frac{\ln(\ln(n))}{n^2}\right)^{(n^2-n)/2} < \mathcal{M}_s^{\text{SSP3}}$$

$$< \left(1 + \frac{\ln(n)}{n^2} - \frac{\ln(\ln(n))}{8n^2}\right)^{(n^2-n)/2} < \frac{\sqrt{n}}{\sqrt[16]{\ln(n)}}.$$

Remark 3.6. As a trivial consequence, for $s = n^2$ with $n \geq 4$ we see that

$$\mathcal{M}_s^{\text{SSP3}} < \sqrt[4]{s} = \sqrt{n}.$$

We illustrate the chain of inequalities in Theorem 3.5 for $s = 10^2$, $s = 10^4$, and $s = 10^{12}$:

$$1.875 < 1.927 < \mathcal{M}_{10^2}^{\text{SSP3}} \approx 2.585 < 2.661 < 3.002,$$

$$4.193 < 4.587 < \mathcal{M}_{10^4}^{\text{SSP3}} \approx 5.757 < 8.887 < 9.090,$$

$$242.135 < 269.038 < \mathcal{M}_{10^{12}}^{\text{SSP3}} \approx 302.551 < 848.642 < 848.647.$$

Moreover, if we *assume* that the asymptotic expansion of ν_n^* starts as

$$\nu_n^* \sim 1 + \frac{\ln(n)}{n^2} - \frac{\ln(\ln(n))}{n^2} + \frac{\ln(\ln(n))}{n^2\ln(n)} - \frac{\ln(\ln(n))}{n^2\ln^2(n)} \qquad (n \to +\infty),$$

a formula obtained as a result of a heuristic argument but not proved rigorously in [16], then we would have, for example,

$$\lim_{n \to +\infty} \frac{\mathcal{M}_{n^2}^{\text{SSP3}}}{\sqrt{\frac{n}{\ln(n)}}} = 1.$$

4. Internal amplification factors of extrapolation methods. In this section we give values, estimates, and bounds for the maximum internal amplification factor (2.18) for two classes of extrapolation methods: explicit Euler (EE) extrapolation and explicit midpoint (EM) extrapolation, both of which can be interpreted as ERK methods. These methods are described in section 4.1. The internal stability polynomials are given explicitly in section 4.2. A brief summary of our findings and techniques is presented in section 4.3. Detailed bounds are provided in sections 4.4 and 4.5. Again, due to space limitations, in this work we only list the main theorems. For complete proofs of all results, see [16].

Throughout this section, we write $\mathcal{M}_p^{\text{EE}}$ ($\mathcal{M}_p^{\text{EM}}$) to denote the maximum internal amplification factor of the order p EE (EM) extrapolation method.

4.1. The extrapolation algorithm. The extrapolation algorithm [10, section II.9] consists of two parts. In the first part, a *base method* and a *step-number sequence* are chosen. In general, the step-number sequence is a strictly monotone increasing sequence of positive integers n_j ($j = 1, 2, \ldots$). The base method is applied to compute multiple approximations to the ODE solution at time t_{n+1} based on the solution value at t_n: the first approximation, denoted by $T_{1,1}$, is obtained by dividing the interval

ALGORITHM 1. EE EXTRAPOLATION.

for $m = 1 \to p$ **do**   ▷ Compute first-order approximations
    $Y_{m,0} = U_n$
    for $j = 1 \to m$ **do**
        $Y_{m,j} = Y_{m,j-1} + \frac{\tau}{m} F(Y_{m,j-1})$
    end for
    $T_{m,1} = Y_{m,m}$
end for
for $k = 2 \to p$ **do**   ▷ Extrapolate to get higher order
    for $j = k \to p$ **do**
        $T_{j,k} = T_{j,k-1} + \frac{T_{j,k-1} - T_{j-1,k-1}}{\frac{j}{j-k+1} - 1}$ ▷ Aitken–Neville formula for extrapolation
to order k
    end for
end for
$U_{n+1} = T_{p,p}$ ▷ New solution value

ALGORITHM 2. EM EXTRAPOLATION.

$r = \frac{p}{2}$
for $m = 1 \to r$ **do**   ▷ Compute second-order approximations
    $Y_{m,0} = U_n$
    $Y_{m,1} = Y_{m,0} + \frac{\tau}{2m} F(Y_{m,0})$ ▷ Initial Euler step
    for $j = 2 \to 2m$ **do**
        $Y_{m,j} = Y_{m,j-2} + \frac{\tau}{m} F(Y_{m,j-1})$ ▷ Midpoint steps
    end for
    $T_{m,1} = Y_{m,2m}$
end for
for $k = 2 \to r$ **do**   ▷ Extrapolate to get higher order
    for $j = k \to r$ **do**
        $T_{j,k} = T_{j,k-1} + \frac{T_{j,k-1} - T_{j-1,k-1}}{\frac{j^2}{(j-k+1)^2} - 1}$ ▷ Aitken–Neville formula for extrapolation
to order $2k$
    end for
end for
$U_{n+1} = T_{r,r}$ ▷ New solution value

$t_{n+1} - t_n$ into n_1 step(s), the second approximation, $T_{2,1}$ is obtained by dividing it into n_2 steps, and so on. In the second part of the extrapolation algorithm, these $T_{m,1}$ values (that is, the low-order approximations) are combined by using the *Aitken–Neville interpolation formula* to get a higher-order approximation to the ODE solution at time t_{n+1}.

Here, the base method is the EE method (Algorithm 1) or the EM rule (Algorithm 2), and the step-number sequence is chosen to be the harmonic sequence $n_j := j$, being the "most economic" [10, formula (9.8)]. In the EM case, we also assume that the desired order of accuracy p is even. We do not consider the effects of smoothing [10].

4.2. The internal stability polynomials and the absolute stability region. By analyzing the perturbed scheme along the lines of Algorithms 1 and 2,

we get the internal stability polynomials $Q_{p,m,\ell}$—this time it is natural to use two indices, say, m and ℓ, to label them, with the dependence on p also indicated. The following results are proved in [16].

LEMMA 4.1. *For any $p \geq 2$ and $z \in \mathbb{C}$, the internal stability polynomials of the EE extrapolation method are given by*

$$Q_{p,m,\ell}^{\mathrm{EE}}(z) = \frac{(-1)^{m+p} m^{p-1}}{(p-m)!(m-1)!} \left(1 + \frac{z}{m}\right)^{\ell-1} \qquad (1 \leq m \leq p, 1 \leq \ell \leq m).$$

LEMMA 4.2. *For any even $p \geq 2$ and $z \in \mathbb{C}$, the internal stability polynomials of the EM extrapolation method are given by*

$$Q_{p,m,\ell}^{\mathrm{EM}}(z) = \frac{2(-1)^{m+r} m^{2r}}{(r-m)!(r+m)!} q_{m,2m-\ell+1}(z) \qquad (1 \leq m \leq r, 1 \leq \ell \leq 2m)$$

with $r = \frac{p}{2}$, and $q_{m,j}$ is defined by $q_{m,0}(z) := 0$, $q_{m,1}(z) := 1$, and

$$q_{m,j}(z) := \frac{z}{m} q_{m,j-1}(z) + q_{m,j-2}(z) \qquad (2 \leq j \leq 2m).$$

Let finally $\mathcal{S}_p$ denote the absolute stability region of the EE or EM extrapolation method of order p. Then one easily verifies (see [16]) that

$$\mathcal{S}_p = \left\{ z \in \mathbb{C} : \left| \sum_{k=0}^{p} \frac{z^k}{k!} \right| \leq 1 \right\}.$$

4.3. Estimates of the maximum internal amplification factors. In sections 4.4 and 4.5 we give estimates of the absolute value of the internal stability polynomials $Q_{p,m,\ell}^{\mathrm{EE}}$ and $Q_{p,m,\ell}^{\mathrm{EM}}$ over three regions, starting with the entire stability region $\mathcal{S}_p$. We know that $\mathcal{S}_p$ may include regions far into the right half-plane. In order to achieve tighter practical bounds, we also focus on the value of the internal stability polynomials over $\mathcal{S}_p \cap \mathbb{C}_{-,0}$, where $\mathbb{C}_{-,0}$ denotes the set of complex numbers with negative or zero real part. Finally, to understand the internal stability of the methods with very small step sizes, we investigate a third quantity, $\mathcal{M}_p^{\mathrm{EE}}(\{0\})$ and $\mathcal{M}_p^{\mathrm{EM}}(\{0\})$, as well.

Using the explicit expressions for the internal stability polynomials, one can employ *Mathematica* to compute exact values of $\mathcal{M}_p(\mathcal{S})$ for small to moderate values of p, for both extrapolation methods and for the sets $\mathcal{S} = \mathcal{S}_p$, $\mathcal{S} = \mathcal{S}_p \cap \mathbb{C}_{-,0}$ or $\mathcal{S} = \{0\}$. Decimal approximations of these exact values are given in tables in the following sections. (In this work we present only two tables; the remaining ones can be found in [16].) The values for Euler extrapolation corroborate the behavior observed with this method in earlier sections. The values for midpoint extrapolation show that it has much better internal stability.

Then we give upper (and sometimes lower) bounds on the amplification factors for arbitrary p values. For p sufficiently large, some improved estimates are also provided. In [16] we indicate certain conjectured asymptotically optimal growth rates as well. Although neither of the proved upper bounds is tight, these theorems again suggest that midpoint extrapolation is much more internally stable than EE extrapolation.

We remark that all the bounds on $\mathcal{M}_p^{\mathrm{EE}}$ and $\mathcal{M}_p^{\mathrm{EM}}$ in the following sections are based on bounds on $\mathcal{S}_p$ and $\mathcal{S}_p \cap \mathbb{C}_{-,0}$ of the form

$$z \in \mathcal{S}_p \implies |z| \leq c_1 p$$

or

$$z \in \mathcal{S}_p \cap \mathbb{C}_{-,0} \implies |z| \leq c_2 p$$

with suitable constants $c_1 > 0$ and $c_2 > 0$ presented in [16].

We make use of the Lambert W-function (`ProductLog` in *Mathematica*): for $x \geq -\frac{1}{e}$, there is a unique $W(x) \geq -1$ such that

(4.1) $$x = W(x)e^{W(x)}.$$

4.4. Bounds on $\mathcal{M}_p^{\mathbf{EE}}(\mathcal{S})$.

4.4.1. The case $\mathcal{S} = \mathcal{S}_p$. The first theorem gives an upper estimate of the maximum internal amplification factor over all the absolute stability region and for all $p \geq 2$.

THEOREM 4.3. *We have*

$$\mathcal{M}_2^{\mathrm{EE}}(\mathcal{S}_2) \leq \frac{13e^2}{10\sqrt{\pi}} < 5.42,$$

while for any $p \geq 3$

$$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p) \leq \frac{\left(\frac{2.6}{W\left(\frac{1}{e}\right)}\right)^p}{5.2\pi\sqrt{p-1}} < \frac{9.34^p}{5.2\pi\sqrt{p-1}}.$$

In [16] we show how one can reduce the constant 9.34 to get the following asymptotically better bound.

THEOREM 4.4. *For any $\varepsilon > 0$ there is a $p(\varepsilon) \in \mathbb{N}^+$ such that for all $p \geq p(\varepsilon)$*

$$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p) \leq \frac{\left(\frac{2+\varepsilon}{W\left(\frac{1}{e}\right)}\right)^p}{(4+2\varepsilon)\pi\sqrt{p-1}} < \frac{(7.19 + 3.6\varepsilon)^p}{(4+2\varepsilon)\pi\sqrt{p-1}}.$$

In the corresponding section of [16], two additional tables are presented: one with the exact values of $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p)$ for small p ($2 \leq p \leq 14$), and its extension to the range $15 \leq p \leq 20$ in the form of upper bounds on $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p)$. These tables contain much better upper estimates than Theorem 4.3. In [16] we also propose a heuristic argument about the conjectured asymptotically optimal upper estimate of $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p)$ being $\approx \frac{6.868^p}{2\pi p}$.

4.4.2. The case $\mathcal{S} = \mathcal{S}_p \cap \mathbb{C}_{-,0}$. The maximum internal amplification factors with z restricted to $\mathcal{S}_p \cap \mathbb{C}_{-,0}$ play an important role in practical computations. The first few exact values of $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S})$ are displayed in Table 4 for $2 \leq p \leq 14$. In [16], we give the details on how these values were obtained. Moreover, by using some simple estimates, we extend Table 4 to the range $15 \leq p \leq 20$ in the form of upper bounds on $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S})$.

An estimate valid for all $p \geq 3$ is cited next.

THEOREM 4.5. *For any $p \geq 3$,*

$$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0}) \leq \frac{\left(\frac{1.95}{W\left(\frac{1}{e}\right)}\right)^p}{3.9\pi\sqrt{p-1}} < \frac{7.01^p}{3.9\pi\sqrt{p-1}}.$$

The counterpart of Theorem 4.4 is the following result.

Maximum internal amplification factors for the first few EE extrapolation methods of order p with respect to the absolute stability region of the method in the left half of the complex plane. For $p = 3$ or $p \geq 6$, the exact maximum values are rounded up. (The exact values corresponding to $p \in \{4, 5\}$ are simple algebraic numbers.)

p	$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$	p	$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$
2	$\sqrt{2\left(1 + \sqrt{2}\right)} \approx 2.198$	9	11631.367
3	6.192	10	46860.486
4	$51/2 = 25.5$	11	98425.587
5	$\left(47 + \sqrt{65}\right)^{3/2}/\sqrt{18} \approx 96.305$	12	336910.368
6	190.163	13	$1.444 \cdot 10^6$
7	631.328	14	$6.561 \cdot 10^6$
8	2549.961		

THEOREM 4.6. *For any $\varepsilon > 0$ there is a $p(\varepsilon) \in \mathbb{N}^+$ such that for all $p \geq p(\varepsilon)$*

$$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0}) \leq \frac{\left(\frac{1 + \frac{1}{e} + \varepsilon}{W\left(\frac{1}{e}\right)}\right)^p}{(2 + \frac{2}{e} + 2\varepsilon)\pi\sqrt{p-1}} < \frac{(4.92 + 3.6\varepsilon)^p}{(2 + \frac{2}{e} + 2\varepsilon)\pi\sqrt{p-1}}.$$

Similarly to the previous section, in [16] we also sketch a heuristic argument about the conjectured asymptotically optimal upper estimate of $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$, now $\approx 0.342 \cdot \frac{3.885^p}{p}$.

4.4.3. The case $\mathcal{S} = \{0\}$. In this case $\mathcal{M}_p^{\mathrm{EE}}(\{0\}) = \max_{1 \leq m \leq p} \frac{m^p}{(p-m)!m!}$. The following theorem gives a two-sided estimate of $\mathcal{M}_p^{\mathrm{EE}}(\{0\})$.

THEOREM 4.7. *For any $p \geq 4$, we have*

$$0.117 \cdot \frac{3.577^p}{p} < \frac{\sqrt{3}}{2e^2} \cdot \frac{\left(\sqrt[4]{3}\, e\right)^p}{p} \leq \mathcal{M}_p^{\mathrm{EE}}(\{0\}) \leq \frac{\left(\frac{1}{W\left(\frac{1}{e}\right)}\right)^p}{2\pi\sqrt{p-1}} < \frac{3.592^p}{2\pi\sqrt{p-1}}.$$

One can also find the first few exact values of $\mathcal{M}_p^{\mathrm{EE}}(\{0\})$ for $2 \leq p \leq 20$ in [16], along with the $p \to +\infty$ asymptotic growth rate of the numbers $\mathcal{M}_p^{\mathrm{EE}}(\{0\})$. In addition, by taking into account the obvious relation

$$\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p) \geq \mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0}) \geq \mathcal{M}_p^{\mathrm{EE}}(\{0\}),$$

Theorem 4.7 also provides us with some lower estimates of $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p)$ or $\mathcal{M}_p^{\mathrm{EE}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$.

4.5. Bounds on $\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S})$. Turning to midpoint extrapolation, in this section $p \geq 2$ denotes an *even* integer and we let $r = p/2$. The corresponding section of [16] contains all the details, proofs, and tables mentioned below.

4.5.1. The case $\mathcal{S} = \mathcal{S}_p$. A theorem for all p and a result with an asymptotically better bound are presented first.

THEOREM 4.8. *For any $1 \leq r = \frac{p}{2} \leq 8$,*

$$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p) < \sqrt{\frac{2}{\pi}} \cdot \frac{4.74^p}{\sqrt{p}},$$

TABLE 5

Maximum internal amplification factors for the first few EM extrapolation methods of order p with respect to the absolute stability region of the method in the left half of the complex plane. For $p \geq 4$, the exact maximum values are rounded up. The algebraic degrees of the entries corresponding to $p = 2, 4, 6, 8$ and returned by Mathematica are 4, 38, 5, 7, respectively.

p	$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$	p	$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p \cap \mathbb{C}_{-,0})$
2	$\sqrt{2\left(1 + \sqrt{2}\right)} \approx 2.198$	6	25.378
4	7.332	8	88.755

while for any $r \geq 9$ we have

$$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p) < \frac{4.986^p}{\pi\sqrt{p-1}}.$$

THEOREM 4.9. *For any $\varepsilon > 0$ there is a $p(\varepsilon) \in \mathbb{N}^+$ such that for all even $p \geq p(\varepsilon)$*

$$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p) < \frac{(3.539 + \varepsilon)^p}{\pi\sqrt{p-1}}.$$

The exact values of $\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p)$ for even $2 \leq p \leq 8$ together with some further upper bounds on $\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p)$ extended up to $p = 20$ are also given in our cited work.

4.5.2. The case $\mathcal{S} = \mathcal{S}_p \cap \mathbb{C}_{-,0}$. As for the estimates in the left half-plane, the exact values of $\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S})$ for even $2 \leq p \leq 8$ are contained in Table 5. As an extension, [16] presents some upper bounds up to $p = 20$, including the estimate $\mathcal{M}_{10}^{\mathrm{EM}}(\mathcal{S}_{10} \cap \mathbb{C}_{-,0}) \leq 836$.

Concerning larger p values, we have the following two theorems.

THEOREM 4.10. *For any even $p \geq 12$,*

$$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p \cap \mathbb{C}_{-,0}) < \frac{3.423^p}{\pi\sqrt{p-1}}.$$

THEOREM 4.11. *For any $\varepsilon > 0$ there is a $p(\varepsilon) \in \mathbb{N}^+$ such that for all even $p \geq p(\varepsilon)$*

$$\mathcal{M}_p^{\mathrm{EM}}(\mathcal{S}_p \cap \mathbb{C}_{-,0}) < \frac{(2.157 + \varepsilon)^p}{\pi\sqrt{p-1}}.$$

4.5.3. The case $\mathcal{S} = \{0\}$. In [16] we list the exact values of $\mathcal{M}_p^{\mathrm{EM}}(\{0\})$ for $2 \leq p \leq 20$, as well as a $p \to +\infty$ asymptotic formula $\mathcal{M}_p^{\mathrm{EM}}(\{0\}) \sim 1.1524 \cdot \frac{1.509^p}{p}$.

5. Conclusion. Roundoff is not usually a significant source of error in traditional RK methods with typical error tolerances and double-precision arithmetic. However, when very high accuracy is required, it is natural to use very high order methods, which necessitates the use of large numbers of stages and can lead to substantial amplification of roundoff errors. This amplification is problematic precisely when high precision is desired. In fact, traditional error bounds and estimates that neglect roundoff error become useless in this situation.

In the past, internal error amplification had been a practical issue only in rare cases. However, current trends toward the use of higher-order methods and higher-accuracy computations suggest that it may be a more common concern in the future.

The analysis and bounds given in this work can be used to accurately estimate at what tolerance roundoff will become important in a given computation. More generally, the *maximum internal amplification factor* that we have defined provides a single useful metric for deciding whether internal stability may be a concern in a given computation.

We have emphasized that internal amplification depends on implementation and that the choice of implementation can be used to modify the internal stability polynomials. However, it is not yet clear whether dramatic improvements in internal stability can be achieved in this manner for methods of interest.

For implicit RK methods, internal stability may become important even when the number of stages is moderate. The numerical solution of the stage equations is usually performed iteratively and stopped when the stage errors are estimated to be "sufficiently small." If the amplification factor is large, then the one-step error also may be large even if the stage errors (and truncation error) are driven iteratively to small values. A study of the amplification of solver errors for practical implicit methods constitutes interesting future work.

REFERENCES

[1] S. ABARBANEL, D. GOTTLIEB, AND M. H. CARPENTER, *On the removal of boundary errors caused by Runge–Kutta integration of nonlinear partial differential equations*, SIAM J. Sci. Comput., 17 (1996), pp. 777–782.

[2] P. BOGACKI AND L. F. SHAMPINE, *An efficient Runge–Kutta (4, 5) pair*, Comput. Math. Appl., 32 (1996), pp. 15–28.

[3] J. C. BUTCHER, *Numerical Methods for Ordinary Differential Equations*, Wiley, New York, 2008.

[4] M. H. CARPENTER, D. GOTTLIEB, S. ABARBANEL, AND W.-S. DON, *The theoretical accuracy of Runge–Kutta time discretizations for the initial boundary value problem: A study of the boundary error*, SIAM J. Sci. Comput., 16 (1995), pp. 1241–1252.

[5] A. DUTT, L. GREENGARD, AND V. ROKHLIN, *Spectral deferred correction methods for ordinary differential equations*, BIT, 40 (2000), pp. 241–266.

[6] E. FEHLBERG, *Klassische Runge–Kutta-Formeln fünfter und siebenter Ordnung mit Schrittweiten-Kontrolle*, Computing, 4 (1969), pp. 93–106.

[7] L. FERRACINA AND M. N. SPIJKER, *An extension and analysis of the Shu–Osher representation of Runge–Kutta methods*, Math. Comp., 74 (2004), pp. 201–220.

[8] R. FRANK, J. SCHNEID, AND C. W. UEBERHUBER, *Stability properties of implicit Runge–Kutta methods*, SIAM J. Numer. Anal., 22 (1985), pp. 497–514.

[9] S. GOTTLIEB, D. I. KETCHESON, AND C.-W. SHU, *Strong Stability Preserving Runge–Kutta and Multistep Time Discretizations*, World Scientific, Singapore, 2011.

[10] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations* I: *Nonstiff Problems, 2nd ed.*, Springer Ser. Comput. Math., Springer, New York, 2010.

[11] K. HEUN, *Neue Methoden zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen*, Z. Math. Phys, 45 (1900), pp. 23–38.

[12] I. HIGUERAS, *Representations of Runge–Kutta methods and strong stability preserving methods*, SIAM J. Numer. Anal., 43 (2005), pp. 924–948.

[13] T. E. HULL, W. H. ENRIGHT, B. M. FELLEN, AND A. E. SEDGWICK, *Comparing numerical methods for ordinary differential equations*, SIAM J. Numer. Anal., 9 (1972), pp. 603–637.

[14] D. I. KETCHESON, *Highly efficient strong stability-preserving Runge–Kutta methods with low-storage implementations*, SIAM J. Sci. Comput., 30 (2008), pp. 2113–2136.

[15] D. I. KETCHESON AND U. BIN WAHEED, *A Theoretical Comparison of High Order Explicit Runge–Kutta, Extrapolation, and Deferred Correction Methods*, CAMCoS, 9 (2014), pp. 175–200.

[16] D. I. KETCHESON, L. LÓCZI, AND M. PARSANI, *Propagation of Internal Errors in Explicit Runge–Kutta Methods and Internal Stability of SSP and Extrapolation Methods*, arXiv:1309.1317, 2013.

[17] W. KUTTA, *Beitrag zur näherungweisen Integration totaler Differentialgleichungen*, Z. Math. Phys., 46 (1901), pp. 435–453.

[18] R. H. MERSON, *An operational method for the study of integration processes*, in Proceedings of Conference on Data Processing and Automatic Computing Machines at Weapons Research Establishments Salisbury, South Australia, 1957, pp. 1–25.

[19] J. NIEGEMANN, R. DIEHL, AND K. BUSCH, *Efficient low-storage Runge–Kutta schemes with optimized stability regions*, J. Comput. Phys., 231 (2011), pp. 372–364.

[20] M. PARSANI, D. I. KETCHESON, AND W. DECONINCK, *Optimized explicit Runge–Kutta schemes for the spectral difference method applied to wave propagation problems*, SIAM J. Sci. Comput., 35 (2013), pp. A957–A986.

[21] P. J. PRINCE AND J. R. DORMAND, *High order embedded Runge–Kutta formulae*, J. Comput. Appl. Math., 7 (1981), pp. 67–75.

[22] J. M. SANZ-SERNA AND J. G. VERWER, *Stability and convergence at the PDE/stiff ODE interface*, Appl. Numer. Math., 5 (1989), pp. 117–132.

[23] J. M. SANZ-SERNA, J. G. VERWER, AND W. H. HUNDSDORFER, *Convergence and order reduction of Runge–Kutta schemes applied to evolutionary problems in partial differential equations*, Numer. Math., 50 (1986), pp. 405–418.

[24] L. F. SHAMPINE, *Stability of explicit Runge–Kutta methods*, Comput. Math. Appl., 10 (1984), pp. 419–432.

[25] C.-W. SHU AND S. OSHER, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, J. Comput. Phys., 77 (1988), pp. 439–471.

[26] P. J. VAN DER HOUWEN AND B. P. SOMMEIJER, *On the Internal Stability of Explicit, m-Stage Runge–Kutta Methods for Large m-Values*, ZAMM Z. Angew. Math. Mech., 60 (1980), pp. 479–485.

[27] J. G. VERWER, *A class of stabilized three-step Runge–Kutta methods for the numerical integration of parabolic equations*, J. Comput. Appl. Math., 3 (1977), pp. 155–166.

[28] J. G. VERWER, *Explicit Runge–Kutta methods for parabolic partial differential equations*, Appl. Numer. Math., 22 (1996), pp. 359–379.

[29] J. G. VERWER, W. H. HUNDSDORFER, AND B. P. SOMMEIJER, *Convergence properties of the Runge–Kutta–Chebyshev method*, Numer. Math., 57 (1990), pp. 157–178.