

STEP SIZES FOR STRONG STABILITY PRESERVATION WITH DOWNWIND-BIASED OPERATORS*

DAVID I. KETCHESON†

Abstract. Strong stability preserving (SSP) integrators for initial value ODEs preserve temporal monotonicity solution properties in arbitrary norms. All existing SSP methods, including implicit methods, either require small step sizes or achieve only first order accuracy. It is possible to achieve more relaxed step size restrictions in the discretization of hyperbolic PDEs through the use of both upwind- and downwind-biased semidiscretizations. We investigate bounds on the maximum SSP step size for methods that include negative coefficients and downwind-biased semi-discretizations. We prove that the downwind SSP coefficient for linear multistep methods of order greater than one is at most equal to two, while the downwind SSP coefficient for explicit Runge–Kutta methods is at most equal to the number of stages of the method. In contrast, the maximal downwind SSP coefficient for second order Runge–Kutta methods is shown to be unbounded. We present a class of such methods with arbitrarily large SSP coefficient and demonstrate that they achieve second order accuracy for large CFL number.

Key words. strong stability-preserving, monotonicity, Runge–Kutta methods

AMS subject classifications. Primary, 65M20; Secondary, 65L06

DOI. 10.1137/100818674

1. Introduction.

1.1. Strong stability preservation. This work is concerned with numerical methods for the initial value problem

$$(1) \quad u'(t) = F(u), \quad u(0) = u_0,$$

where $u \in \mathbb{R}^n$ and $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Numerical methods for (1) compute a sequence of solutions u^1, u^2, \dots approximating the true solution at times $\Delta t, 2\Delta t, \dots$. We are interested in numerical solutions that satisfy the monotonicity property

$$(2) \quad \|u^n\| \leq \|u^{n-1}\|$$

in the case of a one-step method, or more generally

$$(3) \quad \|u^n\| \leq \max \{ \|u^{n-1}\|, \|u^{n-2}\|, \dots, \|u^{n-k}\| \}$$

for a k -step method.

A common approach is to assume that the initial value problem (1) is monotone under forward Euler integration, subject to some step size restriction:

$$(4) \quad \|u + \Delta t F(u)\| \leq \|u\| \quad \text{for all } u, \quad \text{for } 0 \leq \Delta t \leq \Delta t_{\text{FE}}.$$

The term *strong stability preserving* (SSP) is used to denote any method that gives a solution satisfying the monotonicity condition (3) whenever applied to an initial value

*Received by the editors December 20, 2010; accepted for publication (in revised form) May 23, 2011; published electronically August 4, 2011. This publication is based on work supported by award FIC/2010/05 - 2000000231, made by King Abdullah University of Science and Technology (KAUST).
<http://www.siam.org/journals/sinum/49-4/81867.html>

†King Abdullah University of Science and Technology, Box 4700, Thuwal, Saudi Arabia, 23955-6900 (david.ketcheson@kaust.edu.sa).

problem (1) satisfying the forward Euler condition (4). This can be shown to hold under a step size restriction of the form

$$(5) \quad \Delta t \leq \mathcal{C} \Delta t_{\text{FE}},$$

where the factor \mathcal{C} , referred to as the *SSP coefficient*, depends only on the numerical method. For a recent review of SSP methods see [2].

The SSP coefficient is, for most known methods, not very large; for a broad class of explicit general linear methods it is never greater than the number of stages of the method [6], and for many classes of implicit methods it is known or conjectured to be no greater than twice the number of stages [5, 1, 8]. No known methods have $\mathcal{C} > 2$ and exhibit higher than first order convergence for large CFL numbers.

1.2. Downwinding. The study of SSP methods has been motivated by the numerical solution of hyperbolic conservation laws; in one dimension these take the form

$$(6) \quad U_t + f(U)_x = 0.$$

Semidiscretization of the conservation law (6) leads to the initial value problem (1), where u is a finite-dimensional approximation of U and $F(u)$ is an approximation to $-f(U)_x$.

The solution of the scalar conservation law (6) has the property that its total variation does not increase in time. Hence it is desirable for a numerical discretization to satisfy (3) with respect to the total variation seminorm; such schemes are said to be total variation diminishing (TVD). A common approach to development of TVD methods is to employ a semidiscretization for which the TVD property holds under forward Euler integration in time, up to some maximal time step size; this is just condition (4), where $\|\cdot\|$ is taken to be the total variation seminorm.

Such semidiscretizations generally are upwind-biased. By considering corresponding downwind-biased semidiscretizations, one arrives at an operator \tilde{F} that approximates $f(U)_x$ and satisfies

$$(7) \quad \|u + \Delta t \tilde{F}(u)\| \leq \|u\| \quad \text{for all } \Delta t \leq \Delta t_{\text{FE}}.$$

For example, consider the advection equation

$$(8) \quad U_t + U_x = 0$$

discretized via the upwind and downwind discretizations:

$$(9) \quad u'_i(t) = -\frac{u_i - u_{i-1}}{\Delta x} = F_i(u), \quad u'_i(t) = -\frac{u_{i+1} - u_i}{\Delta x} = -\tilde{F}_i(u).$$

Here $u_i(t)$ is an approximation to $U(i\Delta x, t)$. In this case the conditions (4) and (7) hold with $\Delta t_{\text{FE}} = \Delta x$, corresponding to the usual CFL condition.

We say a method is strong stability preserving with *downwind SSP coefficient* $\tilde{\mathcal{C}}$ if the method gives a solution satisfying (3) whenever applied to a system satisfying the forward Euler conditions (4) and (7) under the time step restriction

$$(10) \quad \Delta t \leq \tilde{\mathcal{C}} \Delta t_{\text{FE}}.$$

The idea of using downwinding to achieve the TVD property under larger time steps for explicit time discretizations was originally introduced by Shu and Osher

[13, 14]. These methods are frequently referred to as *methods with downwind-biased discretizations*; in the present work we will refer to them simply as *downwind methods* for brevity, with the understanding that they incorporate both upwind- and downwind-biased discretizations. Optimal explicit downwind Runge–Kutta methods of up to fifth order and ten stages were developed in [12]. Further optimal explicit Runge–Kutta and linear multistep schemes, along with an approach to efficient implementation of upwind- and downwind-based WENO (weighted essentially nonoscillatory) discretizations, are given in [3]. A theory of necessary and sufficient conditions for Runge–Kutta methods with downwinding to be SSP was developed in [4].

The principal reason for studying downwind methods is that the downwind SSP coefficient \tilde{C} is not as restricted as the SSP coefficient C ; for instance, explicit Runge–Kutta methods can have order greater than four and $\tilde{C} > 0$ [12]. In the present work, we consider also *implicit* downwind methods. Generally speaking, implicit numerical methods are used in order to allow the use of time steps based solely on accuracy considerations and not on stability. As discussed above, implicit SSP methods require step sizes not much larger than those allowed by explicit methods. The principal aim of the present is to determine whether implicit downwind SSP methods allow much larger step sizes.

1.3. Downwind methods. Downwind Runge–Kutta methods take the form

$$(11a) \quad y_i = u^{n-1} + \Delta t \sum_{j=1}^s a_{ij} F(t^{n-1} + c_j \Delta t, y_j) \\ + \Delta t \sum_{j=1}^s \tilde{a}_{ij} \tilde{F}(t^{n-1} + c_j \Delta t, y_j), \quad 1 \leq i \leq s$$

$$(11b) \quad u^n = u^{n-1} + \Delta t \sum_{j=1}^s b_j F(t^{n-1} + c_j \Delta t, y_j) + \Delta t \sum_{j=1}^s \tilde{b}_j \tilde{F}(t^{n-1} + c_j \Delta t, y_j).$$

The method is explicit if each stage depends only on previous stages; i.e., if $a_{ij} = 0$, $\tilde{a}_{ij} = 0$ for $j \geq i$.

Downwind linear multistep methods take the form

$$(12) \quad u_n - \Delta t \beta_k F(u_n) - \Delta t \tilde{\beta}_k \tilde{F}(u_n) = \sum_{j=0}^{k-1} \alpha_j u_{n-k+j} + \Delta t \beta_j F(u_{n-k+j}) + \Delta t \tilde{\beta}_j \tilde{F}(u_{n-k+j}).$$

The method is explicit if $\beta_k = \tilde{\beta}_k = 0$.

2. Summary of main results. In this section we present the main results of the paper. The proofs are deferred to later sections.

Our first result concerns explicit downwind Runge–Kutta methods. It is known that the SSP coefficient C for a broad class of general linear methods (without downwinding) cannot be greater than the number of stages [6]. It turns out that the same bound holds for explicit downwind Runge–Kutta methods.

THEOREM 2.1. *The downwind SSP coefficient \tilde{C} of any first order accurate explicit downwind Runge–Kutta method is no greater than the number of stages of the method.*

It is already known (see [6]) that the bound $\tilde{C} \leq 1$ holds for explicit linear multistep methods. Together with the theorem above, this implies that not too much can

be gained by using downwinding in explicit methods, at least in an asymptotic sense. Hence we consider implicit downwind methods. For implicit linear multistep methods of greater than first order accuracy (without downwinding), it is known that the SSP coefficient cannot be greater than two. It turns out that the same bound holds for implicit downwind linear multistep methods.

THEOREM 2.2. *The downwind SSP coefficient of any second order accurate linear multistep method is at most two.*

Given the negative nature of the two theorems above, one might expect that bounds on the SSP coefficient hold for the downwind SSP coefficient for all classes of methods. It seems, however, that this is not the case; for implicit Runge–Kutta methods it is conjectured that the SSP coefficient can be no greater than $2s$, where s is the number of stages. However, it turns out that implicit downwind Runge–Kutta methods can have an arbitrarily large SSP coefficient.

THEOREM 2.3. *For any positive finite r , there exists a two-stage, second-order downwind Runge–Kutta method with downwind SSP coefficient equal to r .*

The rest of the paper proceeds as follows. The proof of Theorem 2.1 is presented in section 3. The proof of Theorem 2.2 is presented in section 4. In section 5, we prove Theorem 2.3 by construction and analyze the resulting methods. Some final remarks on these results are provided in section 7.

3. Explicit Runge–Kutta methods. In this section, we prove Theorem 2.1, which bounds the downwind SSP coefficient for explicit downwind Runge–Kutta methods. SSP methods are most conveniently analyzed by writing them in a certain Shu–Osher form; see [16, 7]. The corresponding form for downwind methods is introduced in the following lemma, which is an easy corollary of Theorems 3.5 and 3.6 of [16].

LEMMA 3.1. *The SSP coefficient of the downwind Runge–Kutta method (11) is the largest $r \geq 0$ such that the method can be written in the form*

$$(13a) \quad y_i = u^{n-1} + \Delta t \sum_{j=1}^s p_{ij} \left(y_j + \frac{\Delta t}{r} F(y_j) \right) + \Delta t \sum_{j=1}^s \tilde{p}_{ij} \left(y_j + \frac{\Delta t}{r} \tilde{F}(y_j) \right),$$

(13b)

$$u^n = u^{n-1} + \Delta t \sum_{j=1}^s p_{s+1,j} \left(y_j + \frac{\Delta t}{r} F(y_j) \right) + \Delta t \sum_{j=1}^s \tilde{p}_{s+1,j} \left(y_j + \frac{\Delta t}{r} \tilde{F}(y_j) \right),$$

with all coefficients nonnegative:

$$(14) \quad r_{ij}, p_{ij}, \tilde{p}_{ij} \geq 0.$$

In the remainder of this section, we consider explicit methods of the form (13); i.e., those for which

$$(15) \quad r_{ij} = p_{ij} = \tilde{p}_{ij} = 0 \quad \text{for all } j \geq i.$$

Consider the application of a method satisfying (13)–(15) to a linear problem with

$$(16) \quad F(u) = \mathbf{L}u, \quad \tilde{F}(u) = \tilde{\mathbf{L}}u,$$

where $\mathbf{L}, \tilde{\mathbf{L}}$ are fixed matrices, and define $z = \Delta t \mathbf{L}$, $\tilde{z} = \Delta t \tilde{\mathbf{L}}$. Then a straightforward calculation shows that the solution can be written as $u_n = \psi(z, \tilde{z})u_{n-1}$, where

$$(17) \quad \psi(z, \tilde{z}) = \sum_{j=0}^s \sum_{l=0}^j \gamma_{jl} \left(1 + \frac{z}{r}\right)^{j-l} \left(1 + \frac{\tilde{z}}{r}\right)^l \quad \text{with } \gamma_{jl} \geq 0 \text{ if } 0 \leq r \leq \tilde{c}.$$

Theorem 2.1 follows by considering the conditions for first-order accuracy of the method and positivity of the coefficients in the form (17).

Proof of Theorem 2.1. For a given downwind Runge–Kutta method, consider (17) with $r = \tilde{\mathcal{C}}$. Since $\tilde{F}(u) \approx -F(u)$, the method approximates the true solution to order p if

$$\psi(z, -z) = \exp(z) + \mathcal{O}(z^{p+1}).$$

In terms of the coefficients γ_{jl} , the conditions for consistency and first order accuracy are

$$\begin{aligned} \sum_{j=0}^s \sum_{l=0}^j \gamma_{jl} &= 1, \\ \sum_{j=0}^s \sum_{l=0}^j \gamma_{jl} (j - 2l) &= \tilde{\mathcal{C}}. \end{aligned}$$

Since the γ 's are positive, we have

$$\tilde{\mathcal{C}} = \sum_{j=0}^s \sum_{l=0}^j \gamma_{jl} (j - 2l) \leq \max_{j \in (0, s), l \in (0, j)} (j - 2l) = s. \quad \square$$

Remark 1. It is possible to show in a similar way that the bound $\tilde{\mathcal{C}} \leq s$ holds even for the broad class of general linear methods considered in [6].

Remark 2. In [7], tighter bounds on $\tilde{\mathcal{C}}$ were computed for specific classes of downwind Runge–Kutta methods. The same approach used there could be applied to find tighter bounds for classes of general linear methods.

4. Implicit linear multistep methods. Lenferink [10] showed that the SSP coefficient \mathcal{C} is no greater than two for implicit linear multistep methods of order greater than one. This bound was shown to hold in an even more general sense in [5]. In this section we consider the downwind SSP coefficient $\tilde{\mathcal{C}}$ for implicit downwind linear multistep methods (12). Our main result, presented already as Theorem 2.2, states that the same bound proved by Lenferink also holds for downwind methods. Even though this is a more general result than those of [10, 5], we provide a simpler proof.

The method (12) is accurate to order p if

$$(18) \quad \sum_{j=0}^{k-1} \alpha_j j^i + \sum_{j=0}^k (\beta_j - \tilde{\beta}_j) i j^{i-1} = k^i \quad (0 \leq i \leq p)$$

and has downwind SSP coefficient $\tilde{\mathcal{C}} \geq r$ if and only if

$$(19a) \quad \beta_j, \tilde{\beta}_j \geq 0 \quad (0 \leq j \leq k-1),$$

$$(19b) \quad \alpha_j - r(\beta_j + \tilde{\beta}_j) \geq 0 \quad (0 \leq j \leq k-1).$$

The following lemma, which appeared in [6], facilitates the proof of Theorem 2.2. We say a method (12) of k steps and order p is *optimal* if there exists no other method of at most k steps and order at least p with larger downwind SSP coefficient $\tilde{\mathcal{C}}$.

LEMMA 4.1. *Any optimal downwind linear multistep method (12) has the property that $\beta_j \tilde{\beta}_j = 0$ for each j .*

Proof of Lemma 4.1. Note that the order conditions (18) depend only on the difference $\beta_j - \tilde{\beta}_j$, while the inequality constraint (19a) can be written as (setting $r = \tilde{\mathcal{C}}$)

$$\tilde{\mathcal{C}} \leq \alpha_j / (\beta_j + \tilde{\beta}_j).$$

Suppose that an optimal method has $\beta_j > \tilde{\beta}_j > 0$ for $j \in J_1 \subset (0, 1, \dots, k-1)$ and $\tilde{\beta}_j > \beta_j > 0$ for $j \in J_2 \subset (0, 1, \dots, k-1)$. Then for $j \in J_1$ define $\beta_j^* = \beta_j - \tilde{\beta}_j$ and $\tilde{\beta}_j^* = 0$; for $j \in J_2$ define $\tilde{\beta}_j^* = \tilde{\beta}_j - \beta_j$ and $\beta_j^* = 0$. Then the coefficients obtained by replacing $\beta, \tilde{\beta}$ with $\beta^*, \tilde{\beta}^*$ satisfy (19) and (18) with a larger value of $\tilde{\mathcal{C}}$, which is a contradiction. \square

We now prove our main result on linear multistep methods. Similar to the proof of Theorem 2.1, this result follows from combining the order conditions and the positivity of the coefficients.

Proof of Theorem 2.2. It is sufficient to prove that $\tilde{\mathcal{C}} \leq 2$ for any optimal method. Hence we consider an optimal method that is at least second order accurate. Applying Lemma 4.1, we define $\beta_j^* = \beta_j + \tilde{\beta}_j$ and $\sigma_j = \text{sgn}(\beta_j - \tilde{\beta}_j)$ so that $\beta_j - \tilde{\beta}_j = \sigma_j \beta_j^*$. Additionally define $\delta_j = \alpha_j - r\beta_j^*$. Then the order conditions for order two can be written (taking $r = \tilde{\mathcal{C}}$)

$$(20a) \quad \sum_{j=0}^{k-1} \delta_j + \tilde{\mathcal{C}} \beta_j^* = 1,$$

$$(20b) \quad \sum_{j=0}^{k-1} j \delta_j + (\tilde{\mathcal{C}} j + \sigma_j) \beta_j^* = k - \sigma_k \beta_k^*,$$

$$(20c) \quad \sum_{j=0}^{k-1} j^2 \delta_j + (\tilde{\mathcal{C}} j^2 + 2j \sigma_j) \beta_j^* = k(k - 2\sigma_k \beta_k^*).$$

Taking $-k^2 \times (20a) + 2k \times (20b) - (20c)$ gives

$$(21) \quad \sum_{j=0}^{k-1} -(k-j)^2 \delta_j + \left(-\tilde{\mathcal{C}}(k-j)^2 + 2\sigma_j(k-j) \right) \beta_j^* = 0.$$

Since all coefficients δ_j, β_j^* are nonnegative, at least one of the terms multiplying them must be nonnegative, for some value of j (else the sum would be negative). Since the terms multiplying δ_j are all negative, this implies that

$$-\tilde{\mathcal{C}}(k-j)^2 + 2\sigma_j(k-j) \geq 0$$

for some j ; i.e.,

$$\tilde{\mathcal{C}} \leq \max_{j \in [0, k-1]} \frac{2\sigma_j(k-j)}{(k-j)^2} \leq 2. \quad \square$$

Remark 3. Using the order conditions and the SSP conditions, it is straightforward to determine (for any particular k, p) the methods with largest $\tilde{\mathcal{C}}$. This was done in [6].

5. Implicit Runge–Kutta methods. It is well known (see, e.g., [15, 9, 2]) that (implicit) Runge–Kutta methods of order higher than one cannot have $\mathcal{C} = \infty$. Furthermore, it is conjectured that they cannot have $\mathcal{C} > 2s$, where s is the number of stages [1, 8].

In [4], the following question was posed: Do there exist high order implicit downwind Runge–Kutta methods with $\tilde{\mathcal{C}} = \infty$? Although we have not answered this question directly, we have found downwind methods with arbitrarily large $\tilde{\mathcal{C}}$. One class of such methods is the two-stage, second order family

$$(22a) \quad y_1 = \frac{2}{r(r-2)}u^{n-1} + \frac{2}{r} \left(y_1 + \frac{\Delta t}{r} F(y_1) \right) + \frac{r^2 - 4r + 2}{r(r-2)} \left(y_2 + \frac{\Delta t}{r} \tilde{F}(y_2) \right),$$

$$(22b) \quad y_2 = y_1 + \frac{\Delta t}{r} F(y_1),$$

$$(22c) \quad u^n = y_2 + \frac{\Delta t}{r} F(y_2).$$

Here we have written the method in the form (13) so that the downwind SSP coefficient is apparent: $\tilde{\mathcal{C}} = r$, for any $r > 2 + \sqrt{2}$. It can be shown that the method (22) is A-stable (i.e., unconditionally stable when $F(u) = -\tilde{F}(u) = \lambda u$ with $\Re(\lambda) \leq 0$).

Numerical searches indicate that other two-stage, second order methods with large $\tilde{\mathcal{C}}$ exist. An obvious drawback of method (22) is that it is fully implicit. However, a search for two-stage, second order diagonally implicit methods yielded no results.

For reference, we include here also the coefficients of the method when written in form (11):

$$(23) \quad \begin{array}{c|cc|cc} \frac{r-2}{r} & \frac{r^2-2r-2}{2r} & 0 & 0 & \frac{r^2-4r+2}{2r} \\ \frac{r-1}{r} & \frac{r-2}{2} & 0 & 0 & \frac{r^2-4r+2}{2r} \\ \hline & \frac{r-2}{2} & \frac{1}{r} & 0 & \frac{r^2-4r+2}{2r} \end{array}.$$

The left part of (23) corresponds to the usual Butcher coefficients, and the rightmost part displays the additional coefficients $\tilde{a}_{ij}, \tilde{b}_j$. In the case $F = -\tilde{F}$, the method reduces to an ordinary Runge–Kutta method; we refer to this method as the *underlying method*:

$$(24) \quad \begin{array}{c|cc} \frac{r-2}{r} & \frac{r^2-2r-2}{2r} & -\frac{r^2-4r+2}{2r} \\ \frac{r-1}{r} & \frac{r-2}{2} & -\frac{r^2-4r+2}{2r} \\ \hline & \frac{r-2}{2} & \frac{4-r}{2} \end{array}.$$

Although the downwind method (23) will in general behave differently from the underlying method (24), analysis of the latter may still give useful insight. Straightforward calculation shows that the stability function of method (24) is

$$(25) \quad \psi(z) = \frac{1 + \frac{2}{r}z + \frac{1}{r^2}z^2}{1 - \frac{r-2}{r}z + \frac{r^2-4r+2}{2r^2}z^2}.$$

Further calculation reveals that the method is A-stable since $|\psi(z)| \leq 1$ for all z with nonpositive real part. It is clearly not L-stable, but it is nearly so for large r since

$$\lim_{|z| \rightarrow \infty} \psi(z) = \frac{2}{r^2 - 4r + 2}$$

is quite small for large r .

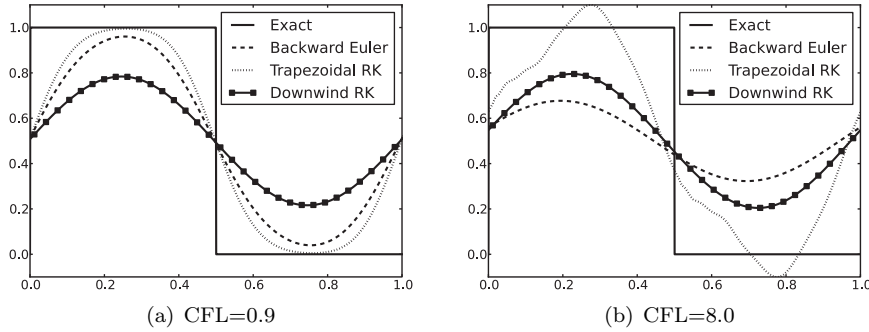


FIG. 1. Square wave advection with the downwind Runge–Kutta method (22), compared to two standard methods.

6. Numerical tests. In this section we conduct a few numerical tests to study the accuracy of the family of methods introduced in the last section. We compare the backward Euler method, the implicit trapezoidal Runge–Kutta method, and the downwind Runge–Kutta method (22). In all tests we take $r = 8$.

In the first test we will see that the temporal error for the downwind method includes a term related to the spatial discretizations. Consequently, some care must be taken when choosing the spatial discretizations. The second and third tests demonstrate that the method performs well when appropriate spatial discretizations are chosen.

6.1. First order upwind advection. As a first test we solve the advection equation (8) using the upwind and downwind differences (9) with $\Delta x = 1/128$. We consider the domain $0 \leq x \leq 1$ with periodic boundary conditions and initial condition $u(x, 0) = 1 - H(x - 1/2)$, where $H(x)$ is the Heaviside function. In Figure 1 we plot the solutions at $t = 1$.

Figure 1(a) shows results obtained with $\Delta t = 0.9\Delta x$. As expected, backward Euler is more dissipative than the second order trapezoidal method. Surprisingly, the (second order accurate) downwind Runge–Kutta method is more dissipative than even the first order backward Euler method. Figure 1(b) shows results obtained with $\Delta t = 8.0\Delta x$, the SSP limit for the downwind method. In this case, the trapezoidal method, which has SSP coefficient $\mathcal{C} = 2$, generates overshoots. The backward Euler method is much more dissipative for this large CFL number. The downwind method is more accurate than backward Euler and maintains maximum norm monotonicity, but the high level of dissipation exhibited is still disappointing.

The behavior of the downwind method can be understood through the following analysis. Write the upwind and downwind discretizations as $F(u) = \mathbf{L}u$, $\tilde{F}(u) = -\tilde{\mathbf{L}}u$, where $\mathbf{L}, \tilde{\mathbf{L}}$ are the matrices

$$(26) \quad \mathbf{L} = \frac{1}{\Delta x} \begin{pmatrix} -1 & & & 1 \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \frac{1}{\Delta x} \begin{pmatrix} 1 & -1 & & \\ & 1 & -1 & \\ & & \ddots & \ddots \\ -1 & & & 1 \end{pmatrix}.$$

Any downwind Runge–Kutta method applied to this problem results in a recurrence

TABLE 1

Max-norm errors for advection of a sine wave using fifth order WENO discretization and a CFL number of 8.

N	Backward Euler	2nd order RK	Downwind RK
32	0.728	0.730	0.436
64	0.603	0.215	0.168
128	0.452	0.054	0.043
256	0.292	0.013	0.011

of the form (see section 3 above)

$$(27) \quad u^n = \psi(\Delta t \mathbf{L}, \Delta t \tilde{\mathbf{L}}) u^{n-1}.$$

For the method (22), we find

$$\begin{aligned} \psi(\Delta t \mathbf{L}, \Delta t \tilde{\mathbf{L}}) &= \left(1 + \frac{z}{r}\right)^2 \left(1 + \frac{r^2 - 4r + 2}{2r} \Delta t \tilde{\mathbf{L}} - \frac{r^2 - 2r - 2}{2r} \Delta t \mathbf{L} \right. \\ &\quad \left. - \frac{r^2 - 4r + 1}{2r^2} \Delta t^2 \mathbf{L} \tilde{\mathbf{L}}\right)^{-1} \\ &= 1 + \Delta t \mathbf{L} + \frac{r^2 - 4r + 2}{2r} \Delta t (\mathbf{L} - \tilde{\mathbf{L}}) + \mathcal{O}(\Delta t^2). \end{aligned}$$

Since $\mathbf{L}u \approx -u_x$, the first two terms ensure that the method is at least first order accurate. However, the term involving $(\mathbf{L} - \tilde{\mathbf{L}})$ is problematic, since

$$(28) \quad \mathbf{L} - \tilde{\mathbf{L}} = \frac{1}{\Delta x} \begin{pmatrix} -2 & 1 & & 1 \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ 1 & & 1 & -2 \end{pmatrix} \approx \Delta x \, u_{xx}.$$

Thus, if F, \tilde{F} are spatial discretizations with order of accuracy q , then the one-step error for the downwind Runge–Kutta method (22) will contain a diffusive term of $\mathcal{O}(r\Delta x^q \Delta t)$. In order to avoid loss of accuracy, the spatial discretization should ensure that $r\Delta x^q \lesssim \Delta t$, so that this term is no larger than the $\mathcal{O}(\Delta t^2)$ term.

6.2. WENO advection. To demonstrate that proper accuracy is achieved when the foregoing condition is satisfied, we use a fifth order WENO interpolation to determine the fluxes. Table 1 shows convergence results in the maximum norm for advection of a sine wave ($u(x, 0) = \sin(2\pi x)$) using a CFL number of 8. Notice that the downwind method achieves its design order of 2 and gives accuracy similar to the trapezoidal Runge–Kutta method. It should be noted that the downwind method is more expensive computationally, since it is not diagonally implicit and requires both upwind and downwind operator evaluations. However, the purpose of this test is only to confirm the theoretically predicted accuracy.

6.3. Burgers equation. In order to provide an initial assessment of the suitability of these methods for application to nonlinear hyperbolic conservation laws, consider Burgers equation:

$$U_t + \left(\frac{1}{2}U^2\right)_x = 0$$

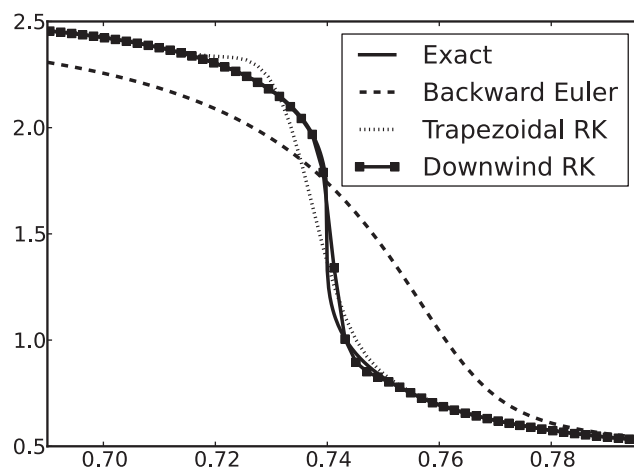


FIG. 2. Closeup view of shock in solution of Burgers equation with various methods using a CFL number of 6.5.

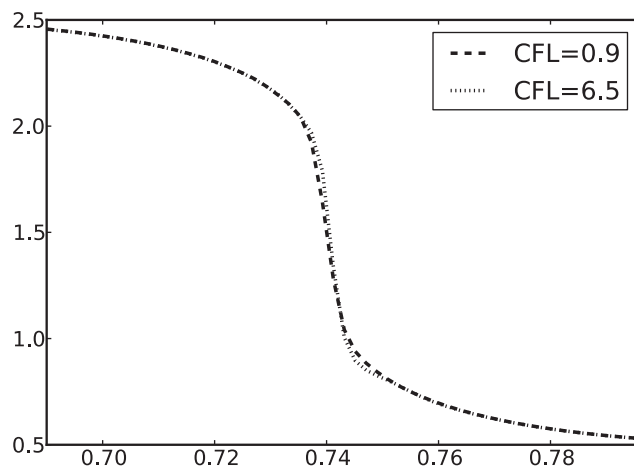


FIG. 3. Closeup view of shock in solution of Burgers equation obtained with the downwind RK method (22) with $r = 8$ and using two different CFL numbers.

on the unit interval with initial condition $U(x, 0) = \sin(2\pi x)$ and periodic boundary conditions. Fifth order WENO interpolation is again used to determine the fluxes. We consider the solution at time $t = 0.16$, just after a shock has formed. Figure 2 shows a closeup around the shock for the exact solution (obtained using characteristics) along with solutions computed by each of the methods compared above, using CFL number 6.5 in each case. A grid with $\Delta x = 1/512$ points is used. As expected, the backward Euler method is much more dissipative than the second order methods. Interestingly, the trapezoidal method solution does not show oscillations *per se*, but clearly has an unphysical bump behind the shock. The downwind method is by far the most accurate. Again, it should be noted that the downwind method is more expensive, since it involves an implicit solve of twice as many equations.

Figure 3 shows a comparison of two solutions obtained with the downwind RK method 22 using two different CFL numbers. The solutions are almost indistin-

guishable, suggesting that the spatial error is dominant, as one might expect. Close inspection reveals that the solution using the larger CFL number is more accurate. Importantly, it seems that the method does not become more dissipative at large CFL numbers.

We note in passing that solution of the nonlinear system of equations for implicit WENO schemes is a significant challenge, especially for large CFL number and in the presence of shocks. The CFL number 6.5 used here allowed the use of an easily accessible nonlinear solver, namely the `newton_krylov` and `fsolve` functions of the SciPy package. Efficient solution of the nonlinear system for larger CFL numbers is an area for future research.

7. Discussion. The new bounds on the downwind SSP coefficient proven here for explicit Runge–Kutta methods and implicit linear multistep methods are disappointing, since they indicate that nothing can be gained—in an asymptotic sense—by including downwind-biased discretizations for integrators in these classes. Previous attempts to find methods that have a large SSP coefficients have been similarly disappointing [11, 8].

The family of second order implicit downwind Runge–Kutta methods we report here is therefore quite remarkable. These are the first methods known to have a large SSP coefficient and to provide higher than first order accuracy in practice for CFL numbers larger than unity. They may be useful for problems with hyperbolic components requiring large time steps and the avoidance of spurious oscillations.

In light of this discovery, the following questions are naturally of interest:

- Do higher (than second) order downwind Runge–Kutta methods exist with large downwind SSP coefficient?
- Do diagonally implicit downwind Runge–Kutta methods exist with large SSP coefficient?
- Can these methods be implemented efficiently in combination with high order space discretizations and large time steps, for problems with shocks?

Preliminary investigation of the first question indicates the answer is affirmative. Work on all three questions is ongoing and will be presented elsewhere.

REFERENCES

- [1] L. FERRACINA AND M. SPJCKER, *Strong stability of singly-diagonally-implicit Runge-Kutta methods*, Appl. Numer. Math., 58 (2008), pp. 1675–1686.
- [2] S. GOTTLIEB, D. I. KETCHESON, AND C.-W. SHU, *High order strong stability preserving time discretizations*, J. Sci. Comput., 38 (2009), p. 251.
- [3] S. GOTTLIEB AND S. J. RUUTH, *Optimal strong-stability-preserving time-stepping schemes with fast downwind spatial discretizations*, J. Sci. Comput., 27 (2006), pp. 289–303.
- [4] I. HIGUERAS, *Representations of Runge-Kutta methods and strong stability preserving methods*, SIAM J. Numer. Anal., 43 (2005), pp. 924–948.
- [5] W. HUNSDORFER AND S. J. RUUTH, *On monotonicity and boundedness properties of linear multistep methods*, Math. Comp., 75 (2005), pp. 655–672.
- [6] D. I. KETCHESON, *Computation of optimal monotonicity preserving general linear methods*, Math. Comp., 78 (2009), pp. 1497–1513.
- [7] D. I. KETCHESON, *High order strong stability preserving time integrators and numerical wave propagation methods for hyperbolic PDEs*, Ph.D. thesis, University of Washington, Seattle, WA, 2009.
- [8] D. I. KETCHESON, C. B. MACDONALD, AND S. GOTTLIEB, *Optimal implicit strong stability preserving Runge-Kutta methods*, Appl. Numer. Math., 52 (2009), p. 373.
- [9] J. F. B. M. KRAAIJEVANGER, *Contractivity of Runge-Kutta methods*, BIT, 31 (1991), pp. 482–528.

- [10] H. W. J. LENFERINK, *Contractivity-preserving implicit linear multistep methods*, Math. Comp., 56 (1991), pp. 177–199.
- [11] C. MACDONALD, S. GOTTLIEB, AND S. RUUTH, *A numerical study of diagonally split Runge-Kutta methods for PDEs with discontinuities*, J. Sci. Comput., 36 (2008), pp. 89–112.
- [12] S. J. RUUTH AND R. J. SPITERI, *High-order strong-stability-preserving Runge-Kutta methods with downwind-biased spatial discretizations*, SIAM J. Numer. Anal., 42 (2004), pp. 974–996.
- [13] C.-W. SHU AND S. OSHER, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, J. Comput. Phys., 77 (1988), pp. 439–471.
- [14] C.-W. SHU, *Total-variation-diminishing time discretizations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 1073–1084.
- [15] M. N. SPIJKER, *Contractivity in the numerical solution of initial value problems*, Numer. Math., 42 (1983), pp. 271–290.
- [16] M. N. SPIJKER, *Stepsize conditions for general monotonicity in numerical initial value problems*, SIAM J. Numer. Anal., 45 (2007), pp. 1226–1245.