

LINEAR INSTABILITY OF THE FIFTH-ORDER WENO METHOD

RONG WANG* AND RAYMOND J. SPITERI†

Abstract. The weighted essentially non-oscillatory (WENO) methods are popular spatial discretization methods for hyperbolic partial differential equations. In this paper we show that the combination of the widely used fifth-order WENO spatial discretization (WENO5) and the forward Euler time integration method is linearly unstable when numerically integrating hyperbolic conservation laws. Consequently it is not convergent. Furthermore we show that all two-stage, second-order explicit Runge–Kutta (ERK) methods are linearly unstable (and hence do not converge) when coupled with WENO5. We also show that all optimal first- and second-order strong-stability-preserving (SSP) ERK methods are linearly unstable when coupled with WENO5. Moreover the popular three-stage, third-order SSP(3,3) ERK method offers no linear stability advantage over non-SSP ERK methods, including ones with negative coefficients, when coupled with WENO5. We give new linear stability criteria for combinations of WENO5 with general ERK methods of *any order*. We find that a sufficient condition for the combination of an ERK method and WENO5 to be linearly stable is that the linear stability region of the ERK method should include the part of the imaginary axis of the form $[-i\mu, i\mu]$, for some $\mu > 0$. The linear stability analysis also provides insight into the behaviour of ERK methods applied to nonlinear problems and problems with discontinuous solutions. We confirm the assertions of our analysis by means of numerical tests.

Key words. Stability analysis, Runge–Kutta methods, WENO method, strong stability preserving.

AMS subject classifications. 65L06, 65M20

1. Introduction. The method of lines (MOL) is a general approach for the treatment of time-dependent partial differential equations (PDEs) [24]. The standard MOL involves two steps. The first step is to discretize the spatial variables of the PDE to obtain a large set of initial-value ordinary differential equations (ODEs). The second step is to integrate the ODEs using a time integration method such as a linear multistep or Runge–Kutta (RK) method [4, 5].

The essentially non-oscillatory (ENO) methods [7, 6] and the weighted essentially non-oscillatory (WENO) methods [15, 11] are popular and effective nonlinear spatial discretizations for hyperbolic PDEs. These methods are adept at handling the non-smooth features that arise in the solutions to hyperbolic PDEs. For example, although these methods are formally first-order accurate once a shock is present, they still have uniform high-order accuracy right up to the location of the shock [11]. Specifically, the fifth-order WENO spatial discretization (WENO5) [11], which uses a convex combination of three third-order ENO stencils, is a widely used and robust spatial discretization for numerical solution of hyperbolic conservation laws.

The three-stage, third-order strong-stability-preserving (SSP) explicit RK (ERK) method, which has most recently been referred to as SSP(3,3) [21], is generally viewed as the time integration method of choice to couple with WENO5; see e.g., [16] and references therein. Numerical results are generically stable and satisfactory [11, 20]. Because of this, it is generally believed that when WENO5 is used as the spatial discretization for hyperbolic PDEs, the SSP property is a necessary (or advantageous) property for the time integration method to possess [16, 20]. However, for ERK methods applied to hyperbolic conservation laws spatially discretized by WENO5, our work shows that there is a different property, i.e., linear stability, that must be considered.

In this paper we analyze the linear stability of some ERK methods when coupled with WENO5 to solve hyperbolic conservation laws. First we show that the forward Euler method is linearly unstable when coupled with WENO5; i.e., the corresponding CFL number is 0. Consequently the numerical solution does

*Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5C9, Canada. The work of this author was partially supported by MITACS and Martec, Inc. rong@cs.usask.ca.

†Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5C9, Canada. The work of this author was partially supported by NSERC Canada, MITACS, and Martec, Inc. spiteri@cs.usask.ca.

not converge to the true solution for any time step. This means that the stability of WENO5 in general is a product of its nonlinear nature and the particular time integration method with which it is coupled. Furthermore SSP ERK methods [18, 3] derived based on the SSP property of forward Euler cannot be SSP when coupled with WENO5. In fact, there is no guarantee for stability of any kind. In Section 3 we show that any optimal s -stage, first- or second-order SSP ERK method [21] is linearly unstable when coupled with WENO5. This result is surprising and contrary to expectation based on existing literature; see e.g., [16, 20, 25] and references therein. Moreover, we show that, in our analysis, the success of the SSP(3,3) method for the time integration of hyperbolic PDEs spatially discretized by WENO5 is not due to its SSP property; indeed any three-stage, third-order ERK method (even with negative coefficients or that is provably non-SSP) possesses the same linear stability properties. We demonstrate all these results by means of numerical experiments.

For linear spatial discretizations of linear, constant-coefficient PDE problems posed on unbounded or periodic domains, linear instability guarantees *global* instability; the instability typically leads to spurious oscillations in the numerical solution that ultimately become unbounded [23]. However, WENO methods are not linear spatial discretizations; see e.g., [9, p.59] for a definition of linear spatial discretizations for periodic advection problems. Therefore, the behaviour of a numerical solution computed from linearly unstable combinations of WENO methods and certain time integration methods for solving hyperbolic conservation laws is more subtle. Such a combination arises for example when using the combination of WENO5 and the forward Euler method; see Theorem 3.1 below. In such cases, the linear instability again manifests itself as spurious oscillations in the numerical solution. However, WENO methods attempt to adapt to the instability by changing the associated ENO stencil weights; see Section 2 below. The spurious oscillations may still grow to large magnitudes; however, they do not necessarily become unbounded. In other words, the nonlinear nature of WENO methods may be successful in controlling potential instabilities, and this process may take a significant length of time to assert itself. It is important to note that because they are linearly unstable, these combined methods are not convergent. Hence although the discussions in this paper are mainly phrased in terms of linear stability analysis, an immediate corollary of every result presented regarding linear instability is the non-convergence of the combined method. We offer further discussion and illustration of this in Section 4; see also Example 1.

The remainder of this paper proceeds as follows. In Section 2 we give a brief review of WENO5. In Section 3 we prove that, when coupled with WENO5, the forward Euler method and all two-stage, second-order ERK methods are linearly unstable. We then provide criteria for the linear stability of general ERK methods of orders 1 and 2. Immediate consequences of these criteria are that the optimal s -stage ($s \geq 2$), first-order or s -stage ($s \geq 3$), second-order SSP ERK methods are linearly unstable when coupled with WENO5. It is also easy to show that all three-stage, third-order ERK methods (including SSP(3,3)) and all four-stage, fourth-order ERK methods (including the classical four-stage, fourth-order ERK method) are linearly stable when coupled with WENO5. Finally we find that a sufficient condition for the combination of any ERK time integration method with WENO5 to be linearly stable is that the linear stability region of the ERK method must include the part of the imaginary axis in the form $[-i\mu, i\mu]$, for some $\mu > 0$. In Section 4 we confirm our theoretical results by means of numerical experiments. We also derive and test four new non-SSP ERK methods (a two-stage, first-order method; a three-stage, second-order method; a three-stage, third-order method with negative coefficients; and a low-storage five-stage, third-order method) that are stable according to our linear stability analysis. Numerical results for both linear and nonlinear problems, as well as problems with continuous and discontinuous solutions, demonstrate the relevance of our analysis.

2. The WENO5 method. WENO methods are widely used for the spatial discretization of hyperbolic conservation laws. They were first introduced in [15] as an improvement to ENO methods. ENO methods are based on polynomial interpolation of solution data to define numerical fluxes. They were originally designed to suppress instabilities that lead to spurious oscillations in other commonly used spatial discretizations. To achieve this, ENO methods choose stencils that are adapted to the directions where the solution has an increased order of smoothness. WENO methods take a convex combination of r candidate ENO stencils of order r to produce a method of order $2r - 1$ in regions where the solution is smooth while retaining the ENO property in regions where the solution exhibits discontinuous behaviour. Specifically, for a given cell, the WENO5 method consists of a convex combination of the three possible third-order ENO stencils containing that cell [11]. Although in practice ENO or WENO methods are very robust, there are very few theoretical

results for them [20]. We note that no proof of the stability of either family of methods has yet been given. The WENO5 method is perhaps the most commonly used of the WENO family of methods. We now give a brief summary of some theoretical aspects of WENO methods, with specific implementational details given for the WENO5 method.

Consider the one-dimensional scalar hyperbolic conservation law

$$u_t = -f_x(u), \quad 0 < x < 1, \quad t > 0. \quad (2.1)$$

Assume a uniform spatial mesh; i.e., $x_j = j\Delta x$, $j = 0, 1, \dots, N$, where $\Delta x = \frac{1}{N}$, and define cells by $I_j = [x_{j-1}, x_j]$, $i = 1, 2, \dots, N$. We use a conservative finite difference scheme in a method-of-lines approach to write

$$\frac{du_j}{dt} = -\frac{1}{\Delta x} \left(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} \right),$$

where $u_j(t) \approx u(x_j, t)$, $j = 0, 1, \dots, N$. The term $\hat{f}_{j+\frac{1}{2}} = \hat{f}(u_{j-R}, \dots, u_{j+S})$ is the numerical flux. The numerical flux must be consistent with $f(u)$; i.e., $\hat{f}(u, \dots, u) = f(u)$. The specification of $\hat{f}(u)$ determines the particular numerical method and its properties. We now derive the specific form of the numerical flux $\hat{f}(u)$ for WENO5.

We first split the flux into positive and negative parts

$$f(u) = f^+(u) + f^-(u). \quad (2.2)$$

This can be accomplished in different ways. In this paper we only consider the Lax–Friedrichs flux splitting [20]

$$f^+(u) = \frac{1}{2} (f(u) + mu), \quad f^-(u) = \frac{1}{2} (f(u) - mu),$$

where $m = \max |f'(u)|$. It is easy to show that $\frac{df^+}{du} \geq 0$, and $\frac{df^-}{du} \leq 0$. However, we note that the same analysis and conclusions apply to other flux splitting methods.

As in [11], we now calculate the indicators of smoothness IS_i , $i = 0, 1, \dots, r-1$, associated with the i th-stencil. For IS_i^+ we use

$$IS_i^+ = \sum_{m=1}^{r-1} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \Delta x^{2m-1} \left(\frac{\partial^m p_i(x)}{\partial x^m} \right)^2 dx, \quad i = 0, 1, \dots, r-1,$$

where $p_i(x)$ is the interpolating polynomial of order r to solution data on r cells; i.e., $I_{j+i-r+1}, \dots, I_{j+i}$. In the case of WENO5, $r = 3$, this leads to

$$IS_0^+ = \frac{13}{12} (f_{j-2}^+ - 2f_{j-1}^+ + f_j^+)^2 + \frac{1}{4} (f_{j-2}^+ - 4f_{j-1}^+ + 3f_j^+)^2, \quad (2.3)$$

$$IS_1^+ = \frac{13}{12} (f_{j-1}^+ - 2f_j^+ + f_{j+1}^+)^2 + \frac{1}{4} (f_{j-1}^+ - f_{j+1}^+)^2, \quad (2.4)$$

$$IS_2^+ = \frac{13}{12} (f_j^+ - 2f_{j+1}^+ + f_{j+2}^+)^2 + \frac{1}{4} (3f_j^+ - 4f_{j+1}^+ + f_{j+2}^+)^2, \quad (2.5)$$

and using an analagous formula for IS_i^- , we have

$$IS_0^- = \frac{13}{12} (f_{j+1}^- - 2f_{j+2}^- + f_{j+3}^-)^2 + \frac{1}{4} (3f_{j+1}^- - 4f_{j+2}^- + f_{j+3}^-)^2,$$

$$IS_1^- = \frac{13}{12} (f_j^- - 2f_{j+1}^- + f_{j+2}^-)^2 + \frac{1}{4} (f_j^- - f_{j+2}^-)^2,$$

$$IS_2^- = \frac{13}{12} (f_{j-1}^- - 2f_j^- + f_{j+1}^-)^2 + \frac{1}{4} (f_{j-1}^- - 4f_j^- + 3f_{j+1}^-)^2.$$

Next the (non-normalized) stencil weights take the form

$$\alpha_i^\pm = \frac{d_i}{(\epsilon + IS_i^\pm)^2}, \quad i = 0, 1, \dots, r-1,$$

where ϵ is a small positive number that is introduced to avoid the denominator becoming zero. In the numerical experiments of this paper, we choose $\epsilon = 10^{-6}$, which is the value recommended in [11]. In the case of the WENO5 method, we have $d_0 = \frac{1}{10}$, $d_1 = \frac{6}{10}$, $d_2 = \frac{3}{10}$ (see e.g., [11] for a derivation of the d_i), and

$$\alpha_0^\pm = \frac{1}{10} \left(\frac{1}{\epsilon + IS_0^\pm} \right)^2, \quad \alpha_1^\pm = \frac{6}{10} \left(\frac{1}{\epsilon + IS_1^\pm} \right)^2, \quad \alpha_2^\pm = \frac{3}{10} \left(\frac{1}{\epsilon + IS_2^\pm} \right)^2. \quad (2.6)$$

In order to achieve a convex combination of ENO stencils, the WENO stencil weights are normalized according to

$$w_i^\pm = \frac{\alpha_i}{\sum_{m=0}^{r-1} \alpha_m}, \quad i = 0, 1, \dots, r-1,$$

to give

$$w_0^\pm = \frac{\alpha_0^\pm}{\alpha_0^\pm + \alpha_1^\pm + \alpha_2^\pm}, \quad w_1^\pm = \frac{\alpha_1^\pm}{\alpha_0^\pm + \alpha_1^\pm + \alpha_2^\pm}, \quad w_2^\pm = \frac{\alpha_2^\pm}{\alpha_0^\pm + \alpha_1^\pm + \alpha_2^\pm}. \quad (2.7)$$

We note that $w_j^\pm \in (0, 1)$, $j = 0, 1, 2$, and $w_0^\pm + w_1^\pm + w_2^\pm = 1$, as required.

The numerical fluxes for WENO5 are given by

$$\begin{aligned} \hat{f}_{j+\frac{1}{2}}^+ &= w_0^+ \left(\frac{2}{6} f_{j-2}^+ - \frac{7}{6} f_{j-1}^+ + \frac{11}{6} f_j^+ \right) + w_1^+ \left(-\frac{1}{6} f_{j-1}^+ + \frac{5}{6} f_j^+ + \frac{2}{6} f_{j+1}^+ \right) \\ &\quad + w_2^+ \left(\frac{2}{6} f_j^+ + \frac{5}{6} f_{j+1}^+ - \frac{1}{6} f_{j+2}^+ \right) \end{aligned} \quad (2.8)$$

and

$$\begin{aligned} \hat{f}_{j+\frac{1}{2}}^- &= w_2^- \left(-\frac{1}{6} f_{j-1}^- + \frac{5}{6} f_j^- + \frac{2}{6} f_{j+1}^- \right) + w_1^- \left(\frac{2}{6} f_j^- + \frac{5}{6} f_{j+1}^- - \frac{1}{6} f_{j+2}^- \right) \\ &\quad + w_0^- \left(\frac{11}{6} f_{j+1}^- - \frac{7}{6} f_{j+2}^- + \frac{2}{6} f_{j+3}^- \right). \end{aligned}$$

Noting (2.2), the WENO5 method takes the final form

$$\frac{du_j}{dt} = -\frac{1}{\Delta x} \left[\left(\hat{f}_{j+\frac{1}{2}}^+ - \hat{f}_{j-\frac{1}{2}}^+ \right) + \left(\hat{f}_{j+\frac{1}{2}}^- - \hat{f}_{j-\frac{1}{2}}^- \right) \right]. \quad (2.9)$$

We refer the interested reader to [20] and references therein for further details and discussion of WENO methods.

3. Linear Stability Analysis. We consider the linear stability properties of various ERK methods when coupled with WENO5 to solve hyperbolic conservation laws (2.1). The CFL number associated with a uniform discretization in both space and time of (2.1) is defined as $\sigma = \left(\max \frac{\partial f}{\partial u} \right) \frac{\Delta t}{\Delta x}$. As is usual when performing linear stability analysis, we linearize and freeze coefficients to write

$$u_t = -\lambda u_x.$$

The corresponding CFL number σ is then $\sigma = \frac{\lambda \Delta t}{\Delta x}$. For the purposes of our analysis, it is sufficient to consider the one-dimensional scalar advection equation

$$u_t = -u_x;$$

i.e., $f(u) = u$; any CFL number appearing in our analysis can then be scaled appropriately for more general interpretations.

We begin this section by showing that the combination of WENO5 and the forward Euler ERK method is linearly unstable.

The Lax–Friedrichs flux splitting yields

$$f^+(u) = u, \quad f^-(u) = 0;$$

i.e., the negative part of the flux is zero. Thus, equation (2.8) takes the form

$$\begin{aligned} \hat{f}_{j+\frac{1}{2}}^+ &= w_0^+ \left(\frac{2}{6}u_{j-2} - \frac{7}{6}u_{j-1} + \frac{11}{6}u_j \right) + w_1^+ \left(-\frac{1}{6}u_{j-1} + \frac{5}{6}u_j + \frac{2}{6}u_{j+1} \right) \\ &+ w_2^+ \left(\frac{2}{6}u_j + \frac{5}{6}u_{j+1} - \frac{1}{6}u_{j+2} \right). \end{aligned} \quad (3.1)$$

When the exact solution is smooth, it is well known that

$$w_0^+ = \frac{1}{10} + \epsilon_1, \quad w_1^+ = \frac{6}{10} + \epsilon_2, \quad w_2^+ = \frac{3}{10} + \epsilon_3, \quad (3.2)$$

where $\epsilon_1, \epsilon_2, \epsilon_3$ are all $\mathcal{O}((\Delta x)^2)$, and $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$ [11]. We now prove the following theorem using von Neumann analysis.

THEOREM 3.1. *The combination of WENO5 and forward Euler is linearly unstable.*

Proof. Assuming periodic boundary conditions, we can expand the approximate solution $\{u_j\}_{j=0}^N$ as a finite Fourier series

$$u_j = \sum_{k=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} \hat{u}_k e^{\iota j \xi_k \Delta x},$$

where ι is the imaginary unit, i.e., $\iota^2 = -1$, and ξ_k is the spatial frequency associated with \hat{u}_k . Because the wave equation is linear and has constant coefficients, it is sufficient to consider only one individual Fourier mode; i.e.,

$$u_j = \hat{u} e^{\iota j \xi \Delta x}.$$

Defining $\phi = \xi \Delta x$, we thus have

$$u_j = \hat{u}e^{\iota j\phi}. \quad (3.3)$$

By using (3.2), (3.1), and (3.3), we now obtain

$$\begin{aligned} \hat{f}_{j+\frac{1}{2}}^+ = u_j \left[w_0^+ \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) + w_1^+ \left(-\frac{1}{6}e^{-\iota\phi} + \frac{5}{6} + \frac{2}{6}e^{\iota\phi} \right) \right. \\ \left. + w_2^+ \left(\frac{2}{6} + \frac{5}{6}e^{\iota\phi} - \frac{1}{6}e^{2\iota\phi} \right) \right]. \end{aligned} \quad (3.4)$$

When the forward Euler method is used, the WENO5 method becomes

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} L(u_{j-3}^n, u_{j-2}^n, \dots, u_{j+2}^n), \quad (3.5)$$

where

$$L(u_{j-3}^n, u_{j-2}^n, \dots, u_{j+2}^n) = \hat{f}_{j+\frac{1}{2}}^{+,n} - \hat{f}_{j-\frac{1}{2}}^{+,n}.$$

From (3.4) we see that $\frac{\hat{f}_{j+\frac{1}{2}}^{+,n} - \hat{f}_{j-\frac{1}{2}}^{+,n}}{u_j^n}$ is a function of ϕ . Thus we define $z(\phi) = \frac{\hat{f}_{j+\frac{1}{2}}^{+,n} - \hat{f}_{j-\frac{1}{2}}^{+,n}}{u_j^n}$ and obtain

$$L(u_{j-3}^n, u_{j-2}^n, \dots, u_{j+2}^n) = z(\phi)u_j^n.$$

Defining the CFL number $\sigma = \frac{\Delta t}{\Delta x}$, (3.5) becomes

$$\begin{aligned} u_j^{n+1} &= u_j^n - \sigma z(\phi)u_j^n \\ &= u_j^n (1 - \sigma z(\phi)). \end{aligned} \quad (3.6)$$

Define the *amplification factor* $g(\sigma z(\phi)) = 1 - \sigma z(\phi)$. In order to prove the theorem, we need to show that for any $\sigma > 0$, there exists a ϕ such that $|g| > 1$. Let ϕ be a small positive number. Using the Taylor expansion of $\sin \phi$ and $\cos \phi$ in (2.3), (2.4), and (2.5), we obtain

$$\begin{aligned} IS_0^+ &= u_j^2 \phi^2 \left(-1 + \frac{5}{12}\phi^2 - \frac{89}{40}\phi^4 + \frac{4889}{4032}\phi^6 - \frac{5}{3}\iota\phi^3 + \frac{17}{9}\iota\phi^5 + \mathcal{O}(\phi^7) \right), \\ IS_1^+ &= u_j^2 \phi^2 \left(-1 + \frac{17}{12}\phi^2 - \frac{9}{40}\phi^4 + \frac{337}{20160}\phi^6 + \mathcal{O}(\phi^7) \right), \\ IS_2^+ &= u_j^2 \phi^2 \left(-1 + \frac{5}{12}\phi^2 - \frac{89}{40}\phi^4 + \frac{4889}{4032}\phi^6 + \frac{5}{3}\iota\phi^3 - \frac{17}{9}\iota\phi^5 + \mathcal{O}(\phi^7) \right), \end{aligned}$$

For $0 < \epsilon \ll 1$, we can choose $\phi = \phi(\epsilon) = \mathcal{O}(\epsilon^{1/9})$ sufficiently small such that the α_i^+ , $i = 0, 1, 2$, in (2.6) can be estimated as follows:

$$\begin{aligned} \alpha_0^+ &= \frac{1}{u_j^4 \phi^4} \left(\frac{1}{10} + \frac{1}{12}\phi^2 - \frac{943}{2400}\phi^4 - \frac{7879}{4320}\phi^6 - \frac{1}{3}\iota\phi^3 - \frac{5}{9}\iota\phi^5 + \mathcal{O}(\phi^7) \right), \\ \alpha_1^+ &= \frac{1}{u_j^4 \phi^4} \left(\frac{6}{10} + \frac{17}{10}\phi^2 + \frac{1337}{400}\phi^4 + \frac{19057}{3600}\phi^6 + \mathcal{O}(\phi^7) \right), \\ \alpha_2^+ &= \frac{1}{u_j^4 \phi^4} \left(\frac{3}{10} + \frac{1}{4}\phi^2 - \frac{943}{800}\phi^4 - \frac{7879}{1440}\phi^6 + \iota\phi^3 + \frac{5}{3}\iota\phi^5 + \mathcal{O}(\phi^7) \right). \end{aligned}$$

Substituting the above expressions into (2.7), we obtain

$$\begin{aligned} w_0^+ &= \frac{1}{10} - \frac{3}{25}\phi^2 - \frac{163}{500}\phi^4 - \frac{30449}{30000}\phi^6 - \frac{2}{5}\iota\phi^3 + \frac{17}{75}\iota\phi^5 + \mathcal{O}(\phi^7), \\ w_1^+ &= \frac{6}{10} + \frac{12}{25}\phi^2 + \frac{163}{125}\phi^4 + \frac{20449}{7500}\phi^6 - \frac{2}{5}\iota\phi^3 - \frac{13}{75}\iota\phi^5 + \mathcal{O}(\phi^7), \\ w_2^+ &= \frac{3}{10} - \frac{9}{25}\phi^2 - \frac{489}{500}\phi^4 - \frac{51347}{30000}\phi^6 + \frac{4}{5}\iota\phi^3 - \frac{4}{75}\iota\phi^5 + \mathcal{O}(\phi^7). \end{aligned}$$

Thus the real and imaginary parts of $\hat{f}_{j+\frac{1}{2}}^{+,n}$ are

$$\operatorname{Re} \hat{f}_{j+\frac{1}{2}}^{+,n} = u_j^n \left(1 - \frac{1}{12}\phi^2 - \frac{1}{720}\phi^4 + \frac{241}{21600}\phi^6 + \mathcal{O}(\phi^8) \right), \quad (3.7)$$

$$\operatorname{Im} \hat{f}_{j+\frac{1}{2}}^{+,n} = u_j^n \left(\frac{1}{2}\phi - \frac{7}{60}\phi^5 + \mathcal{O}(\phi^7) \right). \quad (3.8)$$

Similarly it can be shown that

$$\operatorname{Re} \hat{f}_{j-\frac{1}{2}}^{+,n} = u_j^n \left(1 - \frac{1}{12}\phi^2 - \frac{1}{720}\phi^4 - \frac{2279}{21600}\phi^6 + \mathcal{O}(\phi^8) \right), \quad (3.9)$$

$$\operatorname{Im} \hat{f}_{j-\frac{1}{2}}^{+,n} = u_j^n \left(-\frac{1}{2}\phi - \frac{7}{60}\phi^5 + \mathcal{O}(\phi^7) \right). \quad (3.10)$$

Using (3.7)–(3.10), we obtain

$$z(\phi) = \frac{7}{60}\phi^6 + \mathcal{O}(\phi^8) + \iota(\phi + \mathcal{O}(\phi^7)). \quad (3.11)$$

The amplification factor becomes

$$g = 1 - \sigma(\mathcal{O}(\phi^6) + \iota(\phi + \mathcal{O}(\phi^7))).$$

A simple calculation now shows that

$$\begin{aligned} |g|^2 &= (1 - \sigma\mathcal{O}(\phi^6))^2 + \sigma^2(\phi + \mathcal{O}(\phi^7))^2 \\ &= 1 + \sigma^2\phi^2 + \mathcal{O}(\phi^6) \\ &> 1, \quad \forall \sigma > 0. \end{aligned}$$

This completes the proof. \square

REMARK 1. Equation (3.11) is valid only when ϕ is a small positive number. The general form of $z(\phi)$ for any ϕ is given later in (3.21).

REMARK 2. This form of analysis applies to any linear finite difference method for $\frac{du_j}{dt} = -\frac{1}{\Delta x}L(u_{j-R}, \dots, u_{j+S})$, where $L(u_{j-R}, \dots, u_{j+S})$ can be written in the form

$$L(u_{j-R}, \dots, u_{j+S}) = u_j z(\phi),$$

where $z(\phi)$ is uniquely determined by the spatial operator.

REMARK 3. It should be noted that although this instability argument applies to the WENO spatial discretization, it does not necessarily apply to the ENO spatial discretization. The reason is that this analysis

needs to have a known stencil, and ENO methods may choose any of a number of candidate stencils even if the solution is smooth. Moreover, such a “randomly” chosen stencil may lead to an unstable method. This is why a biased choice for choosing ENO stencils is suggested in [19], and this strategy has been very successful for ENO methods in practice.

REMARK 4. It follows immediately that, for the class of problems considered in this analysis, the combination of the forward Euler method and the WENO5 spatial discretization is not SSP for any step size $\Delta t > 0$. Because every ERK generates its second stage by a forward Euler step, this second stage cannot be SSP, and hence in this framework, no ERK method can be SSP¹. Hence the SSP property offers no stability advantage.

Consider the general s -stage ERK method written in standard form:

$$\begin{aligned} y_n^{(1)} &= y_n, \\ y_n^{(k)} &= y_n + \Delta t \sum_{i=1}^{k-1} a_{k,i} f\left(t_n + c_i \Delta t, y_n^{(i)}\right), \quad k = 2, \dots, s, \\ y_{n+1} &= y_n + \Delta t \sum_{i=1}^s b_i f\left(t_n + c_i \Delta t, y_n^{(i)}\right), \end{aligned}$$

where the c_k satisfy the conditions

$$c_k = a_{k1} + a_{k2} + \dots + a_{k,k-1},$$

for $k = 1, \dots, s$.

Its Butcher tableau is of the form given in Table 3.1.

0						
c_2	a_{21}					
c_3	a_{31}	a_{32}				
c_4	a_{41}	a_{42}	a_{43}			
\vdots	\vdots	\vdots	\vdots	\ddots		
c_s	a_{s1}	a_{s2}	a_{s3}	\dots	$a_{s,s-1}$	
	b_1	b_2	b_3	\dots	b_{s-1}	b_s

TABLE 3.1
Butcher tableau for s -stage ERK methods

We have the following theorem and corollary [12].

THEOREM 3.2. *The amplification factor for an s -stage ERK method is*

$$g(\hat{z}) = 1 + \sum_{l=1}^s \left(\sum_{j=l}^s b_j \left(\sum_{j>m_1>\dots>m_{l-1}\geq 1} a_{j,m_1} a_{m_1,m_2} \dots a_{m_{l-2},m_{l-1}} \right) \right) \hat{z}^l \quad (3.12)$$

$$= 1 + \hat{z} b^T (I - \hat{z} A)^{-1} e, \quad (3.13)$$

¹Recall that an ERK method written in Shu–Osher form is SSP if all its stages i on step n of the numerical solution $Y_n^{(i)}$ satisfy $\|Y_n^{(i)}\| \leq \|Y_n\|$, for all $i = 1, 2, \dots, s$ and $n \geq 1$, for some suitable semi-norm $\|\cdot\|$, where $Y_n^{(1)} = Y_n$, and $Y_n^{(s)} = Y_{n+1}$; see, e.g., [2, 8].

where I is the unit matrix, $A = (a_{ij})_{1 \leq i, j \leq s}$ and $b = (b_1, b_2, \dots, b_s)$ are the coefficients of the Butcher tableau, $e = (1, 1, \dots, 1)$, $z(\phi)$ is determined by the spatial operator, and $\hat{z} = -\sigma z$.

Combining the order conditions for ERK methods with (3.12) or (3.13), we easily obtain the following corollary [14, 22].

COROLLARY 3.3. *The amplification factor of an s -stage, order- p ERK method is*

$$\begin{aligned} g(\hat{z}) &= 1 + \sum_{l=1}^p \frac{1}{l!} \hat{z}^l + \sum_{l=p+1}^s \left(\sum_{j=l}^s b_j \left(\sum_{j > m_1 > \dots > m_{l-1} \geq 1} a_{j, m_1} a_{m_1, m_2} \cdots a_{m_{l-2}, m_{l-1}} \right) \right) \hat{z}^l \quad (3.14) \\ &= 1 + \sum_{l=1}^p \frac{\hat{z}^l}{l!} + \sum_{l=p+1}^s \hat{z}^l b^T A^{l-1} e. \end{aligned}$$

Therefore, a spatial discretization scheme combined with a given ERK method is linearly stable if and only if g in (3.12) satisfies $|g| \leq 1$, $\forall \phi \in [0, 2\pi]$.

We can now prove the following theorem for any two-stage, second-order ERK method.

THEOREM 3.4. *The combination of WENO5 with any two-stage, second-order ERK method is linearly unstable.*

Proof. From (3.14), the amplification factor is given by

$$g(\hat{z}) = 1 - \sigma z + \frac{1}{2}(\sigma z)^2.$$

Choosing ϕ to be a small positive number and using (3.11), a simple calculation shows

$$\begin{aligned} |g|^2 &= \left(1 - \frac{1}{2}\sigma^2\phi^2 + \mathcal{O}(\phi^6) \right)^2 + (-\sigma\phi + \mathcal{O}(\phi^7))^2 \\ &= 1 + \frac{1}{4}\sigma^4\phi^4 + \mathcal{O}(\phi^6) \\ &> 1. \end{aligned}$$

This finishes the proof. \square

Before we derive the theorem for general ERK methods, we give a geometric interpretation for the CFL number σ [1, 10, 14, 22].

DEFINITION 3.5. *Let D_1 denote the classical (linear) stability domain of any ERK method, and let D_2 denote the region interior to the boundary $\{-z(\phi) : 0 \leq \phi \leq 2\pi\}$ in the complex domain. The CFL number σ is the largest nonnegative real number such that the scaled region σD_2 is contained in D_1 .*

It is well known that (3.13) is the stability function for an ERK method; see e.g., [5]. Thus $D_1 = \{\hat{z} : |g(\hat{z})| \leq 1\}$. Note that the set $\{\hat{z} = -\sigma z(\phi) : 0 \leq \phi \leq 2\pi\}$ represents the boundary of the scaled region σD_2 . It is clear that in order to have $\hat{z} \in D_1$, the scaled region σD_2 must be contained in D_1 . We now give two simple examples for the purposes of illustration. The first example is for the upwind spatial discretization; i.e., $du_j/dt = L(u_{j-1}, u_j) = -\frac{1}{\Delta x}(u_j - u_{j-1})$. Using von Neumann analysis, we obtain $z(\phi) = 1 - e^{-i\phi}$. It is easy to see that the set $\{-z(\phi) : 0 \leq \phi \leq 2\pi\}$ represents a circle in the complex plane with center $(-1, 0)$ and radius 1; i.e., D_2 is the shaded area shown in Figure 3.1.

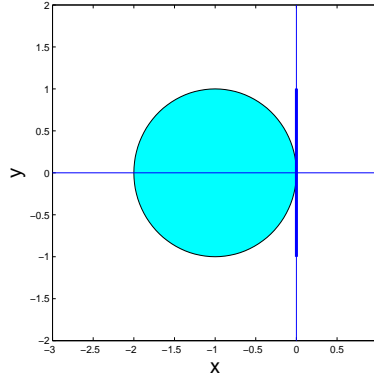


FIG. 3.1. D_2 for the upwind (circle) and central (line segment) spatial discretizations.

If the forward Euler method is used for the time discretization, its classical linear stability domain D_1 is $\{z : |1+z| \leq 1\}$. In other words, D_1 is exactly the same as D_2 in this case. It is trivial to conclude therefore that $\sigma = 1$ is the largest number such that $\sigma D_2 \subseteq D_1$. Hence the CFL number is 1.

The second example is for the central finite difference spatial discretization; i.e., $du_j/dt = L(u_{j-1}, u_j, u_{j+1}) = -\frac{1}{\Delta x} \frac{u_{j+1} - u_{j-1}}{2}$. Using von Neumann analysis, we obtain $z(\phi) = \frac{1}{2}(e^{i\phi} - e^{-i\phi}) = i \sin \phi$. It is easy to see that the set $\{-z(\phi) : 0 \leq \phi \leq 2\pi\} = \{(0, y) : -1 \leq y \leq 1\}$; i.e. D_2 now represents a finite segment of the imaginary axis. If the forward Euler method is used for the time discretization, $\sigma D_2 \not\subseteq D_1$, no matter how small $\sigma > 0$ is chosen. Therefore, the central finite difference spatial discretization is linearly unstable when it is coupled with the forward Euler method.

We now derive the following Lemma for any consistent ERK method. Lemma 3.6 is important for all the theorems in this paper.

LEMMA 3.6. *The classical (linear) stability domain of any consistent ERK method contains a rectangle $[-\eta, 0] \times [-i\hat{\mu}, i\hat{\mu}]$ for some $\eta, \hat{\mu} > 0$ if and only if it has an intersection with the imaginary axis of the form $[-i\mu, i\mu]$, for some $\mu \geq \hat{\mu} > 0$.*

Proof. \implies If the rectangle $[-\eta, 0] \times [-i\hat{\mu}, i\hat{\mu}]$ is inside the classical (linear) stability domain of any consistent ERK method, by definition the part of the imaginary axis $[-i\hat{\mu}, i\hat{\mu}]$ is also inside the stability domain. In other words, the stability domain of the ERK method intersects the imaginary axis at $[-i\mu, i\mu]$, for some $\mu \geq \hat{\mu} > 0$.

\Leftarrow Assume that the stability domain of the ERK method intersects the imaginary axis at $[-i\mu, i\mu]$, for some $\mu > 0$. We first prove that the stability domain of the ERK method intersects the part of the (negative) real axis at $[-\hat{\eta}, 0]$, for some $\hat{\eta} > 0$. Recall that the stability function of any consistent ERK method is of the form

$$g(\hat{z}) = 1 + \hat{z} + \text{higher-order terms.}$$

Let $\hat{z} = -\gamma$, where γ is a small, positive real number. In other words, choose \hat{z} to be close to the origin and on the negative real axis. It is easy to see $g(-\gamma) = 1 - \gamma + \mathcal{O}(\gamma^2)$. It is obvious that $0 < |g(-\gamma)| < 1$ as $\gamma \rightarrow 0+$. That is, there is an intersection of the stability domain with the negative real axis. Assume the intersection is of the form $[-\hat{\eta}, 0]$ for some real $\hat{\eta} > 0$.

Now by contradiction assume that no rectangle $[-\eta, 0] \times [-i\hat{\mu}, i\hat{\mu}]$ is contained inside the stability domain. Using the facts that the stability domain intersects the negative real axis at $[-\hat{\eta}, 0]$, and it intersects the imaginary axis at $[-i\mu, i\mu]$, we give a schematic representation of the topology of a sufficiently small neighborhood of the origin in Figure 3.2. In the figure, areas inside the stability domain are shaded. For simplicity we focus on the second quadrant in the complex plane, and we show only one region not contained in the stability domain. Thus, if we define $\hat{z} = \hat{x} + i\hat{y}$, where \hat{x}, \hat{y} are sufficiently small real numbers, there are at least

two numbers, \hat{x}_1 and \hat{x}_2 , for each \hat{y} such that $|g(\hat{x}_1 + \iota\hat{y})| = |g(\hat{x}_2 + \iota\hat{y})| = 1$. Therefore, the equation $R(\hat{x}, \hat{y}) = |g(\hat{x} + \iota\hat{y})|^2 - 1 = 0$ must have more than one solution $\hat{x} = \hat{x}(\hat{y})$ in the neighbourhood of $(0, 0)$.

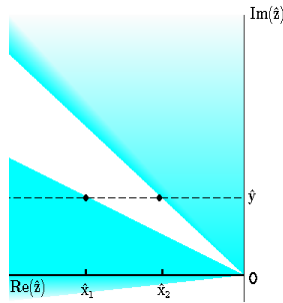


FIG. 3.2. A schematic of the topology of the classical stability domain in a sufficiently small neighbourhood of the origin if a rectangle $[-\nu, 0] \times [-\hat{\mu}, \hat{\mu}]$ is not contained in it. Shaded areas are inside in the stability domain.

We now obtain a contradiction by using the implicit function theorem to prove that in fact $\hat{x}(\hat{y})$ is a unique solution to $R(\hat{x}, \hat{y}) = 0$ in a small neighborhood of the origin. It is easy to see that

$$g(\hat{x} + \iota\hat{y}) = 1 + \hat{x} + \iota\hat{y} + Q_1(\hat{x}, \hat{y}) + \iota Q_2(\hat{x}, \hat{y}), \quad (3.15)$$

where $Q_1(\hat{x}, \hat{y})$ and $Q_2(\hat{x}, \hat{y})$ are two real polynomials of the form

$$Q_1(\hat{x}, \hat{y}) = \sum_{\substack{l, k \geq 0, \\ l+k \geq 2}} \xi_{l, k} \hat{x}^l \hat{y}^k, \quad (3.16)$$

$$Q_2(\hat{x}, \hat{y}) = \sum_{\substack{l, k \geq 0, \\ l+k \geq 2}} \zeta_{l, k} \hat{x}^l \hat{y}^k, \quad (3.17)$$

with real coefficients $\xi_{l, k}$, $\zeta_{l, k}$. Using (3.15)—(3.17), we obtain

$$R(\hat{x}, \hat{y}) = (1 + \hat{x} + Q_1(\hat{x}, \hat{y}))^2 + (\hat{y} + Q_2(\hat{x}, \hat{y}))^2 - 1.$$

A simple calculation shows that

$$\frac{\partial R}{\partial \hat{x}}(0, 0) = 2 \neq 0.$$

From the implicit function theorem, we know that there is a unique solution $\hat{x} = \hat{x}(\hat{y})$ to $R(\hat{x}, \hat{y}) = 0$ in a small neighborhood of the origin, yielding the desired contradiction. \square

REMARK 5. *The result of Lemma 3.6 can be generalized to any consistent one-step method with very little modification of the proof.*

REMARK 6. *Because of this result, the prospect that the spectra of the spatially discretized system may contain negative real components is not problematic.*

REMARK 7. *This result is the equivalent of the result of [13] for local stability. In particular the regions of stability described for ERK methods are equivalent.*

Using Definition 3.5 and Lemma 3.6, we have the following theorem for first-order ERK methods.

THEOREM 3.7. *There exists a CFL number σ such that the combination of WENO5 with a first-order ERK method is linearly stable for $\Delta t/\Delta x \leq \sigma$ provided the first-order ERK method satisfies*

$$\sum_{1 \leq i < j \leq s} b_j a_{ji} > \frac{1}{2}; \quad (3.18)$$

on the other hand, this combination is linearly unstable if

$$\sum_{1 \leq i < j \leq s} b_j a_{ji} < \frac{1}{2}.$$

NOTE 1. *When $\sum_{1 \leq i < j \leq s} b_j a_{ji} = \frac{1}{2}$, the ERK method is second order. The corresponding results are given in Theorem 3.9.*

Proof. We first prove the linearly unstable case. Let $\tau_2 = \sum_{1 \leq i < j \leq s} b_j a_{ji}$; i.e., τ_2 is the coefficient of \hat{z}^2 in (3.12). For $\phi > 0$ sufficiently small, from (3.11) it is easy to show that

$$\begin{aligned} z^l &= l^l \phi^l + \mathcal{O}(\phi^6) + \iota \mathcal{O}(\phi^6), & \text{if } 2 \leq l < 6, \\ z^l &= \mathcal{O}(\phi^6), & \text{if } l \geq 6. \end{aligned}$$

From (3.12), we obtain

$$\begin{aligned} g(-\sigma z(\phi)) &= 1 - \sigma z + \tau_2 (\sigma z)^2 + \mathcal{O}(z^3) \\ &= (1 - \tau_2 \sigma^2 \phi^2 + \mathcal{O}(\phi^4)) - \iota (\sigma \phi + \mathcal{O}(\phi^3)). \end{aligned}$$

A simple calculation shows that

$$\begin{aligned} |g|^2 &= (1 - \tau_2 \sigma^2 \phi^2 + \mathcal{O}(\phi^4))^2 + (\sigma \phi + \mathcal{O}(\phi^3))^2 \\ &= 1 + (1 - 2\tau_2) \sigma^2 \phi^2 + \mathcal{O}(\phi^4) \\ &> 1, \end{aligned}$$

if $\tau_2 < \frac{1}{2}$. Therefore, the combined method is linearly unstable.

Now let us assume $\tau_2 > \frac{1}{2}$. We first show that the stability domain D_1 of the corresponding ERK intersects the imaginary axis. Later (in Theorem 3.14) we show that in fact this is a sufficient condition for linear stability of any ERK method when combined with WENO5. Recall that the stability function of the ERK method is of the form

$$g(\hat{z}) = 1 + \hat{z} + \tau_2 \hat{z}^2 + \dots$$

Let $\hat{z} = \iota \gamma$, where γ is a small real number. In other words, choose \hat{z} to be close to the origin and on the imaginary axis. It is easy to see

$$\begin{aligned} |g(\iota \gamma)|^2 &= (1 - \tau_2 \gamma^2 + \mathcal{O}(\gamma^4))^2 + (\gamma + \mathcal{O}(\gamma^3))^2 \\ &= (1 - 2\tau_2 \gamma^2 + \mathcal{O}(\gamma^4)) + (\gamma^2 + \mathcal{O}(\gamma^4)) \\ &= 1 + (1 - 2\tau_2) \gamma^2 + \mathcal{O}(\gamma^4). \end{aligned}$$

Using the condition $\tau_2 > \frac{1}{2}$, we obtain that $|g(\iota \gamma)| < 1$ as $\gamma \rightarrow 0+$. That is, there is an intersection of D_1 with the imaginary axis. Assume the intersection is the interval $[-\mu, \mu]$ for some real $\mu > 0$. Then from Lemma 3.6 there exists a rectangle $D_3 = [-\eta, 0] \times [-\hat{\mu}, \hat{\mu}] \subseteq D_1$ for some $\eta > 0$, $0 < \hat{\mu} \leq \mu$.

It is easy to derive the expression for $\hat{f}_{j+\frac{1}{2}}^+$ from (3.4) and (3.2):

$$\begin{aligned} \hat{f}_{j+\frac{1}{2}}^+ &= u_j \left[\left(\frac{2}{60}e^{-2\iota\phi} - \frac{13}{60}e^{-\iota\phi} + \frac{47}{60} + \frac{27}{60}e^{\iota\phi} - \frac{3}{60}e^{2\iota\phi} \right) + \epsilon_1 \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) \right. \\ &\quad \left. + \epsilon_2 \left(-\frac{1}{6}e^{-\iota\phi} + \frac{5}{6} + \frac{2}{6}e^{\iota\phi} \right) + \epsilon_3 \left(\frac{2}{6} + \frac{5}{6}e^{\iota\phi} - \frac{1}{6}e^{2\iota\phi} \right) \right]. \end{aligned} \quad (3.19)$$

Similarly we obtain the expression for $\hat{f}_{j-\frac{1}{2}}^+$:

$$\begin{aligned} \hat{f}_{j-\frac{1}{2}}^+ &= u_j \left[\left(\frac{2}{60}e^{-3\iota\phi} - \frac{13}{60}e^{-2\iota\phi} + \frac{47}{60}e^{-\iota\phi} + \frac{27}{60} - \frac{3}{60}e^{\iota\phi} \right) + \epsilon_4 \left(\frac{2}{6}e^{-3\iota\phi} - \frac{7}{6}e^{-2\iota\phi} + \frac{11}{6}e^{-\iota\phi} \right) \right. \\ &\quad \left. + \epsilon_5 \left(-\frac{1}{6}e^{-2\iota\phi} + \frac{5}{6}e^{-\iota\phi} + \frac{2}{6} \right) + \epsilon_6 \left(\frac{2}{6}e^{-\iota\phi} + \frac{5}{6} - \frac{1}{6}e^{\iota\phi} \right) \right], \end{aligned} \quad (3.20)$$

where ϵ_4 , ϵ_5 , and ϵ_6 are all $\mathcal{O}((\Delta x)^2)$, and $\epsilon_4 + \epsilon_5 + \epsilon_6 = 0$.

Using (3.19), (3.20), and the definition $z(\phi) = \frac{\hat{f}_{j+\frac{1}{2}}^{+,n} - \hat{f}_{j-\frac{1}{2}}^{+,n}}{u_j^n}$, we obtain

$$z(\phi) = \tilde{z} + M(\epsilon_1, \epsilon_2, \dots, \epsilon_6, \phi), \quad (3.21)$$

where

$$\tilde{z} = -\frac{1}{30}e^{-3\iota\phi} + \frac{1}{4}e^{-2\iota\phi} - e^{-\iota\phi} + \frac{1}{3} + \frac{1}{2}e^{\iota\phi} - \frac{1}{20}e^{2\iota\phi}, \quad (3.22)$$

and

$$\begin{aligned} M(\epsilon_1, \epsilon_2, \dots, \epsilon_6, \phi) &= \epsilon_1 \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) + \epsilon_2 \left(-\frac{1}{6}e^{-\iota\phi} + \frac{5}{6} + \frac{2}{6}e^{\iota\phi} \right) \\ &\quad + \epsilon_3 \left(\frac{2}{6} + \frac{5}{6}e^{\iota\phi} - \frac{1}{6}e^{2\iota\phi} \right) - \epsilon_4 \left(\frac{2}{6}e^{-3\iota\phi} - \frac{7}{6}e^{-2\iota\phi} + \frac{11}{6}e^{-\iota\phi} \right) \\ &\quad - \epsilon_5 \left(-\frac{1}{6}e^{-2\iota\phi} + \frac{5}{6}e^{-\iota\phi} + \frac{2}{6} \right) - \epsilon_6 \left(\frac{2}{6}e^{-\iota\phi} + \frac{5}{6} - \frac{1}{6}e^{\iota\phi} \right). \end{aligned} \quad (3.23)$$

We note that M is made up of two pairs of three terms, corresponding to each of the ENO stencils associated with each of the flux terms. We now bound each of the terms that comprise M . Because $\text{Re} \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) = \frac{2}{6} \cos 2\phi - \frac{7}{6} \cos \phi + \frac{11}{6}$, and $-1 \leq \cos \phi \leq 1$, $\forall \phi$, we can write

$$\frac{1}{3} \leq \text{Re} \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) \leq \frac{10}{3}.$$

(In fact, the lower bound can be tightened to 95/96, but the proof is not sensitive to this value.)

Thus

$$\text{Re} \left| \epsilon_1 \left(\frac{2}{6}e^{-2\iota\phi} - \frac{7}{6}e^{-\iota\phi} + \frac{11}{6} \right) \right| \leq \frac{10}{3} |\epsilon_1|$$

Similarly we can bound the remaining terms of $M(\epsilon_1, \epsilon_2, \dots, \epsilon_6, \phi)$. Finally we can obtain an expression of the form

$$|\operatorname{Re} M(\epsilon_1, \epsilon_2, \dots, \epsilon_6, \phi)| \leq \Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m|, \quad (3.24)$$

where Γ_1 is a positive constant that is determined by the stencils. Applying the same analysis, we can write

$$|\operatorname{Im} M(\epsilon_1, \epsilon_2, \dots, \epsilon_6, \phi)| \leq \Gamma_2 \max_{1 \leq m \leq 6} |\epsilon_m|, \quad (3.25)$$

where Γ_2 is a positive constant that is determined by the stencils.

We now examine the real and imaginary parts of \tilde{z} :

$$\begin{aligned} \operatorname{Re} \tilde{z} &= -\frac{1}{30} \cos 3\phi + \frac{1}{5} \cos 2\phi - \frac{1}{2} \cos \phi + \frac{1}{3} \\ &= -\frac{1}{30} (4 \cos^3 \phi - 3 \cos \phi) + \frac{1}{5} (2 \cos^2 \phi - 1) - \frac{1}{2} \cos \phi + \frac{1}{3} \\ &= \frac{2}{15} (1 - \cos \phi)^3; \end{aligned} \quad (3.26)$$

$$\operatorname{Im} \tilde{z} = \frac{1}{30} \sin 3\phi - \frac{3}{10} \sin 2\phi + \frac{3}{2} \sin \phi. \quad (3.27)$$

Let $D_4 = [-\frac{31}{15}, 0] \times [-\frac{17}{6}, \frac{17}{6}]$, and let $\sigma_0 > 0$ be such that the rectangle $\sigma_0 D_4 \subseteq D_3 = [-\eta, 0] \times [-\hat{\mu}, \hat{\mu}]$ defined previously. We now use D_3 and D_4 to prove that $\sigma_0 D_2 \subseteq D_1$; i.e., the combination of an s -stage, first-order ERK method where $\tau_2 > \frac{1}{2}$ with WENO5 is linearly stable if $\frac{\Delta t}{\Delta x} \leq \sigma_0$. (Note that σ_0 may not be the same as the CFL number σ ; in fact we know only that $0 < \sigma_0 \leq \sigma$. However, this proves the existence of $\sigma > 0$.)

Using a similar analysis to the first part of this theorem, we conclude that, given σ_0 , we can choose ϕ small enough such that $|g| < 1$. In other words, $\exists \alpha > 0$, such that the scaled domain $\{-\sigma_0 z(\phi) : 0 \leq \phi \leq \alpha \text{ or } 2\pi - \alpha \leq \phi \leq 2\pi\} \subseteq D_1$. We now complete the proof by showing that $D_5 = \{-\sigma_0 z(\phi) : \alpha \leq \phi \leq 2\pi - \alpha\} \subseteq D_1$.

Using (3.24), (3.26), we obtain

$$\frac{2}{15} (1 - \cos \phi)^3 - \Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m| \leq \operatorname{Re} z(\phi) \leq \frac{2}{15} (1 - \cos \phi)^3 + \Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m|.$$

Because $\alpha \leq \phi \leq 2\pi - \alpha$, we see

$$\frac{2}{15} (1 - \cos \alpha)^3 - \Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m| \leq \operatorname{Re} z(\phi) \leq \frac{16}{15} + \Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m|.$$

Note the ϵ_m , $m = 1, 2, \dots, 6$, are $\mathcal{O}((\Delta x)^2)$, and Γ_1 is a constant. We can choose Δx small enough such that $\Gamma_1 \max_{1 \leq m \leq 6} |\epsilon_m| \leq \min(\frac{2}{15} (1 - \cos \alpha)^3, 1)$. Therefore,

$$0 \leq \operatorname{Re} z(\phi) \leq \frac{16}{15} + 1 = \frac{31}{15}. \quad (3.28)$$

From (3.27), using the fact that $-1 \leq \sin \phi \leq 1$, $\forall \phi$, we see that $-\frac{11}{6} \leq \operatorname{Im} \tilde{z} \leq \frac{11}{6}$. Again, we can choose Δx small enough such that $\Gamma_2 \max_{1 \leq m \leq 6} |\epsilon_m| \leq 1$. Using (3.25), (3.27), we have

$$-\frac{17}{6} = -\frac{11}{6} - 1 \leq \operatorname{Im} z(\phi) \leq \frac{11}{6} + 1 = \frac{17}{6}. \quad (3.29)$$

From (3.28), (3.29), we conclude that $D_5 \subseteq D_4$. Because $\sigma_0 D_4 \subseteq D_3 = [-\eta, 0] \times [-\hat{\mu}, \hat{\mu}]$, and $D_3 \subseteq D_1$, we conclude that $\sigma_0 D_5 \subseteq D_1$.

This completes the proof. \square

COROLLARY 3.8. *The combination of WENO5 and any optimal, s -stage, first-order SSP ERK method as in [21] is linearly unstable.*

Proof. The Butcher tableau of the optimal, s -stage, first-order SSP ERK method is of the form shown in Table 3.2.

0					
$\frac{1}{s}$	$\frac{1}{s}$				
$\frac{2}{s}$	$\frac{1}{s}$	$\frac{1}{s}$			
$\frac{3}{s}$	$\frac{1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$		
\vdots	\vdots	\vdots	\vdots	\ddots	
$\frac{s-1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$	\dots	$\frac{1}{s}$
	$\frac{1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$	\dots	$\frac{1}{s}$ $\frac{1}{s}$

TABLE 3.2

Butcher tableau for optimal s -stage, order-1 SSP ERK methods.

The corresponding stability function is

$$\left(1 + \frac{1}{s}\hat{z}\right)^s = 1 + \hat{z} + \frac{s-1}{2s}\hat{z}^2 + \dots$$

Because $\frac{s-1}{2s} < \frac{1}{2}$ for all $s \geq 1$, linear instability follows from Theorem 3.7. \square

Using the same analysis of Theorem 3.7, we obtain the following theorem for s -stage, second-order ERK methods.

THEOREM 3.9. *Assume the ERK method is at least second order. Its stability function is of the form*

$$1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \tau_3\hat{z}^3 + \tau_4\hat{z}^4 + \tau_5\hat{z}^5 + \tau_6\hat{z}^6 + \dots \tag{3.30}$$

Then there exists a CFL number σ such that the combination of this ERK method and WENO5 is linearly stable for $\Delta t/\Delta x \leq \sigma$ if the ERK method satisfies

$$\tau_3 - \tau_4 > \frac{1}{8}; \tag{3.31}$$

on the other hand, the combination is linearly unstable if

$$\tau_3 - \tau_4 < \frac{1}{8}.$$

Proof. First we note that (3.30) is the amplification factor g if $\hat{z} = -\sigma z$. Now choosing ϕ to be a small

positive number, we use (3.11), and have

$$\begin{aligned}
z &= \left(\frac{7}{60} \phi^6 + \mathcal{O}(\phi^8) \right) + \iota (\phi + \mathcal{O}(\phi^7)), \\
z^2 &= (-\phi^2 + \mathcal{O}(\phi^8)) + \iota \mathcal{O}(\phi^7), \\
z^3 &= \mathcal{O}(\phi^8) + \iota (-\phi^3 + \mathcal{O}(\phi^9)), \\
z^4 &= (\phi^4 + \mathcal{O}(\phi^{10})) + \iota \mathcal{O}(\phi^9), \\
z^5 &= \mathcal{O}(\phi^{10}) + \iota (\phi^5 + \mathcal{O}(\phi^{11})), \\
z^6 &= (-\phi^6 + \mathcal{O}(\phi^{12})) + \iota \mathcal{O}(\phi^{11}),
\end{aligned}$$

and $z^l = \mathcal{O}(\phi^7)$, $l \geq 7$. We now calculate $|g|^2$.

$$\begin{aligned}
|g|^2 &= \left(1 - \frac{1}{2} \sigma^2 \phi^2 + \tau_4 \sigma^4 \phi^4 - \frac{7}{60} \sigma \phi^6 - \tau_6 \sigma^6 \phi^6 + \mathcal{O}(\phi^8) \right)^2 + (-\sigma \phi + \tau_3 \sigma^3 \phi^3 - \tau_5 \sigma^5 \phi^5 + \mathcal{O}(\phi^7))^2 \\
&= \left(1 - \sigma^2 \phi^2 + \left(\frac{1}{4} + 2\tau_4 \right) \sigma^4 \phi^4 - \frac{7}{30} \sigma \phi^6 - (\tau_4 + 2\tau_6) \sigma^6 \phi^6 + \mathcal{O}(\phi^8) \right) \\
&\quad + (\sigma^2 \phi^2 - 2\tau_3 \sigma^4 \phi^4 + (\tau_3^2 + 2\tau_5) \sigma^6 \phi^6 + \mathcal{O}(\phi^8)) \\
&= 1 + \left(\frac{1}{4} + 2\tau_4 - 2\tau_3 \right) \sigma^4 \phi^4 + \left(-\frac{7}{30} \sigma + (\tau_3^2 + 2\tau_5 - \tau_4 - 2\tau_6) \sigma^6 \right) \phi^6 + \mathcal{O}(\phi^8) \tag{3.32}
\end{aligned}$$

If $(\frac{1}{4} + 2\tau_4 - 2\tau_3) > 0$, i.e., $\tau_3 - \tau_4 < \frac{1}{8}$, we have $|g| > 1$ as $\phi \rightarrow 0+$. Therefore, in this case the combination is linearly unstable.

However, if $(\frac{1}{4} + 2\tau_4 - 2\tau_3) < 0$; i.e., $\tau_3 - \tau_4 > \frac{1}{8}$, we have $|g| < 1$ as $\phi \rightarrow 0+$. Moreover, there is an intersection between the stability domain of the ERK method and the imaginary axis. The rest of the proof of stability is similar to that of Theorem 3.7. \square

COROLLARY 3.10. *The combination of WENO5 with any optimal s -stage, second-order SSP ERK method as in [21] is linearly unstable.*

Proof. The Butcher tableau of the optimal s -stage, second-order SSP ERK method is of the form given in Table 3.3.

0					
$\frac{1}{s-1}$	$\frac{1}{s-1}$				
$\frac{2}{s-1}$	$\frac{1}{s-1}$	$\frac{1}{s-1}$			
$\frac{3}{s-1}$	$\frac{1}{s-1}$	$\frac{1}{s-1}$	$\frac{1}{s-1}$		
\vdots	\vdots	\vdots	\vdots	\ddots	
$\frac{s-1}{s-1}$	$\frac{1}{s-1}$	$\frac{1}{s-1}$	$\frac{1}{s-1}$	\dots	$\frac{1}{s-1}$
	$\frac{1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$	\dots	$\frac{1}{s}$

TABLE 3.3
Butcher tableau for optimal s -stage, second-order SSP ERK methods.

The corresponding stability function is

$$\frac{1}{s} + \frac{s-1}{s} \left(1 + \frac{1}{s-1} \hat{z} \right)^s = 1 + \hat{z} + \frac{1}{2} \hat{z}^2 + \frac{s-2}{6(s-1)} \hat{z}^3 + \frac{(s-2)(s-3)}{24(s-1)^2} \hat{z}^4 + \dots$$

Because $\tau_3 - \tau_4 = \frac{s(s-2)}{(s-1)^2} \left(\frac{1}{8} - \frac{1}{24s} \right) < \frac{1}{8}$, the linear instability follows from Theorem 3.9. \square

Note that the stability function of any three-stage, third-order ERK method is $1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3$. Because $\tau_3 - \tau_4 = \frac{1}{6} > \frac{1}{8}$, the combination of WENO5 and any three-stage, third-order ERK method is linearly stable. Furthermore, we have the following theorem for s -stage, third-order ERK methods.

THEOREM 3.11. *Assume the ERK method is third order. Its stability function is of the form*

$$1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \tau_4\hat{z}^4 + \dots$$

Then there exists a CFL number σ such that the combination of WENO5 and the ERK method is linearly stable for $\Delta t/\Delta x \leq \sigma$ if the ERK method satisfies

$$\tau_4 < \frac{1}{24}; \tag{3.33}$$

on the other hand, the combination is linearly unstable if

$$\tau_4 > \frac{1}{24}.$$

When $\tau_4 = \frac{1}{24}$, the ERK method has the same linear stability as the fourth-order ERK methods. The corresponding results are given in Theorem 3.12.

Proof. The proof is similar to that of Theorem 3.9. \square

Note if the ERK method is at least fourth order, we have $\tau_3 - \tau_4 = \frac{1}{8}$. From (3.32), we obtain

THEOREM 3.12. *Assume that the ERK method is at least order 4. Its stability function is of the form*

$$1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \frac{1}{24}\hat{z}^4 + \tau_5\hat{z}^5 + \tau_6\hat{z}^6 + \dots$$

Then there exists a CFL number σ such that the combination of the ERK method and WENO5 is linearly stable for $\Delta t/\Delta x \leq \sigma$ if the ERK method satisfies

$$\tau_5 - \tau_6 < \frac{1}{144}.$$

Proof. From (3.32) we conclude that $-\frac{7}{30}\sigma + (\tau_3^2 + 2\tau_5 - \tau_4 - 2\tau_6)\sigma^6 < 0$ for σ sufficiently small. Therefore, $|g|^2 < 1$ for ϕ sufficiently small and positive. In other words, $\exists \sigma_0, \alpha > 0$, such that the scaled domain $\{-\sigma_0 z(\phi) : 0 \leq \phi \leq \alpha \text{ or } 2\pi - \alpha \leq \phi \leq 2\pi\}$ is inside the stability domain D_1 of the ERK method.

The stability function of the ERK method is now of the form

$$g(\hat{z}) = 1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \frac{1}{24}\hat{z}^4 + \tau_5\hat{z}^5 + \tau_6\hat{z}^6 + \dots$$

Let $\hat{z} = \iota\gamma$, where γ is a small real number. In other words, choose \hat{z} to be close to the origin and on the imaginary axis. It is easy to see

$$|g|^2 = 1 + 2 \left(\tau_5 - \tau_6 - \frac{1}{144} \right) \gamma^6 + \mathcal{O}(\gamma^8).$$

Using the condition $\tau_5 - \tau_6 < \frac{1}{144}$, we obtain that $|g(\iota\gamma)| < 1$ as $\gamma \rightarrow 0+$. In other words, there is an intersection between the stability domain D_1 of the ERK method and the imaginary axis if $\tau_5 - \tau_6 < \frac{1}{144}$. The rest of the proof is similar to the proof of linear stability for second-order ERK methods from Theorem 3.9. \square

From Theorem 3.12, we see that $\tau_5 = \tau_6 = 0$ for the classical four-stage, fourth-order ERK method. Therefore, its combination with WENO5 is linearly stable.

REMARK 8. *There is no corresponding result for linear instability as in Theorem 3.9 if the ERK method is fourth order. This is because, regardless the values of τ_5 and τ_6 , $|g|^2 < 1$ whenever ϕ is a small positive number, and σ is sufficiently small. On the other hand, if the stability domain D_1 of the ERK method does not intersect the imaginary axis (i.e., $\tau_5 - \tau_6 > \frac{1}{144}$), linear instability cannot be proved as in Theorem 3.9.*

Following immediately from Theorem 3.12 with $\tau_5 = \frac{1}{120}$, we have the following theorem for s -stage, fifth-order ERK methods.

THEOREM 3.13. *Assume that the ERK method is at least order 5. Its stability function is of the form*

$$1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \frac{1}{24}\hat{z}^4 + \frac{1}{120}\hat{z}^5 + \tau_6\hat{z}^6 + \dots$$

Then there exists a CFL number σ such that the combination of the ERK method and WENO5 is linearly stable for $\Delta t/\Delta x \leq \sigma$ if the ERK method satisfies

$$\tau_6 > \frac{1}{720}.$$

Finally we have the more general result for s -stage, order- $p \geq 4$ ERK methods as follows.

THEOREM 3.14. *Assume that the ERK method is at least order $p \geq 4$. Its stability function is of the form*

$$1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \frac{1}{24}\hat{z}^4 + \dots$$

Then there exists a CFL number σ such that the combination of the ERK method and WENO5 is linearly stable for $\Delta t/\Delta x \leq \sigma$ if the stability domain of the ERK method includes the part of the imaginary axis of the form $[-\iota\mu, \iota\mu]$, for some $\mu > 0$.

The proof is similar that of Theorem 3.12.

REMARK 9. *We note that the presence of an intersection of the stability domain of an ERK method with the imaginary axis of the form $[-\iota\mu, \iota\mu]$, for some $\mu > 0$ has been used to prove stability of s -stage methods of orders 1, 2, and 3 already. Combining this with the result in Theorem 3.14 allows us to conclude that the intersection of the stability domain with the imaginary axis in the form $[-\iota\mu, \iota\mu]$, for some $\mu > 0$ is a sufficient condition for linear stability of any ERK method when coupled with WENO5.*

4. Numerical Results. In this section, we study two classical scalar hyperbolic conservation laws: the (linear) advection equation and the (nonlinear) inviscid Burgers equation. In both cases, the problems are posed in one dimension, and WENO5 is employed as the spatial discretization. We use a uniform mesh with N spatial subintervals. Both SSP and non-SSP ERK time integration methods are considered. We illustrate the linear instability of some well-known first- and second-order SSP ERK methods by plotting the solution at a given time T_{out} with a specified Courant number, $\sigma = \frac{\Delta t}{\Delta x}$. Extensive numerical tests have shown that smaller values of σ require larger values of T_{out} for the effect of the instability to clearly manifest itself.

4.1. ERK methods. In order to illustrate our theory, we consider the following four well-known ERK methods. Under appropriate assumptions, these methods can be SSP. The first three are linearly unstable when coupled with WENO5 and used to solve hyperbolic conservation laws; the fourth is arguably the most widely used time integration method used with WENO spatial discretizations.

- (1) The forward Euler (FE) method.
- (2) The optimal two-stage, second-order SSP ERK method (which we call SSP(2,2)) with Butcher tableau (cf. Table 3.1)

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

- (3) The optimal three-stage, second-order SSP ERK method [21] (which we call SSP(3,2)) with Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}$$

- (4) The (optimal) three-stage, third-order SSP ERK method, SSP(3,3) [3, 21], with Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \hline & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array}$$

We also consider four ERK methods that are linearly stable according to our analysis when coupled with WENO5. These methods are provably not SSP for any time step $\Delta t > 0$. Stated differently, these methods have a radius of contractivity of 0 (see, e.g., [12]); this fact is obvious because each method has a 0 in its \mathbf{b} vector (or *quadrature weights*) of the Butcher tableau.

- (5) A two-stage, order-1 non-SSP ERK method (which we call NSSP(2,1)) with Butcher tableau

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{3}{4} & \frac{3}{4} & 0 \\ \hline & 0 & 1 \end{array}$$

Because $b_2a_{21} = \frac{3}{4}$, it is a stable ERK method for WENO5 according to Theorem 3.7. Its CFL number can be directly estimated to be $\sigma = 0.80$; see e.g., [9, p.150].

- (6) A three-stage, second-order non-SSP ERK method (which we call NSSP(3,2)) with Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \hline & \frac{1}{2} & 0 & \frac{1}{2} \end{array}$$

It is easy to show that the linear stability function (and hence the amplification factor) of this ERK method is the same as SSP(3,3) and indeed all three-stage, third-order ERK methods. According to our analysis, it has the same linear stability properties, and in particular CFL number $\sigma = 1.43$, as SSP(3,3); see [11].

- (7) A three-stage, third-order non-SSP ERK method (which we call NSSP(3,3)) with Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ -\frac{4}{9} & -\frac{4}{9} & 0 & 0 \\ \frac{2}{3} & \frac{7}{6} & -\frac{1}{2} & 0 \\ \hline & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

This method has negative coefficients. It is sometimes necessary to specially treat right-hand side function evaluations $\hat{f}(u)$ that correspond to negative coefficients, e.g., by downwinding [18, 17]. However, according to our analysis, the linear stability properties of NSSP(3,3) are identical to SSP(3,3) (and all other three-stage, third-order ERK methods). We show that even for a nonlinear problem (Example 2), it has the same stability performance as SSP(3,3).

- (8) A five-stage, third-order non-SSP ERK method (which we call NSSP(5,3)) with Butcher tableau

$$\begin{array}{c|ccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ \frac{3}{16} & 0 & \frac{3}{16} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & 0 & 0 & \frac{2}{3} & 0 \\ \hline & \frac{1}{4} & 0 & 0 & 0 & \frac{3}{4} \end{array}$$

This is a new, low-storage ERK method, whose stability function is $1 + \hat{z} + \frac{1}{2}\hat{z}^2 + \frac{1}{6}\hat{z}^3 + \frac{1}{32}\hat{z}^4 + \frac{1}{224}\hat{z}^5$. It is a linearly stable ERK method for WENO5 according to Theorem 3.11. Its CFL number can be directly estimated to be $\sigma = 2.56$. Defining the effective CFL number to be $\frac{\sigma}{s}$, it is easy to see that a larger effective CFL number leads to more efficient time integration. The effective CFL number is 0.512 for NSSP(5,3), which is larger than 0.477 for SSP(3,3). We choose this scheme to illustrate that the theoretical principles described in this paper give us the ability to develop more efficient schemes than the popular SSP(3,3). We report on the results for the optimal ERK schemes for WENO5 elsewhere.

EXAMPLE 1. *The first example is the linear advection equation*

$$u_t + u_x = 0, \quad 0 < x < 2, \quad t > 0,$$

with periodic boundary conditions. We consider three different initial conditions: (a) the smooth initial condition, $u(x,0) = \sin(\pi x)$; (b) the smooth but more spatially varying initial condition, $u(x,0) = \sin^9(\pi x)$; (c) the discontinuous initial condition,

$$u(x,0) = \begin{cases} 1, & \text{if } 0 < x < 0.5 \text{ or } 1.5 < x < 2, \\ 0, & \text{if } 0.5 \leq x \leq 1.5. \end{cases} \quad (4.1)$$

(a) Figures 4.1 and 4.2 show the performance of FE and NSSP(2,1) for the problem with smooth initial condition $u(x,0) = \sin(\pi x)$. The solid lines in the figures are the exact solutions, and the dashed lines are the computed solutions. Both solutions are computed with $N = 200$, $\sigma = 0.5$, and they are plotted at $T_{out} = 2$.

As expected, spurious oscillations due to linear instability are present when the FE method is used, whereas there is no instability exhibited for NSSP(2,1). Although we only show the numerical result for $\sigma = 0.5$, we emphasize that the linear instability of the FE method appears for every $\sigma > 0$, no matter how small. We

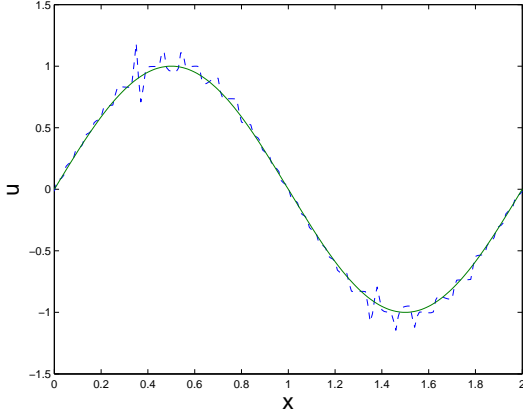


FIG. 4.1. FE for $u(x,0) = \sin(\pi x)$.

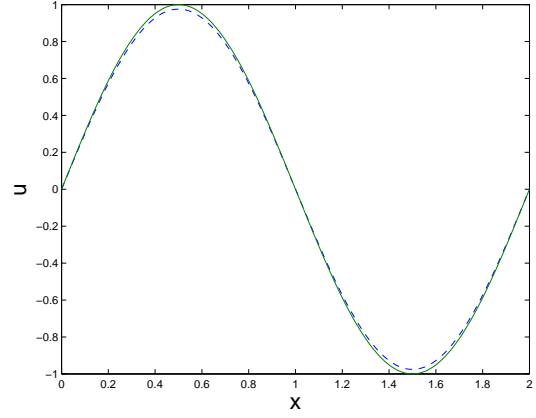


FIG. 4.2. N SSP(2,1) for $u(x,0) = \sin(\pi x)$.

observe the linear instability of the FE method for any of the later problems, whereas N SSP(2,1) is stable when $\sigma \leq 0.8$, which agrees with our expectation. However, the dissipation of N SSP(2,1) is very strong; this makes it unsuitable for computation in practice.

We now show an example of the non-convergence effect of linear instability. Figures 4.3 and 4.4 show the result of the numerical integration using WENO5 coupled with SSP(2,2) at two different output times. The solutions are computed with $N = 200$ and $\sigma = 1.32$. Figure 4.3 gives the solution at $T_{out} = 16$, while Figure 4.4 gives it at $T_{out} = 25$. We note that oscillations are generated almost immediately at the start of the integration. By $T_{out} = 16$, Figure 4.3 shows significant oscillation. However, the solution does not become unbounded, as can be seen in Figure 4.4 for $T_{out} = 25$; i.e., WENO5 has successfully adapted to the oscillations, not allowing them to become unbounded. However, with $N = 1000$ and the same σ , the solution quickly becomes unbounded. In other words, the spurious oscillations confirm the linear instability, but due to the nonlinear nature of WENO5, the numerical solution does not necessarily become unbounded. However, we point out that the linear instability of the combined method precludes convergence to the true solution; i.e., irrespective of the long-term boundedness, the error of the numerical solution in such cases can at best be expected to be $\mathcal{O}(1)$.

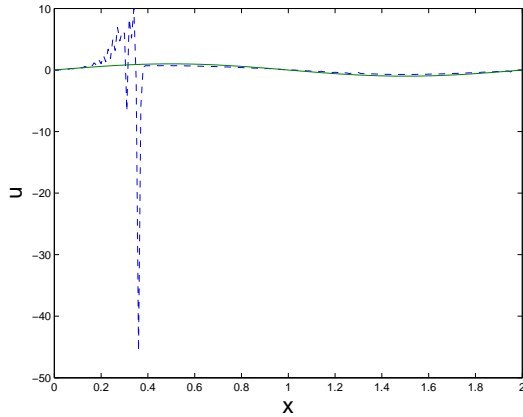


FIG. 4.3. SSP(2,2) for $u(x,0) = \sin(\pi x)$ at $T_{out} = 16$.

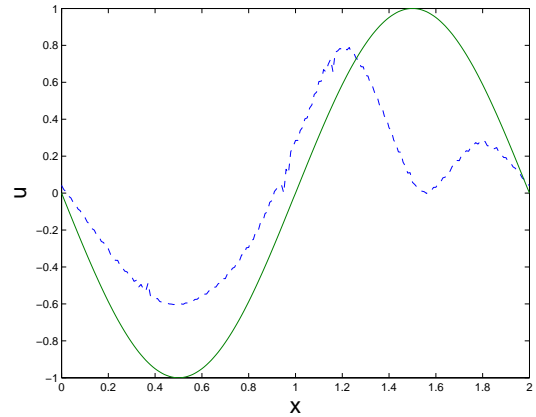


FIG. 4.4. SSP(2,2) for $u(x,0) = \sin(\pi x)$ at $T_{out} = 25$.

We now compare the performance of SSP(3,3) and N SSP(5,3). Figure 4.5 shows the performance of SSP(3,3) with $\sigma = 1.5$ for the problem just described, while Figure 4.6 shows the performance of N SSP(5,3) with $\sigma = 2.5$. The solid lines in the figures are the exact solutions, and the dashed lines are the computed solutions. Both solutions are computed with $N = 200$, and plotted at $T_{out} = 30$. We choose the two CFL numbers to make the computational costs equal for both experiments. Note that the difference between the

computed solution by $NSSP(5,3)$ and the exact solution is negligible. On the other hand, spurious oscillations appear for $SSP(3,3)$ because it is linearly stable only for CFL numbers less than about 1.43. Once again we note that the solution does not become unbounded. Another interesting observation is that when we choose $\sigma = 1.4$ for $SSP(3,3)$, the difference between the computed solution and the exact solution is also negligible. This means that the solution by $SSP(3,3)$ with $\sigma = 1.5$ is inaccurate due only to linear instability. The experiment clearly favours $NSSP(5,3)$ for its larger effective CFL number.

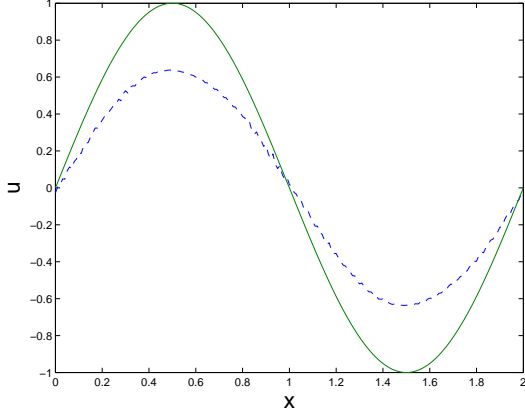


FIG. 4.5. $SSP(3,3)$ with $\sigma = 1.5$ for $u(x,0) = \sin(\pi x)$.

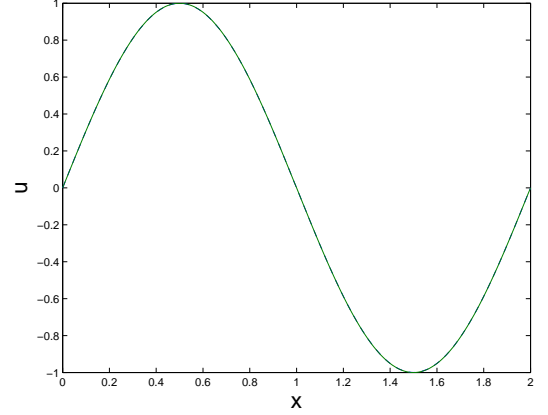


FIG. 4.6. $NSSP(5,3)$ with $\sigma = 2.5$ for $u(x,0) = \sin(\pi x)$.

Our experiments show that when $SSP(2,2)$ and $SSP(3,2)$ are used for the above problem with a small value for σ , the instability requires a long time to develop. In many cases the oscillations are not conspicuous in a relatively short time.

(b) We can introduce more spatial difficulty by using $u(x,0) = \sin^9(\pi x)$. Figures 4.7, 4.8, 4.9, and 4.10 show the performance of $SSP(2,2)$, $SSP(3,2)$, $NSSP(3,2)$, and $NSSP(3,3)$, respectively. The solid lines in the figures are the exact solutions, and the dashed lines are the computed solutions. All solutions are computed with $N = 200$, $\sigma = 0.5$, and they are plotted at $T_{out} = 150$.

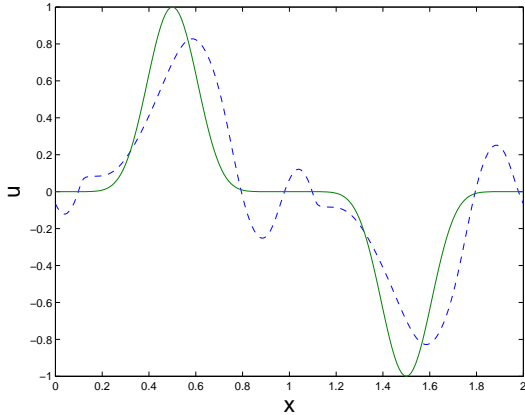


FIG. 4.7. $SSP(2,2)$ for $u(x,0) = \sin^9(\pi x)$.

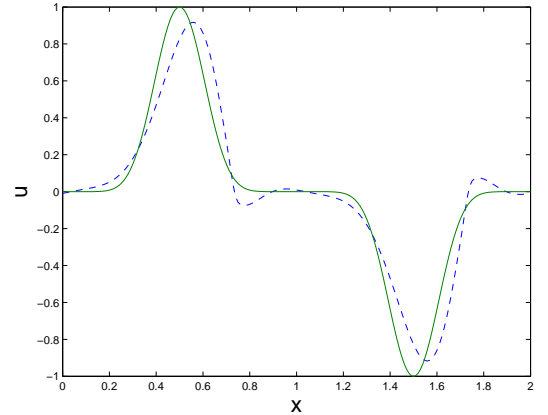


FIG. 4.8. $SSP(3,2)$ for $u(x,0) = \sin^9(\pi x)$.

The linear instability of WENO5 coupled with $SSP(2,2)$ or $SSP(3,2)$ takes a long time to become conspicuous for this problem, whereas $NSSP(3,2)$ and $NSSP(3,3)$ are stable. Moreover smaller values of σ tend to delay the manifestation of the instability even further. We also observe that the linear instability of the combination of WENO5 and $SSP(3,2)$ develops more slowly than that of WENO5 and $SSP(2,2)$. This can be explained by the fact that the classical (linear) stability domain of $SSP(3,2)$ includes the classical stability domain of $SSP(2,2)$.

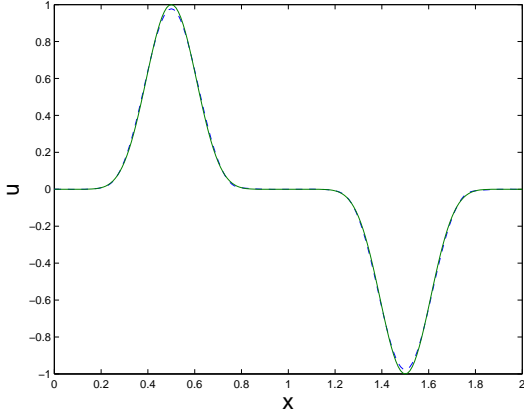


FIG. 4.9. $NSSP(3,2)$ for $u(x,0) = \sin^9(\pi x)$.

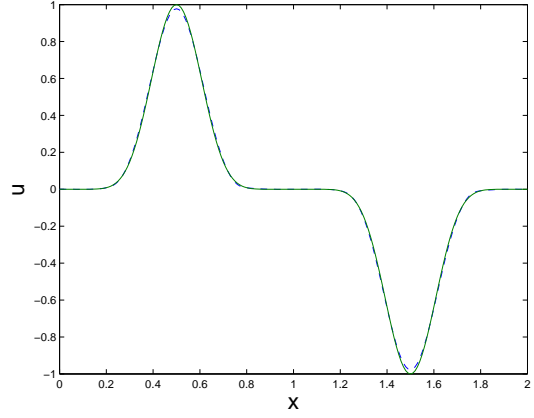


FIG. 4.10. $NSSP(3,3)$ for $u(x,0) = \sin^9(\pi x)$.

(c) The final initial condition is the step function (4.1). As we know, $WENO5$ is widely used in the numerical simulations of discontinuous solutions of hyperbolic PDEs. If the solution only has a few discontinuities and the wave speed is not zero, i.e., there is no stationary shock, the $WENO5$ stencil weights at any given point are the same as for the continuous case (3.2) for the majority of the time. That is, only the points close to the discontinuity use different discretizations because some stencil weights approach zero. However, the discretization at these points returns to those of the continuous case after the discontinuity passes through. Therefore, we expect that our analysis for the continuous case is also relevant to such discontinuous problems. Figures 4.11, 4.12, 4.13, and 4.14 show the performance of $SSP(2,2)$, $SSP(3,2)$, $NSSP(3,2)$, and $NSSP(3,3)$ respectively for the discontinuous initial condition (4.1). The solid lines in the figures are the exact solutions, and the dashed lines are the computed solutions. All solutions are computed with $N = 200$, $\sigma = 0.5$, and they are plotted at $T_{out} = 50$.

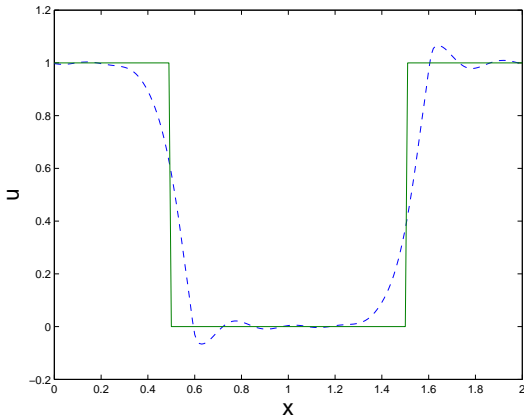


FIG. 4.11. $SSP(2,2)$ for (4.1).

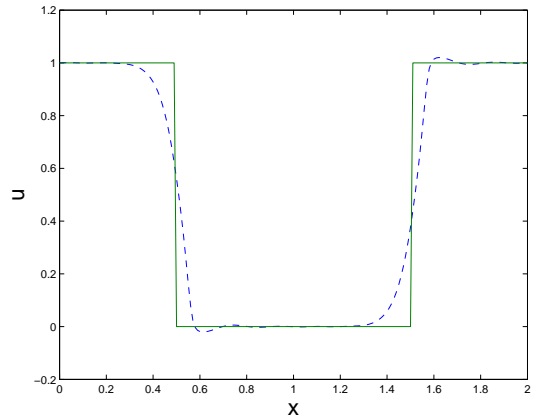


FIG. 4.12. $SSP(3,2)$ for (4.1).

We again make the observation that $NSSP(3,2)$ and $NSSP(3,3)$ are stable, whereas $SSP(2,2)$ and $SSP(3,2)$ exhibit oscillations.

We now compare $SSP(3,3)$ with $NSSP(5,3)$ for the discontinuous initial condition (4.1). Figures 4.15 and 4.16 show the performance of $SSP(3,3)$ with $\sigma = 1.11$, and $NSSP(5,3)$ with $\sigma = 1.85$. We choose the two CFL numbers to make the computational costs equal for both experiments. The solid lines in the figures are the exact solutions, and the dashed lines are the computed solutions. All solutions are computed with $N = 200$, and they are plotted at $T_{out} = 10$. Spurious oscillations appear when we use $SSP(3,3)$ with $\sigma = 1.11$, whereas there is no problem with $NSSP(5,3)$ with $\sigma = 1.85$. These numerical experiments show that $WENO5$ coupled with $SSP(3,3)$ is stable when $\sigma \leq 1$ for the discontinuous initial condition (4.1), whereas $NSSP(5,3)$ is stable

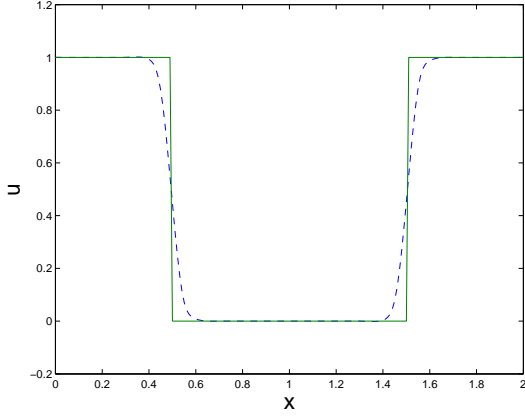


FIG. 4.13. $NSSP(3,2)$ for (4.1).

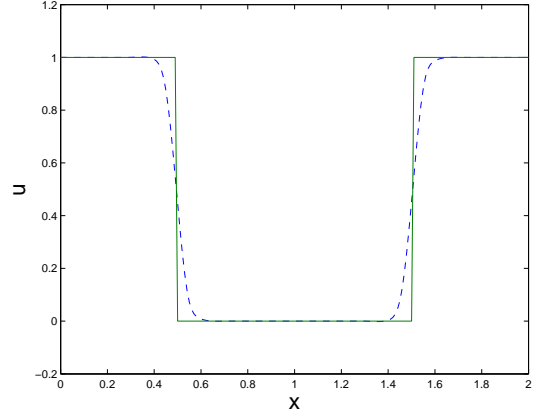


FIG. 4.14. $NSSP(3,3)$ for (4.1).

when $\sigma \leq 1.9$. This example shows that larger time steps can be used with $NSSP(5,3)$ than with $SSP(3,3)$ even when the solutions have discontinuities.

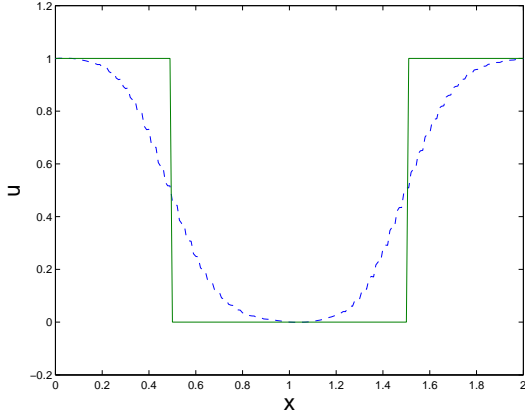


FIG. 4.15. $SSP(3,3)$ with $\sigma = 1.11$ for (4.1).

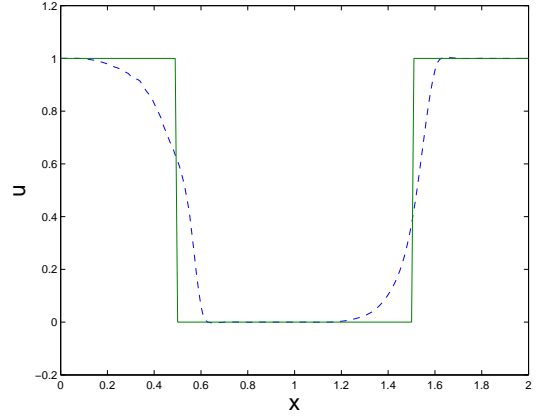


FIG. 4.16. $NSSP(5,3)$ with $\sigma = 1.85$ for (4.1).

EXAMPLE 2. The second example is the inviscid Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad 0 < x < 2, \quad t > 0,$$

with periodic boundary conditions and the initial condition $u(x, 0) = 2 + \sin^9(\pi x)$. This is a nonlinear problem with $f(u) = u^2/2$. Thus $\partial f/\partial u = u$. For a given σ , Δt is chosen as $\Delta t = \sigma \Delta x(\max_j u_j)$. Figures 4.17, 4.18, 4.19, and 4.20 show the performance of $SSP(2,2)$, $SSP(3,2)$, $NSSP(3,2)$, and $NSSP(3,3)$ respectively. The solid lines in the figures represent the reference solution, which is generated using WENO5 with $SSP(3,3)$, $N = 1000$, and $\sigma = 0.5$. The dashed lines are the computed solutions, all of which have $N = 100$, $\sigma = 0.5$, and they are plotted at $T_{out} = 40$. Again we see that $NSSP(3,2)$ and $NSSP(3,3)$ are stable, whereas $SSP(2,2)$ and $SSP(3,2)$ exhibit oscillations. This is a compelling illustration that our analysis accurately predicts the linear stability of ERK methods even with negative coefficients and even when applied to a nonlinear problem whose solution develops a discontinuity.

5. Conclusions. In this paper we employ a linear stability analysis for ERK time integration methods coupled with the WENO5 spatial discretization. We prove the forward Euler method, all two-stage, second-order ERK methods, and all optimal SSP ERK methods of up to second order are linearly unstable when

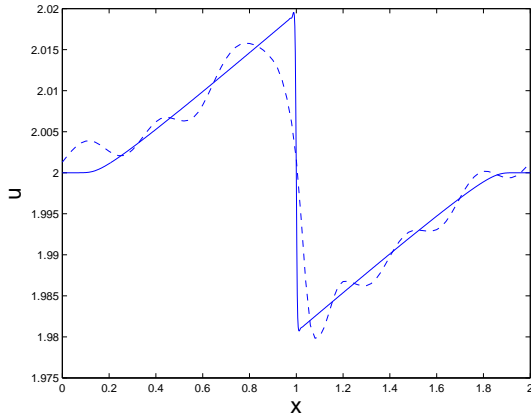


FIG. 4.17. $SSP(2,2)$ for the Burgers equation.

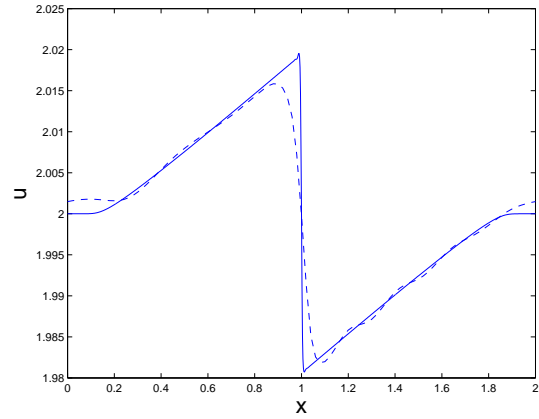


FIG. 4.18. $SSP(3,2)$ for the Burgers equation.

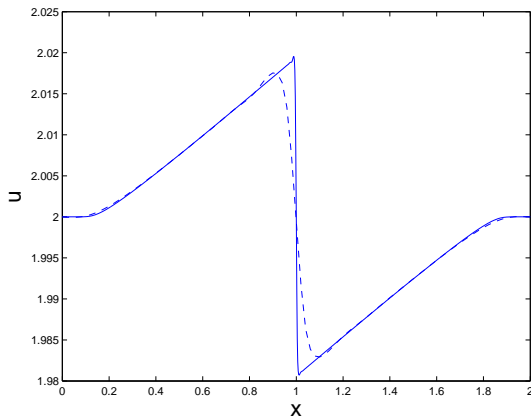


FIG. 4.19. $NSSP(3,2)$ for the Burgers equation.

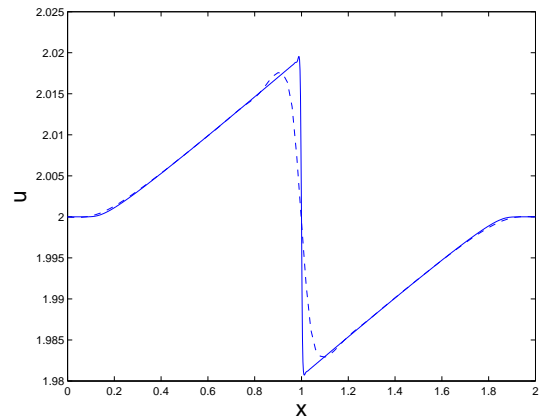


FIG. 4.20. $NSSP(3,3)$ for the Burgers equation.

coupled with WENO5 and used for solving hyperbolic conservation laws. Hence all of these combined methods are also not convergent. Moreover we show that, in our analysis, the success of the popular $SSP(3,3)$ method is not due to the SSP property; indeed all three-stage, third-order ERK methods, including those with negative coefficients or those that are provably non-SSP, have precisely the same linear stability performance according to our analysis, and this has translated to very similar performance in the examples we have presented. New stability criteria are derived for general ERK methods *of any order*. Numerical experiments demonstrate that although the analysis is strictly valid only for linear constant-coefficient problems with continuous initial conditions and periodic boundary conditions, it is relevant to both linear and nonlinear problems with continuous and discontinuous solutions. It is also relevant to ERK methods with negative coefficients without a special treatment (downwinding) of the spatial operator. For linear stability of an ERK time integration method coupled with WENO5, we show that it is sufficient that the classical linear stability region of the ERK method include a piece of the imaginary axis. The analysis techniques described in this paper apply to other WENO methods such as the seventh- or higher-order WENO methods. From this analysis it is also possible to derive optimal ERK methods in terms of the CFL number for WENO5. In particular it is possible to derive methods such as $NSSP(5,3)$ that are more efficient (i.e., have larger effective CFL numbers) than the benchmark method $SSP(3,3)$. We report on these results elsewhere.

Acknowledgments. The authors wish to express their thanks to the editor and the referees for their comments. R.J.S. also wishes to thank I. Higuera and L. Ferracina for useful discussions.

REFERENCES

- [1] J. BLAZEK, *Computational fluid dynamics: principles and applications*, Elsevier Science Ltd, Oxford, 2001.
- [2] L. FERRACINA AND M. N. SPIJKER, *Stepsize restrictions for the total-variation-diminishing property in general Runge-Kutta methods*, SIAM J. Numer. Anal., 42 (2004), pp. 1073–1093 (electronic).
- [3] S. GOTTLIEB, C.-W. SHU, AND E. TADMOR, *Strong stability-preserving time discretization methods*, SIAM Review, 43 (2001), pp. 89–112.
- [4] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving ordinary differential equations. I. Nonstiff problems*, Springer Series in Computational Mathematics, 8. Springer-Verlag, Berlin, 1987.
- [5] E. HAIRER AND G. WANNER, *Solving ordinary differential equations. II. Stiff and differential-algebraic problems*, Springer Series in Computational Mathematics, 14. Springer-Verlag, Berlin, 1991.
- [6] A. HARTEN, B. ENGQUIST, S. OSHER, AND S. CHAKRAVARTHY, *Uniformly high order Essentially Non-Oscillatory schemes III*, J. Comput. Phys., 71 (1987), pp. 231–303.
- [7] A. HARTEN AND S. OSHER, *Uniformly high order Essentially Non-Oscillatory schemes I*, SIAM J. Numer. Anal., 24 (1987), pp. 270–309.
- [8] I. HIGUERAS, *On strong stability preserving time discretization methods*, J. Sci. Comput., 21 (2004), pp. 193–223.
- [9] W. HUNSDORFER AND J. VERWER, *Numerical solution of time-dependent advection-diffusion-reaction equations*, vol. 33 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2003.
- [10] A. JAMESON, *Solution of the Euler equations for two dimensional transonic flow by a multigrid method*, Appl. Math. Comput., 13 (1983), pp. 327–355.
- [11] G.-S. JIANG AND C.-W. SHU, *Efficient implementation of Weighted ENO schemes*, J. Comput. Phys., 126 (1996), pp. 202–228.
- [12] J. F. B. M. KRAAIJEVANGER, *Contractivity of Runge-Kutta methods*, BIT, 31 (1991), pp. 482–528.
- [13] H.-O. KREISS AND G. SCHERER, *Method of lines for hyperbolic differential equations*, SIAM J. Numer. Anal., 29 (1992), pp. 640–646.
- [14] D. LEVY AND E. TADMOR, *From semidiscrete to fully discrete: stability of Runge-Kutta schemes by the energy method*, SIAM Review, 40 (1998), pp. 40–73.
- [15] X.-D. LIU, S. OSHER, AND T. CHAN, *Weighted essentially non-oscillatory schemes*, J. Comput. Phys., 115 (1994), pp. 200–212.
- [16] S. OSHER AND R. FEDKIW, *Level set methods and dynamic implicit surfaces*, Applied Mathematical Sciences. Springer-Verlag, New York, 2003.
- [17] 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.
- [18] C.-W. SHU, *Total-variation-diminishing time discretizations*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 1073–1084.
- [19] ———, *Numerical experiments on the accuracy of ENO and modified ENO schemes*, J. Sci. Comput., 5 (1990), pp. 127–149.
- [20] ———, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, in Advanced numerical approximation of nonlinear hyperbolic equations, Lecture Notes in Mathematics, vol. 1697, Berlin, 1998, Springer-Verlag, pp. 325–432.
- [21] R. J. SPITERI AND S. J. RUUTH, *A new class of optimal high-order strong-stability-preserving time discretization methods*, SIAM J. Numer. Anal., 40 (2002), pp. 469–491.
- [22] E. TADMOR, *Stability analysis of finite-difference, pseudospectral and Fourier-Galerkin approximations for time-dependent problems*, SIAM Review, 29 (1987), pp. 525–555.
- [23] J. W. THOMAS, *Numerical partial differential equations: Finite difference methods*, Texts in Applied Mathematical. Springer-Verlag, New York, 1995.
- [24] A. VANDE WOUWER, P. SAUCEZ, AND W. E. SCHIESSER, eds., *Adaptive method of lines*, Chapman-Hall/CRC, Boca Raton, 2001.
- [25] Z. XU AND C.-W. SHU, *Anti-diffusive flux corrections for high order finite difference WENO schemes*, J. Comput. Phys., 205 (2005), pp. 458–485.