

# Spurious Behavior of Shock-Capturing Methods by the Fractional Step Approach: Problems Containing Stiff Source Terms and Discontinuities<sup>☆</sup>

H. C. Yee<sup>a,\*</sup>, D. V. Kotov<sup>b</sup>, Wei Wang<sup>c</sup>, Chi-Wang Shu<sup>d</sup>

<sup>a</sup>NASA Ames Research Center, Moffett Field, CA, 94035

<sup>b</sup>Stanford Center for Turbulence Research, Stanford, CA 94305-3035

<sup>c</sup>Department of Mathematics & Statistics, Florida International University, Miami, FL, 33199, USA

<sup>d</sup>Division of Applied Mathematics, Brown University, Providence, RI, 02912, USA

---

## Abstract

The goal of this paper is to relate numerical dissipations that are inherited in high order shock-capturing schemes with the onset of wrong propagation speed of discontinuities. For pointwise evaluation of the source term, previous studies indicated that the phenomenon of wrong propagation speed of discontinuities is connected with the smearing of the discontinuity caused by the discretization of the advection term. The present study focuses only on solving the reactive system by the fractional step method using the Strang splitting. Studies shows that the degree of wrong propagation speed of discontinuities is highly dependent on the accuracy of the numerical method. The manner in which the smearing of discontinuities is contained by the numerical method and the overall amount of numerical dissipation being employed play major roles. Depending on the numerical method, time step and grid spacing, the numerical simulation may lead to (a) the correct solution (within the truncation error of the scheme), (b) a divergent solution, (c) a wrong propagation speed of discontinuities solution or (d) other spurious solutions that are solutions of the discretized counterparts but are not solutions of the governing equations. The findings might shed some light on the reported difficulties in numerical combustion and problems with stiff nonlinear (homogeneous) source terms and discontinuities in general.

*Keywords:* High order numerical methods, Numerical combustion, Chemical reacting flows, Nonequilibrium flows, Stiff source terms, Numerical methods for stiff source terms with shocks, Wrong propagation speed of discontinuities

---

## 1. Introduction

Consider 3D reactive Euler equations of the form

$$U_t + F(U)_x + G(U)_y + H(U)_z = S(U), \quad (1)$$

where  $U$ ,  $F(U)$ ,  $G(U)$ ,  $H(U)$  and  $S(U)$  are vectors. Here,  $S(U)$  is restricted to be homogeneous in  $U$ ; that is,  $(x, y, z)$  and  $t$  do not appear explicitly in  $S(U)$ . If the time scale of the ordinary differential equation (ODE)  $U_t = S(U)$  for the source term is orders of magnitude smaller than the time scale of the homogeneous conservation law  $U_t + F(U)_x + G(U)_y + H(U)_z = 0$ , then the problem is said to be stiff due to the source terms. In combustion or high speed chemical reacting flows the source term represents the chemical reactions which may be much faster than the gas flow. This leads to problems of numerical stiffness due to chemical reactions. Insufficient spatial/temporal resolution may cause an incorrect propagation speed of discontinuities and nonphysical states for standard dissipative numerical methods that were developed for non-reacting flows.

This numerical phenomenon was first observed by Colella et al. [6] in 1986 who considered both the reactive Euler equations and a simplified system obtained by coupling the inviscid Burgers equation with a single convection/reaction

---

<sup>☆</sup>Expanded version of the proceedings of ICCFD7, July 9-13, 2012, The Big Island, Hawaii.

\*Corresponding author

Email address: Helen.M.Yee@nasa.gov (H. C. Yee)

equation. LeVeque and Yee [24] showed that a similar spurious propagation phenomenon can be observed even with scalar equations, by properly defining a model problem with a stiff source term. They introduced and studied the simple one-dimensional scalar conservation law with an added nonhomogeneous parameter dependent source term

$$u_t + u_x = S(u), \quad (2)$$

$$S(u) = -\mu u(u - \frac{1}{2})(u - 1). \quad (3)$$

When the parameter  $\mu$  is very large, a wrong propagation speed of discontinuity phenomenon by dissipative numerical methods will be observed in coarse grids. In reacting flows,  $\frac{1}{\mu}$  can be described as the reaction time. In order to isolate the problem, LeVeque and Yee solved 2 and 3 by the fractional step method using the Strang splitting [37]. For this particular source term, the reaction (ODE) step of the fractional step method can be solved exactly. In their study using pointwise evaluation of the source term ( $S(u)$  is evaluated at the  $j$  grid point index, i.e.,  $S(u_j)$  for each time evolution), the phenomenon of wrong propagation speed of discontinuities is connected with the smearing of the discontinuity caused by the spatial discretization of the advection term. They found that the propagation error is due to the numerical dissipation contained in the scheme, which smears the discontinuity front and activates the source term in a nonphysical manner. The smearing introduces a nonequilibrium state into the calculation. Thus as soon as a nonequilibrium value is introduced in this manner, the source term turns on and immediately restores equilibrium, while at the same time shifting the discontinuity to a cell boundary. By increasing the spatial resolution by an order of magnitude, they were able to improve towards the correct propagation speed. It is remarked here that in a general stiff source term problem, a sufficient spatial resolution is as important as temporal resolution when the reaction step of the fractional step method cannot be solved exactly. As will be shown in the present study, on one hand, the degree of wrong propagation speed of discontinuities is highly dependent on the accuracy of the numerical method. On the other hand, the manner in which the smearing of discontinuities is contained by the numerical method and the overall amount of numerical dissipation being employed play major roles. Moreover, employing finite time steps and grid spacings that are below the standard Courant-Friedrich-Levy (CFL) limit on shock-capturing methods for compressible Euler and Navier-Stokes equations containing stiff reacting source terms and discontinuities reveals surprising counter-intuitive results.

Based on the work of [24, 48, 8, 20, 21, 56], in addition to the incorrect propagation speed of discontinuities, other spurious numerics, that are directly tied to the amount of numerical dissipation contained in the chosen scheme and the numerical treatment of source terms may result in

- **Possible spurious steady-state numerical solutions and spurious standing waves [48, 8, 20, 21]:** It was shown in Lafon & Yee [20, 21] and Griffiths et al. [48] that various ways of discretizing the nonlinear reaction terms can affect the stability of, and convergence to, the spurious numerical steady states and/or the exact steady states. Pointwise evaluation of the source terms appears to be the least stable. The studies of Lafon & Yee [20, 21] indicated that numerical phenomena of incorrect propagation speeds of discontinuities may be linked to the existence of some stable spurious steady-state numerical solutions. More importantly, the different combination of time step, grid spacing and initial condition plays a major role in obtaining the correct solution. In addition, it was shown in Yee et al. [48] and Griffiths et al. [8] that spurious discrete traveling waves can exist, depending on the method of discretizing the source term. Recently, Wang et al. [43] indicated that a well-balanced scheme for reacting flows can minimize certain spurious numerics.

Studies linking spurious numerical standing waves for Eq. (2) and Eq. (3) by first and second-order spatial and temporal discretizations can be found in Lafon and Yee [20, 21] and Griffiths, Stuart and Yee [8, 51].

- **Possible wrong prediction of transition point Reynolds number by DNS due to spurious bifurcation that created a false transition point:** Inaccuracy of the scheme or insufficient grid points might lead to possible spurious bifurcation as well as creating wrong propagation speed of discontinuities and smearing of turbulent fluctuations. See [51] for a discussion.

The term “spurious (numerical) solutions” here refers to computed solutions that are solutions of the discretized counterparts but are not solutions of the considered governing equation. Pointwise evaluation of the source term here

means that, for each time evolution,  $S(U)$  is evaluated at the single grid point  $S(U_{j,k,l})$ , where  $(j, k, l)$  is the grid point index.

For the last two decades, the wrong speed phenomenon has attracted a large volume of research work in the literature (see, e.g., [2, 30, 3, 39, 5, 51, 23, 13, 4, 26, 1, 10, 40]). Various strategies have been proposed to overcome this wrong speed difficulty for one to two species cases with a single reaction. Since numerical dissipation that spreads the discontinuity front is the cause of the wrong propagation speed of discontinuities, a natural strategy is to avoid any numerical dissipation in the scheme. In combustion, level set and front tracking methods were used to track the wave front to minimize this spurious behavior [23, 13, 4, 26]. See Wang et al. [42] for a comprehensive overview of the last two decades of development. Wang et al. also proposed a new high order finite difference method with subcell resolution for advection equations with stiff source terms for a single reaction for Eq. (eq:3D-reacting-Euler) to overcome the difficulty. Research for multi-species (3 or more species and multi-reactions) is forthcoming.

### 1.1. Objective

This is a follow on work to Wang, Shu, Yee & Sjögren, Yee, Kotov & Sjögren and related earlier work [42, 45, 48, 8, 20, 21]. The objective of this paper is to study spurious behavior of high order shock-capturing methods using the pointwise evaluation of stiff homogenous source terms for problems containing discontinuities. Pointwise evaluation is used in the current study in spite of the fact that Lafon & Yee [20, 21] and Griffiths et al. [8] indicated two decades ago that pointwise evaluation of the source term (for first and second-order schemes) appears to be the least stable. They suggested using non-pointwise evaluation of the source term that is more compatible with the convection difference operator. The current study presents a more in-depth understanding of the pointwise evaluation approach as the majority of the schemes in use for numerical combustion and problems containing stiff sources and discontinuities employ this approach. In addition, spurious behavior in this type of highly nonlinear coupling system cases using finite time steps and grid spacings is not fully understood.

Special focus is on the behavior of the recently developed high order finite difference method with subcell resolution [42], and the filter counterparts [47, 54] of the new high order subcell resolution method as time step and grid spacing are refined. The finite difference method with subcell resolution method solve the reactive system by the fractional step approach using the Strang splitting. The study also accounts for the scheme behavior as the stiffness of the source term increases. Early and less extensive study on the subject has been reported in [18]. Comparison with the performance of the Harten & Yee second-order TVD method [50, 44], and standard fifth-order and seventh-order WENO schemes (WENO5 and WENO7) [15] are included. Although the subcell resolution idea and its filter counterparts are applicable to any high order shock-capturing method, here the study is focused on the class of WENO schemes. From here on, the subcell resolution counterparts of WENO5 and WENO7 will be denoted by WENO5/SR and WENO7/SR [42], whereas their filter counterparts will be denoted by WENO5fi/SR and WENO7fi/SR.

### 1.2. Outline

The outline of this paper is as follows: A practical stiff hypersonic chemical nonequilibrium viscous computation is illustrated in Section 2 to motivate the current study. The high order methods with subcell resolution and their filter counterparts [42, 47, 54] are summarized in Section 3. The problem setup for the two stiff detonation test cases with numerical results comparing the performance among WENO5, WENO7, and the associated filter version of WENO5 (WENO5fi) [47, 54], WENO5/SR and WENO5fi/SR are then presented in Section 4. The present investigation for three very different system cases confirms the findings of Lafon & Yee and LeVeque & Yee for a model scalar PDE. In all of the computations, the classical fourth-order Runge-Kutta method (RK4) and the Roe flux with Roe's average states [31] are used. Performance using the third-order TVD Runge-Kutta [32] is similar but with a slightly smaller CFL limit. All the WENO schemes are the original form of Jiang & Shu [15], except for one case where the finite difference form of the recently developed positive WENO scheme [58] using the Lax-Friedrichs was tested.

## 2. Motivation: An Unsteady Nonequilibrium Navier-Stokes Computation [17]

In general, the reacting terms that arise from nonequilibrium flows in hypersonic aeronautics are less stiff than their counterparts in combustion. However, there are stiff chemical nonequilibrium flows that are due to the reaction

terms. Before the study of two stiff detonation test cases, a stiff 13-species, one temperature nonequilibrium model related to the NASA Ames Electric Arc Shock Tube (EAST) experiment is briefly investigated. Detailed study is reported in [17]. See [25] for a brief introduction and earlier simulations. The reason for this introductory example is to illustrate that it is unlike earlier work in [24, 48, 8, 20, 21], where detailed analysis using dynamical system theory were possible. A complex high Mach number and high temperature problem like EAST is very costly even for a 3D coarse grid complete unsteady simulation. The length of the EAST shock tube experiment is very long and the associated flow physics is multiscale with multi-reaction terms [25].

## 2.1. Governing Equations

In component form of 1, a 3D nonequilibrium Navier-Stokes system for the 8.5m (meter) EAST problem (with the thermo-nonequilibrium part neglected) for a preliminary study is given by:

$$\frac{\partial \rho_s}{\partial t} + \frac{\partial}{\partial x_j} (\rho_s u_j + \rho_s d_{sj}) = \Omega_s \quad (4)$$

$$\frac{\partial}{\partial t} (\rho u_i) + \frac{\partial}{\partial x_j} (\rho u_i u_j + p \delta_{ij} - \tau_{ij}) = 0 \quad (5)$$

$$\frac{\partial}{\partial t} E + \frac{\partial}{\partial x_j} \left[ u_j (E + p) + q_j + \sum_s \rho_s d_{sj} h_s - u_i \tau_{ij} \right] = 0, \quad (6)$$

where  $U = (\rho_s, \rho u_i, E)$  are the conservative variables,  $\rho_s$  are the partial densities with  $k = 1, \dots, N_s$  for a mixture of  $N_s$  species. Here  $i = 1, 2, 3$  for 3D.  $u_i, i = 1, 2, 3$  are the mixture  $x, y$  and  $z$ -velocities,  $E$  is the mixture total energy per unit volume,  $p$  is the pressure,  $K(T)$  is the chemical reaction rate and  $T$  is the temperature. The mixture total density, the pressure and the total energy per unit volume are

$$\rho = \sum_s \rho_s, \quad p = RT \sum_{s=1}^{N_s} \frac{\rho_s}{M_s}, \quad E = \sum_{s=1}^{N_s} \rho_s (e_s(T) + h_s^0) + \frac{1}{2} \rho v^2, \quad (7)$$

where  $R$  is the universal gas constant,  $h_s^0$  are the species formation enthalpies, and  $M_s$  indicates the species molar masses.

The viscous stress tensor is given by:

$$\tau_{ij} = \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \mu \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij}. \quad (8)$$

The diffusion flux is given by:

$$d_{sj} = -D_s \frac{\partial X_s}{\partial x_j}, \quad (9)$$

where  $D_s$  is the diffusion coefficient and  $X_s$  is the mole fraction of species  $s$ .

The conduction heat flux is given by:

$$q_j = -\lambda \frac{\partial T}{\partial x_j}, \quad (10)$$

where  $\lambda$  is the thermal conductivity of the mixture. The chemical source term is given by:

$$\Omega_s = M_s \sum_{r=1}^{N_r} (b_{s,r} - a_{s,r}) \left[ k_{f,r} \prod_{m=1}^{N_s} \left( \frac{\rho_m}{M_m} \right)^{a_{m,r}} - k_{b,r} \prod_{m=1}^{N_s} \left( \frac{\rho_m}{M_m} \right)^{b_{m,r}} \right], \quad (11)$$

where  $a$  and  $b$  are the stoichiometric coefficients, and the forward reaction rates  $k_{f,r}$  coefficients are given by Arrhenius' law:

$$k_{f,r} = A_{f,r} T^{n_{f,r}} \exp(-E_{f,r}/kT). \quad (12)$$

The backward reactions rates coefficients are computed as  $k_{b,r} = k_{f,r}/K_{c,r}^{eq}$ , where  $K_{c,r}^{eq}$  is the equilibrium constant.

$\rho$	1.10546 $kg/m^3$	$\rho$	$3.0964 \times 10^{-4} kg/m^3$
$T$	6000 $K$	$T$	300 $K$
$p$	12.7116 $MPa$	$p$	26.771 $Pa$
$Y_{He}$	0.9856	$Y_{O_2}$	0.21
$Y_{N_2}$	0.0144	$Y_{N_2}$	0.79

Table 1: High (left) and low (right) pressure region initial data

Due to the multiscale and multi-stiffness of the problem [25], numerical simulations for 1D ( $i = 1$ ) and 2D ( $i = 1, 2$ ) are considered first in [17]. Numerical study of grid size and numerical method dependence of the computed shear and shock locations as the grid is refined for 1D and 2D simplifications of the 3D EAST problem will be illustrated here. All the computations employ a multi-D high order single/overset grid nonequilibrium code ADPDIS3D [22]. Due to high computational cost, only single grid results for a very early stage of the unsteady flow development are presented. The desired simulation requires that the shock wave propagates to a 8.5 meter distance. The MUTATION library [29], developed by Thierry Magin and Marco Panesi, is used for the numerical experiment to provide reaction rate and transport properties. Here, for this viscous simulation, all the CFL values are based on the convection and viscous part of the PDEs.

## 2.2. 1D 13 Species EAST Test Case

The computational domain has a total length of 8.5m. The left part of the domain with length 0.1m is a high pressure region. The right part of the domain with length 8.4m is a low pressure region. The gas mixture consists of 13 species:

$$e^-, He, N, O, N_2, NO, O_2, N_2^+, NO^+, N^+, O_2^+, O^+, He^+.$$

The initial conditions of the high and low pressure regions are listed in the Table 1. For the left-side boundary the Euler (slip) wall condition is applied, and for the right-side, the zero gradient condition is applied for all variables.

Figure 1 shows the results from the computation using the Harten-Yee second-order TVD scheme [50, 44] for four grids with  $\Delta x = 10^{-3} m$ ,  $5 \times 10^{-4} m$ ,  $5 \times 10^{-5} m$  and  $2.5 \times 10^{-5} m$  at time  $t_{end} = 0.325 \times 10^{-4} sec$ . One can observe a significant shift in the shear (left discontinuity) and the shock (right discontinuity) locations as the grid is refined. The distance between the shear and the shock shrinks as the grid is refined. The difference between shock locations obtained on the grids with  $\Delta x = 5 \times 10^{-5} m$  and  $2.5 \times 10^{-5} m$  is less than 0.3%. Thus the solution using  $\Delta x = 5 \times 10^{-5} m$  can be considered as the reference solution.

The left subfigure of Fig. 2 shows a comparison among five methods obtained on a coarse grid ( $\Delta x = 10^{-3} m$ ) with the reference solution. The scheme's labels are defined as follows:

- ACMTVDfi: Second-order central base scheme using ACM flow sensor. See [52] for further information on filter schemes.
- WENO5-llf: Fifth-order WENO (WENO5) using the local Lax-Friedrichs flux.
- WENO5P-llf: Positive WENO5 of [58] using the local Lax-Friedrichs flux.
- WENO5PH-llf: Positive WENO5 of [12] using the local Lax-Friedrichs flux.

The right subfigure of Fig. 2 shows a comparison of ACMTVDfi using a different weight  $\kappa$  parameter of the ACM flow sensor. The smaller the  $\kappa$ , the smaller the amount of TVD dissipation that is used. Among the considered schemes, Fig. 2 indicates that the least dissipative scheme predicts the shear and shock locations best when compared with the reference solution. The results indicate that ACMTVDfi is slightly more accurate than WENO5-llf. This is due to the fact that ACMTVDfi reduces the amount of numerical dissipation away from high gradient regions. Using the subcell resolution method of [42] for one reaction case by applying it to only one of the reactions in this multireaction flow does not improve the performance over standard schemes. Further research on the generalization of subcell resolution to multi-reactions needs to be explored.

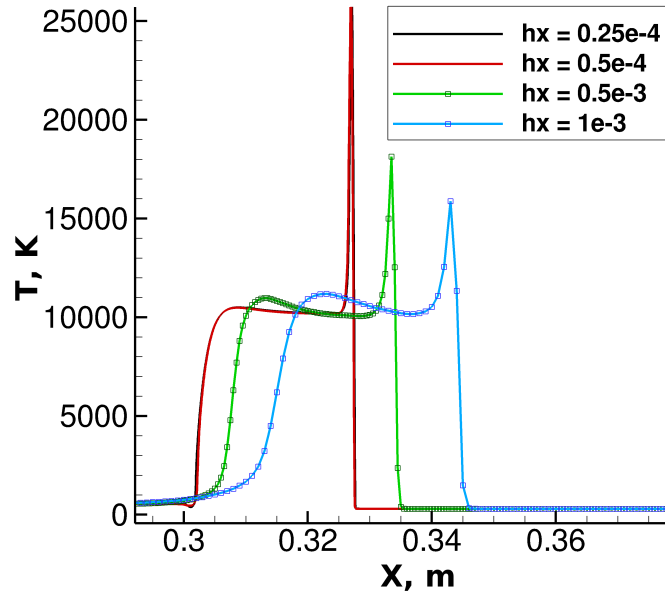


Figure 1: 13 species 1D EAST problem: Second-order Harten-Yee TVD simulation for three grids:  $\Delta x = 10^{-3} m$  (line 1),  $5 \times 10^{-4} m$  (line 2),  $5 \times 10^{-5} m$  (line 3),  $2.5 \times 10^{-5} m$  (line 4), and  $T_{end} = 0.325 \times 10^{-4} s$ , with  $CFL = 0.8$ .

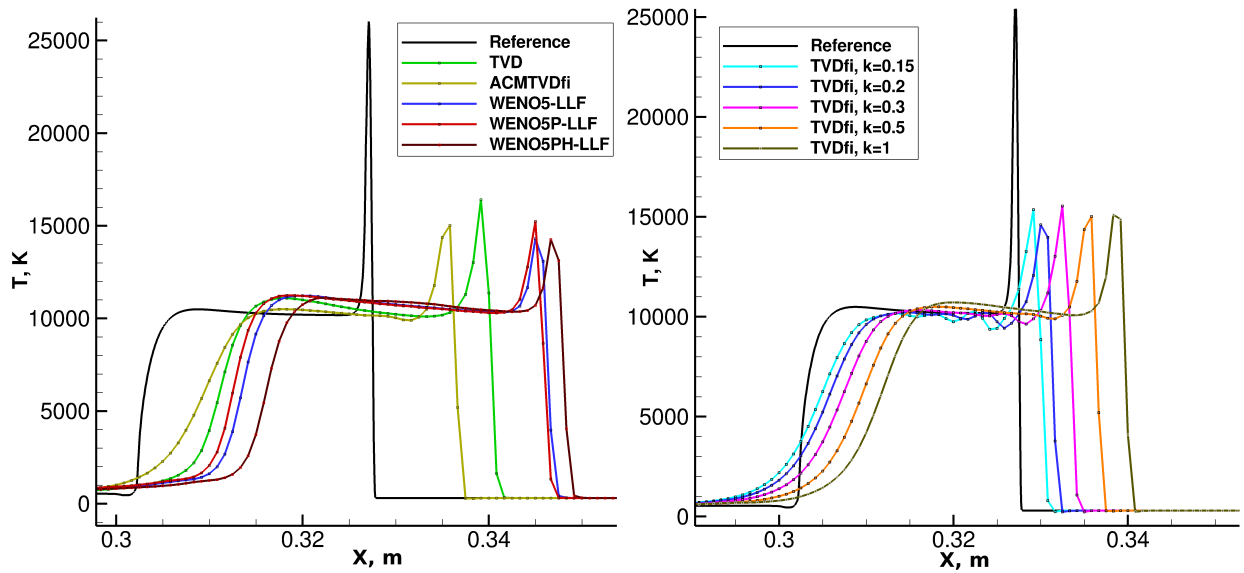


Figure 2: 1D, 13 species EAST problem: Comparison among methods using 601 point grids with  $CFL = 0.6$  and  $t_{end} = 3.25 \times 10^{-5} sec$ . Left subfigure: Reference solution (TVD on a 10,001 point grid) (line 1), TVD (line 2), ACMTVDfi (TVDfi) using  $\kappa = 0.5$  (line 3), WENO5-llf (line 4), WENO5P-llf (line 5), WENO5PH-llf (line 6). Right: ACMTVDfi,  $\kappa = 0.15$  (line 2),  $\kappa = 0.2$  (line 3),  $\kappa = 0.3$  (line 4),  $\kappa = 0.5$  (line 5),  $\kappa = 1$  (line 6). See text for method notation.

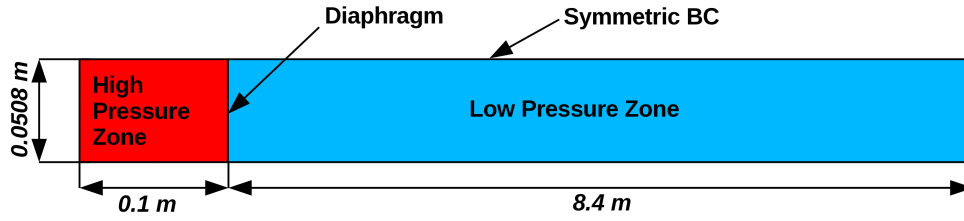


Figure 3: Schematic of a 13 species 2D EAST problem.

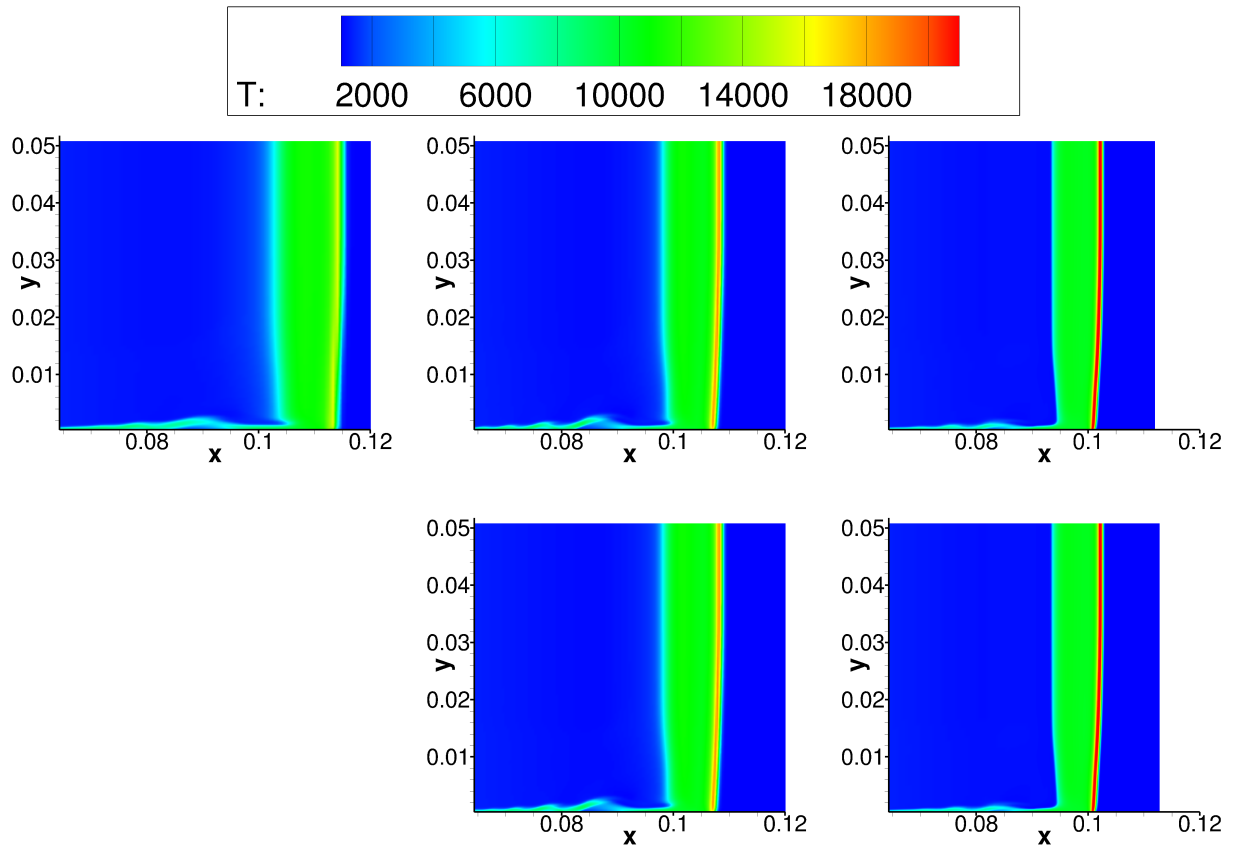


Figure 4: 2D 13 species EAST simulation by TVD for  $CFL = 0.7$  and  $T_{end} = 10^{-5}s$ : Top Row - Three  $x$ -direction grid refinement  $601 \times 121$ ,  $1201 \times 121$  and grid clustering between shear and shock in the  $x$ -direction of  $691 \times 121$ . All  $y$ -grid use boundary grid stretching with a minimum of  $\Delta y = 10^{-5}$ . Bottom Row: Two  $x$ -direction grid refinement  $1201 \times 121$  and grid clustering between shear and shock in the  $x$ -direction of  $691 \times 121$ . All  $y$ -grid use boundary grid stretching with a minimum of  $\Delta y = 5 \times 10^{-6}$ .

<b>grid Nx</b>	601	1201	1201	691	691
<b>refine</b>	no	no	no	yes	yes
<b>min <math>h_y</math></b>	1.e-5	1.e-5	5.e-6	1.e-5	5.e-6
<b>Shock <math>T_{max}, K</math></b>	15,846	18,851	18,848	25,098	25,015
<b>Shear <math>T_{max}, K</math></b>	11,301	11,203	11,203	10,598	10,598

Table 2: Shock and Shear maximum temperature grid dependence at time  $T_{end} = 10^{-5}s$ .  $N_x$  indicates the grid spacing in the  $x$ -direction. The last two columns are for the grid clustering results for two different minimum  $y$ -grid stretching.

### 2.3. 2D 13 Species EAST Test Case

The computational domain is half of the 2D shock tube  $y$ -height with total length  $8.5m$ , height  $0.0508m$  and symmetry boundary condition imposed on the top. The left part of the  $x$ -domain with length  $0.1m$  is a high pressure region. The right part of the domain with length  $8.4m$  is a low pressure region. The gas mixture consists of the same 13 species as the 1D simulation:

$$e^-, He, N, O, N_2, NO, O_2, N_2^+, NO^+, N^+, O_2^+, O^+, He^+.$$

The initial conditions of the high and low pressure regions are listed in the Table 1.

For the left boundary the slip (Euler) wall condition is applied. For the right-side the zero gradient condition is applied for all variables. The bottom boundary is treated as an isothermal wall with the constant temperature  $T_{wall} = 300 K$ . The top boundary is treated as a symmetric boundary condition. Figure 3 shows the schematic of the 2D EAST simulation.

For this 2D test case a very accurate reference solution is not practical to obtain due to the CPU intensive nature of the problem. Here, three levels of refinement are conducted. Figure 3 shows the schematic of the 2D EAST simulation at time  $T_{end} = 10^{-5}s$  using  $CFL = 0.7$  by TVD. Figure 4 shows the computed temperature contour results by TVD for three levels  $x$ - and  $y$ -direction grid refinement simulations. The top row shows three  $x$ -direction grid refinements of  $601 \times 121$ ,  $1201 \times 121$  and grid clustering between shear and shock in the  $x$ -direction of  $691 \times 121$ . All  $y$  grids use boundary grid stretching with a minimum of  $\Delta y = 10^{-5}$ . The bottom row shows the same two  $x$ -direction grid refinements  $1201 \times 121$  and grid clustering between shear and shock in the  $x$ -direction of  $691 \times 121$ . All  $y$  grids use boundary grid stretching with a minimum of  $\Delta y = 5 \times 10^{-6}$ . Comparing the two rows of grid refinement study indicates that by refining the  $x$ -direction grid with the  $y$ -direction the same has a big effect on the locations of the shear/shock. This is due to the fact that aside from the boundary layer, the shear and shock are nearly one dimensional. However, comparing the last two columns of the grid refinement study indicates that by refining the  $y$ -direction grid with the  $x$ -direction the same has no effect on the locations of the shear/shock, but increases the boundary layer prediction. As in the 1D EAST simulation, the discontinuity locations shift as the  $x$ -direction grid is refined. The width of the distance between the shear and the shock shrank as the grid was refined. The shear and shock strength are also different. Table 2 indicates the maximum shear and contact temperature for each set of grids. For the minimum grid stretching of  $\Delta y = 10^{-5}$ , the maximum shear temperature is  $11,301K$ , and maximum shock temperature is  $15,846K$  for the  $601 \times 121$  grid. However, the shear and shock strength are with maximum shear temperature =  $11,203K$ , and maximum shock temperature =  $18,851K$  for the  $1201 \times 121$  grid. For the stretched grid the shear and shock strength are with maximum shear temperature =  $10,598K$ , and maximum shock temperature =  $25,098K$ . As we decrease the minimum grid stretching to  $\Delta y = 5 \times 10^{-6}$ , the shear and shock strength are with maximum shear temperature =  $11,203K$ , and maximum shock temperature =  $18,848K$  for the  $1201 \times 121$  grid. For the stretched grid the shear and shock strength are with maximum shear temperature =  $10,598K$ , and maximum shock temperature =  $25,015K$ . Aside from the different shock/shear locations the result indicated in the last column shows the maximum temperature at the shock location is higher than the result indicated in the the middle and the first columns. Results comparing with WENO5-llf, and with further grid refinement and longer time evolution are reported in [17].

These results indicate that the numerical method and grid dependence of the shear and shock locations are related to the stiffness of the source terms. Note that for non-reacting flows, numerical method and grid dependence of the solution normally do not affect the location of the discontinuities, but rather affect the degree of the smearing of the discontinuities. However for extreme non-reacting flows there are studies reporting the grid dependence of



the shear and shock locations. Studies comparing the previous 1D EAST test case with  $S(U) = 0$  (using the same flow condition) indicate a similar spurious behavior. See Kotov et al. [18] and references cited therein for details. The implication of the EAST computation exercise is to illustrate the danger of practical numerical simulation for problems containing stiff source terms where there is no reliable means of assessing the accuracy of the computed result other than by extreme grid refinement, which is beyond the capability of the current super computer. Before the detailed numerical study for the two stiff detonation test cases, the next section gives a brief description of our recently developed high order shock-capturing method with specific numerical dissipation controls [42, 47, 54].

### 3. Overview of Two Recently Developed High Order Shock-Capturing Schemes

Here only the newly developed high order finite difference method with subcell resolution for advection equations with stiff source terms ([42]) in 2D is briefly summarized. The key aspects of the filter counterpart of the WENO schemes are included at the end of the section. For simplicity of discussion only 2D reactive Euler equations are considered. It is noted that the considered schemes are applicable to 3D reactive flows. The Wang et al. high order scheme with subcell resolution [42] is only developed for a single reaction case and further extension to multi-reaction cases is still under investigation. The regular WENO scheme and the Yee & Sjogreen and Sjogreen & Yee high order nonlinear filter scheme [47, 54, 35, 55, 43] are applicable for any number of species and reactions. The high order nonlinear filter scheme with local flow sensor is applied to further control the amount of numerical dissipation being used for turbulence with strong shocks.

#### 3.1. 2D Reactive Euler Equations

Consider a 2D inviscid combustion flow containing two species

$$(\rho_1)_t + (\rho_1 u)_x + (\rho_1 v)_y = K(T)\rho_2 \quad (13)$$

$$(\rho_2)_t + (\rho_2 u)_x + (\rho_2 v)_y = -K(T)\rho_2 \quad (14)$$

$$(\rho u)_t + (\rho u^2 + p)_x + (\rho uv)_y = 0 \quad (15)$$

$$(\rho v)_t + (\rho uv)_x + (\rho v^2 + p)_y = 0 \quad (16)$$

$$E_t + (u(E + p))_x + (v(E + p))_y = 0 \quad (17)$$

where  $\rho_1$  is the density of burned gas,  $\rho_2$  is the density of unburned gas,  $u$  and  $v$  are the mixture  $x$ - and  $y$ -velocities,  $E$  is the mixture total energy per unit volume,  $p$  is the pressure,  $K(T)$  is the chemical reaction rate and  $T$  is the temperature. The pressure is given by

$$p = (\gamma - 1)(E - \frac{1}{2}\rho(u^2 + v^2) - q_0\rho_2), \quad (18)$$

where the temperature  $T = p/\rho$  and  $q_0$  is the chemical heat released in the reaction.

The mass fraction of the unburnt gas is  $z = \rho_2/\rho$ . The mixture density is  $\rho = \rho_1 + \rho_2$ .

The reaction rate  $K(T)$  is modeled by an Arrhenius law

$$K(T) = K_0 \exp\left(\frac{-T_{ign}}{T}\right), \quad (19)$$

where  $K_0$  is the reaction rate constant and  $T_{ign}$  is the ignition temperature. The reaction rate may be also modeled in the Heaviside form

$$K(T) = \begin{cases} K_0 & T \geq T_{ign} \\ 0 & T < T_{ign}. \end{cases} \quad (20)$$

### 3.2. High Order Finite Difference Methods with Subcell Resolution for Advection Equations with Stiff Source Terms

The general fractional step approach based on Strang-splitting [37] for the 2D reactive Euler equations written in vector notation

$$U_t + F(U)_x + G(U)_y = S(U) \quad (21)$$

is as follows. The numerical solution at time level  $t_{n+1}$  is approximated by

$$U^{n+1} = A\left(\frac{\Delta t}{2}\right)R(\Delta t)A\left(\frac{\Delta t}{2}\right)U^n. \quad (22)$$

The reaction operator  $R$  is over a time step  $\Delta t$  and the convection operator  $A$  is over  $\Delta t/2$ . The two half-step reaction operations over adjacent time steps can be combined to save cost. The convection operator  $A$  is defined to approximate the solution of the homogeneous part of the problem on the time interval, i.e.,

$$U_t + F(U)_x + G(U)_y = 0, \quad t_n \leq t \leq t_{n+1}. \quad (23)$$

The reaction operator  $R$  is defined to approximate the solution on a time step of the reaction problem:

$$\frac{dU}{dt} = S(U), \quad t_n \leq t \leq t_{n+1}. \quad (24)$$

Here, the convection operator consists of, e.g., WENO5 with Roe flux and RK4 for time discretization. If there is no smearing of discontinuities in the convection step, any ODE solver can be used as the reaction operator. However, all the standard shock-capturing schemes will produce a few transition points in the shock when solving the convection equation. These transition points are usually responsible for causing incorrect numerical results in the stiff case. Thus, a direct application of a standard ODE solver at these transition points will create incorrect shock speed. To avoid this, here the Harten's subcell resolution technique [9] in the reaction step is employed. The general idea is as follows. If a point is considered a transition point of the shock, information from its neighboring points which are deemed not transition points will be used instead. In 2D case we apply the subcell resolution procedure dimension by dimension. Here,  $U^T = (\rho_1, \rho_2, \rho u, \rho v, E)$  and we select the mass fraction  $z$  as the stiffness indicator. The algorithm proceeds as follows.

(1) Use a "shock indicator" to identify cells in which discontinuities are believed to be situated. One can use any indicator suitable for the particular problem. Here the minmod-based shock indicator in [9, 33] is considered. Identify "troubled cell"  $I_{ij}$  in both  $x$ - and  $y$ -directions by applying the shock indicator to, e.g., the mass fraction  $z$ . Define the cell  $I_{ij}$  as troubled in the  $x$ -direction if  $|s_{ij}^x| \geq |s_{i-1,j}^x|$  and  $|s_{ij}^x| \geq |s_{i+1,j}^x|$  with at least one strict inequality, where

$$s_{ij}^x = \text{minmod}\{z_{i+1,j} - z_{ij}, z_{ij} - z_{i-1,j}\}. \quad (25)$$

Similarly we can define  $s_{ij}^y$ , the cell  $I_{ij}$  as troubled in the  $y$ -direction.

If  $I_{ij}$  is only troubled in one direction, we apply the subcell resolution along this direction. If  $I_{ij}$  is troubled in both directions, we choose the direction which has a larger jump. Namely, if  $|s_{ij}^x| \geq |s_{ij}^y|$ , subcell resolution is applied along the  $x$ -direction, otherwise it is done along the  $y$ -direction. In the following steps (2)-(3), without loss of generality, we assume the subcell resolution is applied in the  $x$ -direction. Assuming  $I_{ij}$  is troubled in the  $x$ -direction, we apply subcell resolution along the  $x$ -direction.

In a troubled cell identified above, we continue to identify its neighboring cells. For example, we can define  $I_{i+1,j}$  as troubled if  $|s_{i+1,j}^x| \geq |s_{i-1,j}^x|$  and  $|s_{i+1,j}^x| \geq |s_{i+2,j}^x|$  and similarly define  $I_{i-1,j}$  as troubled if  $|s_{i-1,j}^x| \geq |s_{i-2,j}^x|$  and  $|s_{i-1,j}^x| \geq |s_{i+1,j}^x|$ . If the cell  $I_{i-s,j}$  and the cell  $I_{i+r,j}$  ( $s, r > 0$ ) are the first good cells from the left and the right (i.e.,  $I_{i-s+1,j}$  and  $I_{i+r-1,j}$  are still troubled cells), we compute the fifth-order ENO interpolation polynomials  $p_{i-s,j}(x)$  and  $p_{i+r,j}(x)$  for the cells  $I_{i-s,j}$  and  $I_{i+r,j}$ , respectively.

(2) Modify the point values  $z_{ij}$ ,  $T_{ij}$  and  $\rho_{ij}$  in the troubled cell  $I_{ij}$  by the ENO interpolation polynomials

$$\begin{cases} \tilde{z}_{ij} = p_{i-s,j}(x_i; z), & \tilde{T}_{ij} = p_{i-s,j}(x_i; T), & \tilde{\rho}_{ij} = p_{i-s,j}(x_i; \rho), & \text{if } \theta \geq x_i \\ \tilde{z}_{ij} = p_{i+r,j}(x_i; z), & \tilde{T}_{ij} = p_{i+r,j}(x_i; T), & \tilde{\rho}_{ij} = p_{i+r,j}(x_i; \rho), & \text{if } \theta < x_i \end{cases}, \quad (26)$$

where the location  $\theta$  is determined by the conservation of energy  $E$

$$\int_{x_{i-1/2}}^{\theta} p_{i-s,j}(x; E)dx + \int_{\theta}^{x_{i+1/2}} p_{i+r,j}(x; E)dx = E_{ij}\Delta x. \quad (27)$$

Under certain conditions, it can be shown that there is a unique  $\theta$  satisfying Eq. (27), which can be solved using, for example, a Newton's method. If there is no solution for  $\theta$  or there is more than one solution, we choose  $\tilde{z}_{ij} = z_{i+r,j}$ ,  $\tilde{T}_{ij} = T_{i+r,j}$  and  $\tilde{\rho}_{ij} = \rho_{i+r,j}$ . For particular problems one can choose any other suitable method for the reconstruction.

(3) Use  $\tilde{U}_{ij}$  instead of  $U_{ij}$  in the ODE solver if the cell  $I_{ij}$  is a troubled cell. For simplicity, explicit Euler is used as the ODE solver.

$$(\rho z)_{ij}^{n+1} = (\rho z)_{ij}^n + \Delta t S(\tilde{T}_{ij}, \tilde{\rho}_{ij}, \tilde{z}_{ij}). \quad (28)$$

Here we would like to remark that, implicit temporal discretization cannot be used in this step because the troubled values need to be modified explicitly. However, there is no small time step restriction in the explicit method used here, because once the stiff points have been modified, the modified source term  $S(\tilde{T}_{ij}, \tilde{\rho}_{ij}, \tilde{z}_{ij})$  is no longer stiff. Therefore, a regular CFL number is allowed in the explicit method. (Note that if however, a linearized form of a two-level implicit time discretization might be suitable for the reaction step operator. This will be investigated in the future.)

Earlier study reported in [42], in general, a regular  $CFL = 0.1$  using the explicit Euler to solve the reaction operator step can be used in the subcell resolution scheme to produce a stable solution. But the solution is very coarse in the reaction zone because of the underresolved mesh in time. In order to obtain more accurate results in the reaction zone, we evolve one reaction step via  $N_r$  sub steps, i.e.,

$$u^{n+1} = A\left(\frac{\Delta t}{2}\right)R\left(\frac{\Delta t}{N_r}\right)\cdots R\left(\frac{\Delta t}{N_r}\right)A\left(\frac{\Delta t}{2}\right)u^n \quad (29)$$

in some numerical examples studied in [42]. For the present numerical experiment for the 1D detonation problem,  $N_r = 2$ . See [18] for additional  $N_r$  value studies. The study in [18] using  $N_r = 1, 2, 4, 5, 10, 1000$  indicated that  $N_r \geq 2$  is a reasonable choice for the 1D test case. For the 2D detonation problem, a higher  $N_r$  value is desirable.

### 3.3. Well-Balanced High Order Filter Schemes for Reacting Flows ([47, 54, 35, 55, 43])

The high order nonlinear filter scheme of [47, 54, 35], if used in conjunction with a dissipative portion of a well-balanced shock-capturing scheme as the nonlinear numerical flux, is a well-balanced scheme [43]. The well-balanced high order nonlinear filter scheme for reacting flows, solving the fully coupled system 1 consists of three steps.

#### 3.3.1. Preprocessing Step

Before the application of a high order non-dissipative spatial base scheme, the pre-processing step to improve stability had split inviscid flux derivatives of the governing equation(s) in the following three ways, depending on the flow types and the desire for rigorous mathematical analysis or physical argument.

- Entropy splitting of [28] and [49, 46]: The resulting form is non-conservative and the derivation is based on entropy norm stability with boundary closure for the initial value boundary problem.
- The system form of the Ducros et al. splitting [7]: This is a conservative splitting and the derivation is based on physical arguments.
- Tadmor entropy conservation formulation for systems [36]: The derivation is based on mathematical analysis. It is a generalization of Tadmor's entropy formulation to systems and has not been fully tested on complex flows.

See Honein [11] for a comparison of the entropy splitting and other earlier momentum conservation methods.

#### 3.3.2. Base Scheme Step

A full time step is advanced using a high order non-dissipative (or very low dissipation) spatially central scheme on the split form of the governing partial differential equations (PDEs). Summation-by-parts (SBP) boundary operator [27, 34] and matching order conservative high order free stream metric evaluation for curvilinear grids [41] are used. High order temporal discretization such as the third-order or fourth-order Runge-Kutta (RK3 or RK4) temporal is used. It is remarked that other temporal discretizations can be used for the base scheme step. Numerical experiments only focused on RK4 using Roe's approximate Riemann solver.

### 3.3.3. Post-Processing (Nonlinear Filter Step)

After the application of a non-dissipative high order spatial base scheme on the split form of the governing equation(s), to further improve nonlinear stability from the non-dissipative spatial base scheme, the post-processing step of [47, 54, 35] nonlinearly filtered the solution by a dissipative portion of a high order shock-capturing scheme with a local flow sensor. The flow sensor provides locations and amounts of built-in shock-capturing dissipation that can be further reduced or eliminated. The idea of these nonlinear filter schemes for turbulence with shocks is that, instead of solely relying on very high order high-resolution shock-capturing methods for accuracy, the filter schemes [52, 49, 35, 47, 53] take advantage of the effectiveness of the nonlinear dissipation contained in good shock-capturing schemes as stabilizing mechanisms (a post-processing step) at locations where needed. At each grid point, a local flow sensor, e.g., a multi-resolution wavelet would be employed to analyze the regularity of the computed flow data. Only the discontinuity locations would receive the full amount of shock-capturing dissipation. At smooth regions, no shock-capturing dissipation would be added. At turbulent regions, a small fraction of the shock-capturing dissipation would be added to improve stability. The nonlinear dissipative portion of a high-resolution shock-capturing scheme can be any shock-capturing scheme. For reacting flow, it is best to employ the dissipative portion of a well-balanced shock-capturing scheme. By design, the flow sensors, spatial base schemes and nonlinear dissipation models are standalone modules. Unlike standard shock-capturing and/or hybrid shock-capturing methods, the nonlinear filter method requires one Riemann solve per dimension per time step, independent of time discretizations. The nonlinear filter method is more efficient than its shock-capturing method counterparts employing the same order of the respective methods. See [54] for the recent improvements of the work [52, 49, 35, 47] that are suitable for a wide range of flow speed with minimal tuning of scheme parameters. For all the computations shown, if the pre-processing step is used, the Ducros et al. splitting is employed. This is due to the fact that for the subject test cases we need a robust conservative splitting as the preprocessing step. Some attributes of the high order filter approach are:

- Spatial Base Scheme: High order and conservative (no flux limiter or Riemann solver)
- Physical Viscosity: Contribution of physical viscosity, if it exists, is automatically taken into consideration by the base scheme in order to minimize the amount of numerical dissipation to be used by the filter step
- Efficiency: One Riemann solve per dimension per time step, independent of time discretizations (less CPU time and fewer grid points than their standard shock-capturing scheme counterparts)
- Accuracy: Containment of numerical dissipation via a local wavelet flow sensor
- Well-balanced scheme: These nonlinear filter schemes are well-balanced schemes for certain chemical reacting flows [43]
- Stiff Combustion with Discontinuities: For some stiff reacting flow test cases, it appears that the high order filter scheme is able to obtain the correct propagation speed of discontinuities on coarse, underresolved meshes for which the standard high order shock-capturing (e.g., WENO) schemes cannot (see the result below)
- Parallel Algorithm: Suitable for most current supercomputer architectures

The nonlinear filter counterpart of the subcell resolution method employing, e.g., WENO5 or WENO7 as the dissipative portion of the filter numerical flux (WENO5fi or WENO7fi) can be obtained in a similar manner with the convection operator replaced by the nonlinear filter scheme and will be denoted by WENO5fi/SR or WENO7fi/SR.

## 4. Numerical Results

Here “coarse grids” means standard mesh density requirement for accurate simulation of typical non-reacting flows of similar problem setup. The two well known stiff detonation test cases consist of the Arrhenius 1D Chapman-Jouguet (C-J) detonation wave [10, 40] and a 2D Heaviside detonation wave [1]. These are the same two test cases considered in [42]. The considered six schemes are WENO5, “WENO5/SR” [42], “WENO5fi” (the Yee & Sjögren nonlinear filter version of WENO5 using a local flow sensor to further limit the amount of WENO5 numerical dissipation), “WENO5fi+split” (the Ducros et al. splitting of the governing equations [7] of WENO5fi [47, 54]), and

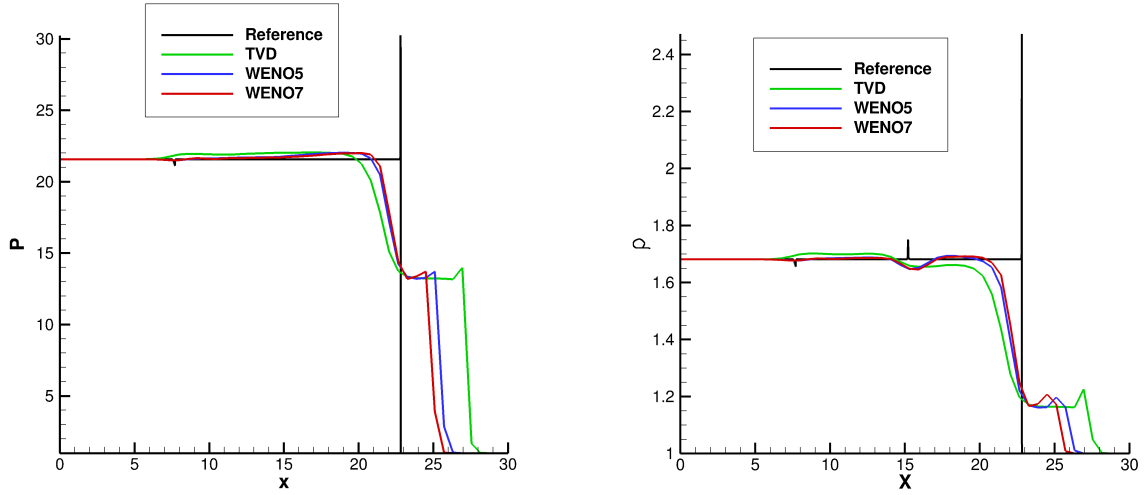


Figure 5: 1D C-J detonation problem, Arrhenius case for the original stiffness  $K_0$  at  $t = 1.8$ : Pressure and density comparison among three standard shock-capturing methods (TVD, WENO5, WENO7) using 50 uniform grid points with  $CFL = 0.05$ .

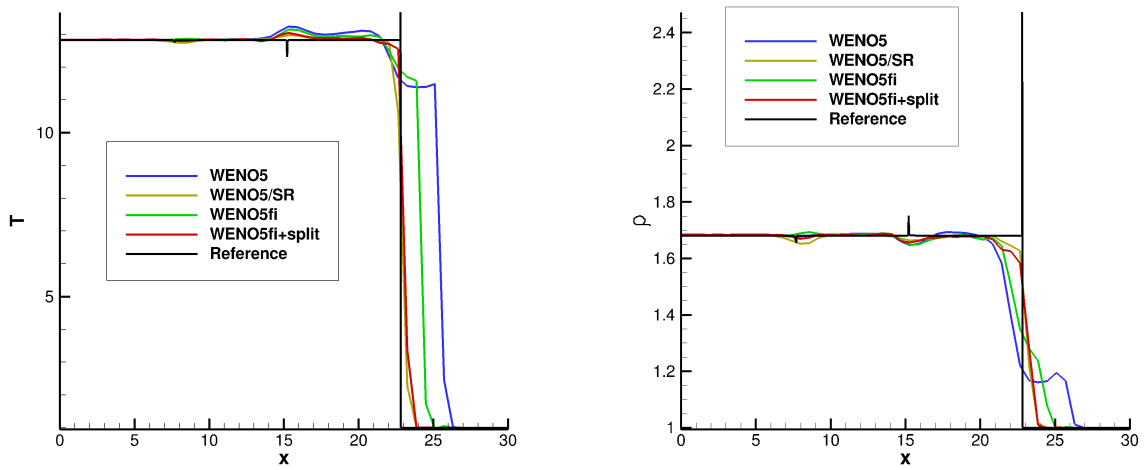


Figure 6: 1D C-J detonation problem, Arrhenius case for the original stiffness  $K_0$  at  $t = 1.8$ : Temperature and density comparison among standard high order shock-capturing methods and low dissipative methods (WENO5, WENO5/SR, WENO5fi+split and WENO5fi/SR+split) using 50 uniform grid points with  $CFL = 0.05$ .

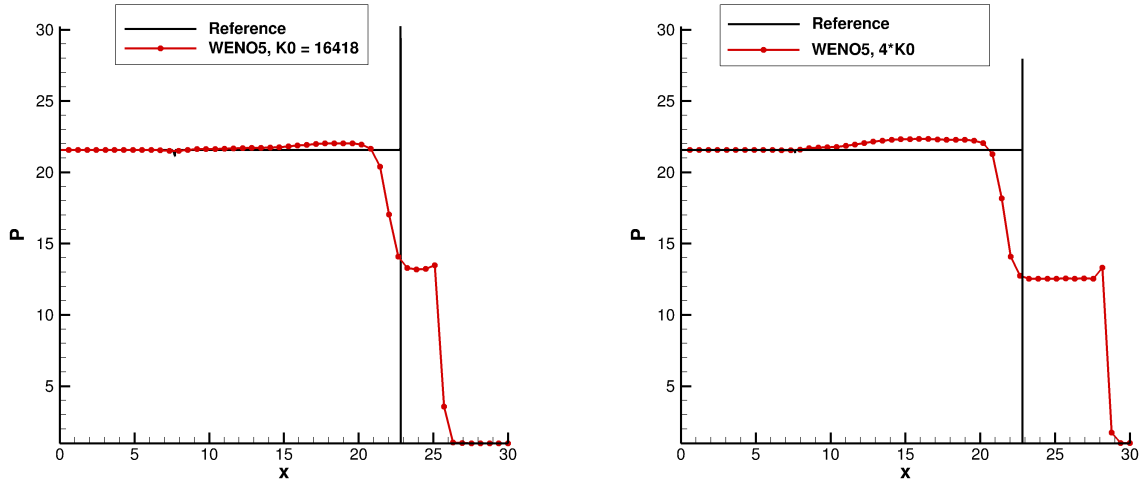


Figure 7: 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Pressure comparison between the original stiffness  $K_0$  and  $4K_0$  of the source term computed by WENO5 using 50 uniform grid points. All the CFL values for the inviscid simulations are based on the convection part of the PDEs.

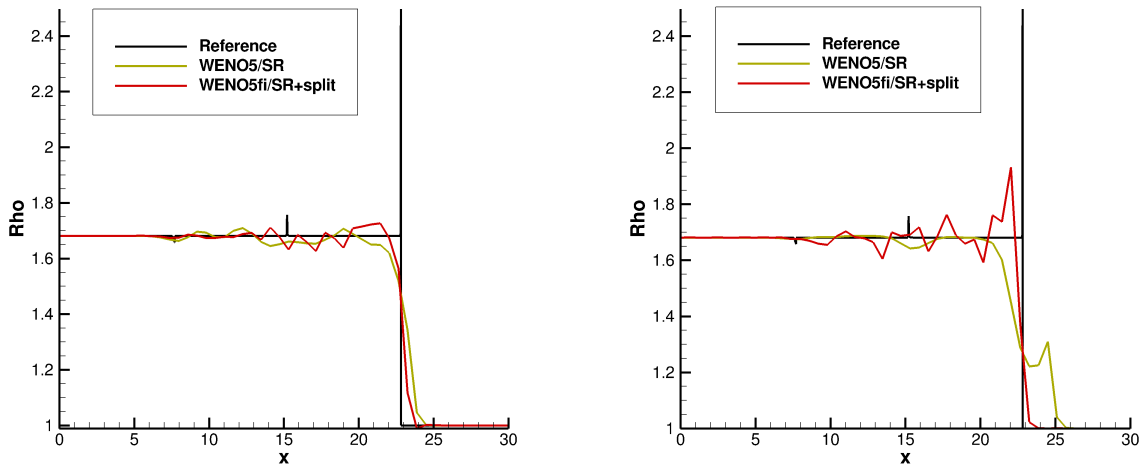


Figure 8: C-J detonation problem, Arrhenius case at  $t = 1.8$ : Density comparison between WENO5/SR and WENO5fi/SR+split for  $100K_0$  and  $1000K_0$  using 50 uniform grid points with  $CFL = 0.05$ .

”WENO5fi/SR+split ” (the nonlinear filter version of WENO5/SR with Ducros et al. splitting of the governing equations). All of the considered methods use the Roe’s average states. For the temporal discretization the classical fourth-order Runge-Kutta method (RK4) is used since the TVD RK3 has lower CFL limit than RK4 but with a similar behavior as RK4. The results by RK3 are not considered here. Note that all the CFL values for the inviscid simulations are based on the convection part of the PDEs. In addition, the computed solutions and their spurious behavior by the studied schemes presented here could be slightly different from the results presented in [42] due to the minor differences in the formulation of the governing equation; e.g., different choice of variables.

**Remark: The following study (also the EAST simulations shown earlier) focuses only on solving the reactive system using the Strang splitting.** In addition, the high order new subcell resolution method utilizes the Strang splitting procedure and it is natural to compare among methods using the same procedure in solving the governing equations. Moreover, for the following 1D and 2D detonation test cases, all of the results include a cut off safeguard if densities are outside the permissible range. Spurious behavior of the same schemes by solving the fully coupled reactive system without the Strang splitting is reported in our companion study [18]. Some comparisons between the two approaches are briefly summarized here. The main reason for the present focus study on the Strang splitting is due to the fact that it is widely used in combustion and reactive flow simulations. The simple cut off safeguard procedure is also commonly used by practitioners in computational physics and engineering simulations.

#### 4.1. 1D Chapman-Jouguet (C-J) Detonation Wave (Arrhenius Case)

The test case is the 1D C-J detonation wave (Arrhenius case) [10, 40]. The initial values consist of totally burnt gas on the left-hand side and totally unburnt gas on the right-hand side. The density, velocity, and pressure of the unburnt gas are given by  $\rho_u = 1$ ,  $u_u = 0$  and  $p_u = 1$ .

The initial state of the burnt gas is calculated from C-J condition:

$$p_b = -b + (b^2 - c)^{1/2}, \quad (30)$$

$$\rho_b = \frac{\rho_u[p_b(\gamma + 1) - p_u]}{\gamma p_b}, \quad (31)$$

$$S_{CJ} = [\rho_u u_u + (\gamma p_b \rho_b)^{1/2}]/\rho_u, \quad (32)$$

$$u_b = S_{CJ} - (\gamma p_b / \rho_b)^{1/2}, \quad (33)$$

where

$$b = -p_u - \rho_u q_0 (\gamma - 1), \quad (34)$$

$$c = p_u^2 + 2(\gamma - 1)p_u \rho_u q_0 / (\gamma + 1). \quad (35)$$

The heat release  $q_0 = 25$  and the ratio of specific heats is set to  $\gamma = 1.4$ . The ignition temperature  $T_{ign} = 25$  and  $K_0 = 16,418$ . The computation domain is  $[0, 30]$ . Initially, the discontinuity is located at  $x = 10$ . At time  $t = 1.8$ , the detonation wave has moved to  $x = 22.8$ . The reference solution is computed by the regular WENO5 scheme with 10,000 uniform grid points and  $CFL=0.05$ .

##### 4.1.1. Initial Study of Scheme Behavior [45]

Figure 5 shows the pressure and density comparison among the standard TVD, WENO5 and WENO7 using 50 uniform grid points and  $CFL = 0.05$  for the same stiffness  $K_0 = 16,418$  used in [45]. Figure 6 shows the pressure and density comparison among the standard WENO5 scheme, WENO5/SR, WENO5fi and WENO5fi+split using 50 uniform grid points. For this particular problem and grid size, all standard TVD WENO5 and WENO7 exhibit wrong shock speed of propagation with the lower order and more dissipative schemes exhibiting the largest error. WENO5fi+split compares well with WENO5/SR for the computed pressure solution. WENO5/SR and WENO5fi+split can capture the correct structure using fewer grid points than those in [10] and [40]. A careful examination of the 50 coarse grid mass fraction solutions indicates that WENO5fi+split is 0.7 grid point ahead of WENO5/SR at the discontinuity location when compared to the reference solution. Since WENO5fi+split is less dissipative than WENO5, the restriction of the shock-capturing dissipation using the wavelet flow sensor helps to improve the wrong propagation speed of discontinuities without the subcell resolution procedure. It is interesting to see that all of the methods

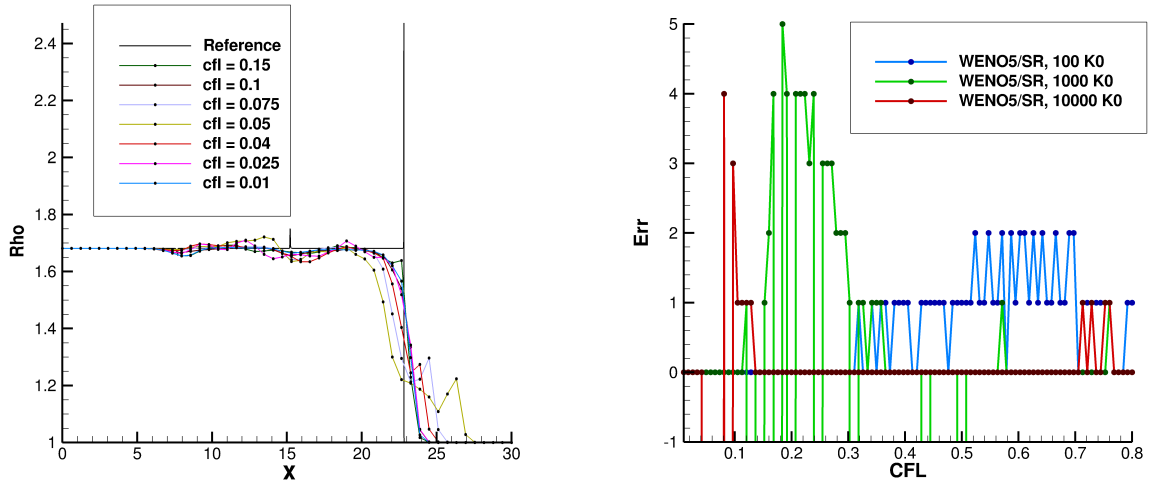


Figure 9: 1D C-J detonation problem, Arrhenius case using 50 uniform grid points: Density comparison for seven CFL numbers by WENO5/SR (left). Number of grid point away from the reference solution (Err) as a function of the CFL number (128 CFL values with  $6.316455696 \times 10^{-3}$  equal increment) for three stiffness coefficients ( $100K_0$ ,  $1000K_0$ ,  $10000K_0$ ) by WENO5/SR. A negative "Err" value indicates the number of grid points behind the reference shock solution. For certain values of CFL, divergent solutions might occur that are outside the plotting area. See e.g., the red and green negative values of Err. All the CFL values for the inviscid simulations are based on the convection part of the PDEs.

(except WENO5) produce oscillatory solutions in the vicinity of the reaction front (figures not shown). This behavior prompted us to perform a systematic six levels of uniform grid refinements (200, 400, 800, 1600, 3200 and 6400, figures not shown). As the number of grid points increases, this oscillatory behavior in the vicinity of the reaction front becomes more pronounced. However, for the more dissipative scheme WENO5, as we refine the grid, the computed solution gets closer and closer to the reference solution. The spurious oscillation might be contributed by the use of the Roe's average state without any correction for reacting flows. See Jenny et al. [14] and related development for details.

#### 4.1.2. Scheme Behavior with Increase in Stiffness of the Source Terms

Figure 7 indicates the behavior of WENO5 for two stiffness coefficients of the reaction rate using 50 grid points and  $CFL = 0.05$ . As the stiffness of the source term increases, the wrong shock location gets further and further away from the reference solution. It seems that the reference solution is independent of the stiffness coefficient. The subcell resolution technique developed in [42] is only designed for the spatial discretization and is frozen during the time step evolution. As the stiffness increases, the CFL number in WENO5/SR needs to decrease in order to obtain sufficient temporal resolution to capture the correct shock location. Figure 8 indicates that as we increase the stiffness coefficient further, WENO5/SR needs a smaller  $CFL=0.01$  in order to obtain the correct shock speed for the  $1000K_0$  stiffness, whereas WENO5fi/SR+split was able to maintain the correct shock speed for this grid with  $CFL=0.05$ . For this problem it is indicated in Eq. (4.15) in Bao & Jin [1], that the shock speed depends on the initial condition and  $\gamma$  has a closed form solution. It appears that the shock location is independent of the stiffness coefficient for this problem. We use that formula to judge if the reference solution is close to the true shock location. For the original  $K_0$  case the distance between the reference (10,000 grid) and the exact solution is 5 points which is 0.025 point on the 50 grid point spacing. Due to the high cost of obtaining a closer to the exact reference solution, we consider the current reference solution as the reference shock location. For the stiffer cases we also use the reference solution for the  $K_0$  (although the spike at the detonation front is not the same, the shock location should be within one grid point of the coarse grid solution). It would be too costly to obtain a better detonation front spike value for the stiffer case as all of the coarse grid solutions are far removed from resolving the detonation front.



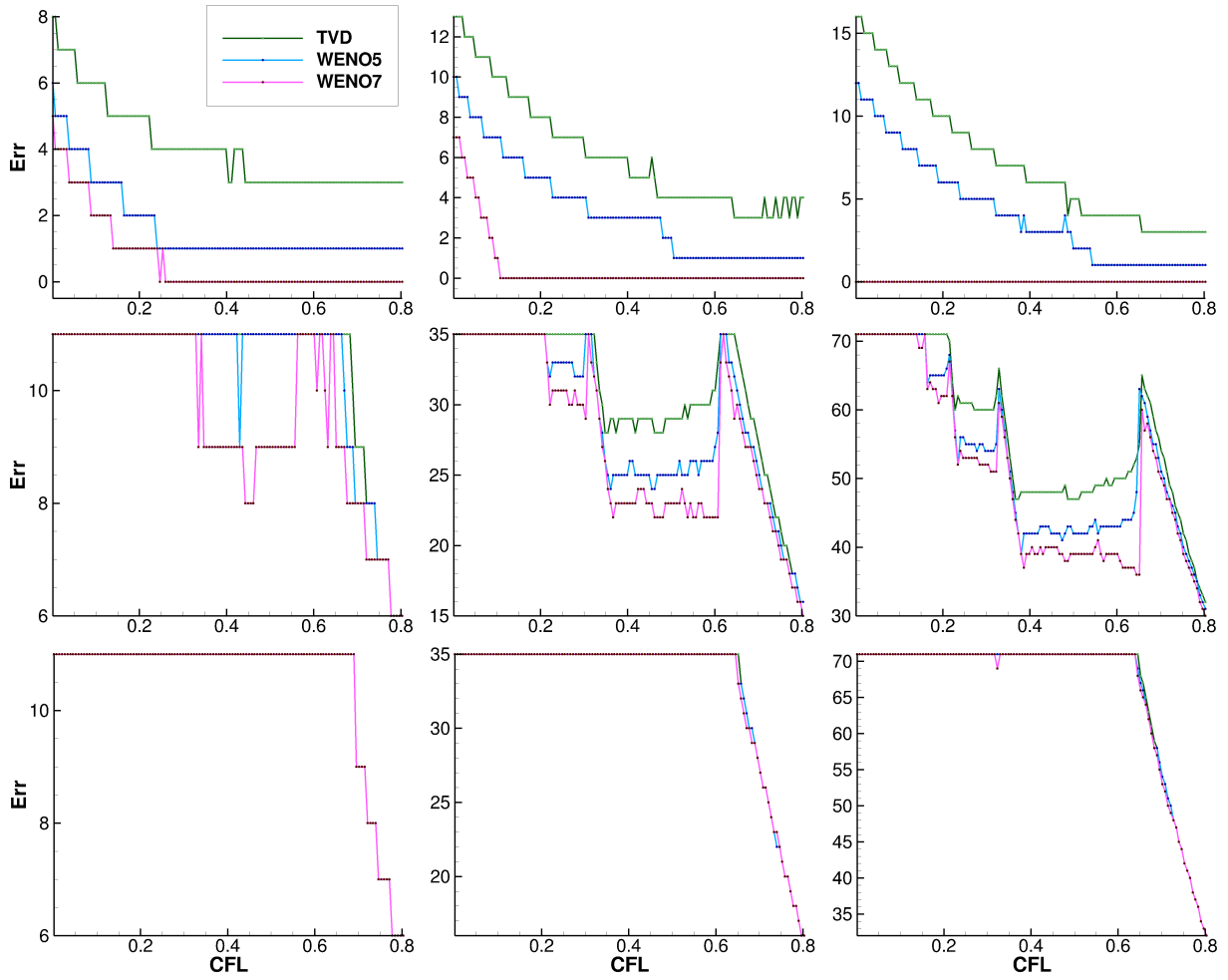


Figure 10: 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Number of grid points away from the reference shock solution (Err) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increment) for three standard shock-capturing methods using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ ,  $100K_0$ ,  $1000K_0$  (top to bottom). See Fig. 9 for additional captions.

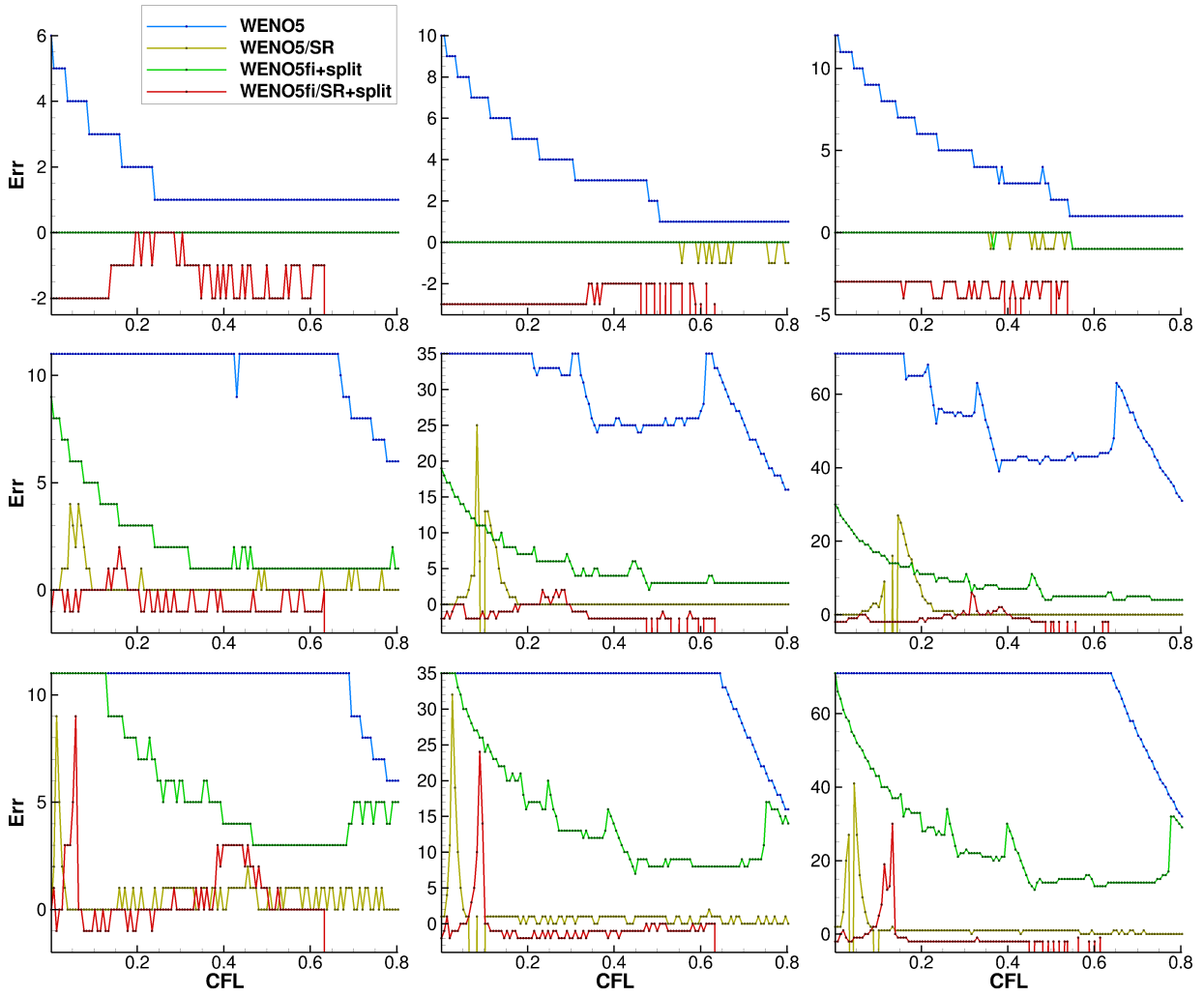


Figure 11: 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Number of grid point away from the reference shock solution (Err) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increment) for three low dissipative shock-capturing methods using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ ,  $100K_0$ ,  $1000K_0$  (top to bottom). See Fig. 9 for additional captions

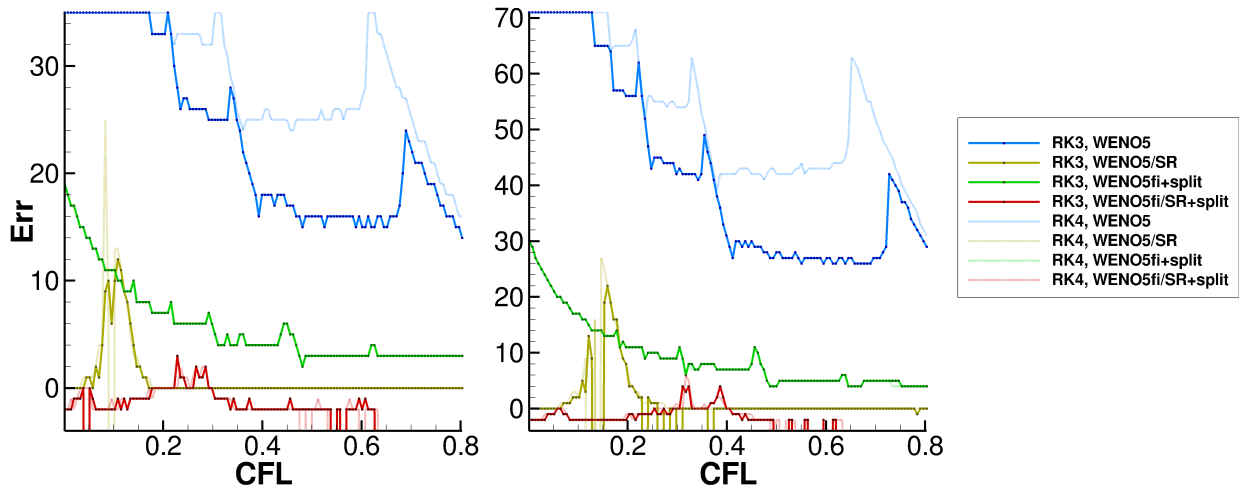


Figure 12: 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Comparison of the same spatial discretization with RK4 and RK3 temporal discretization for three low dissipative shock-capturing methods using 150, 300 uniform grid points and for stiffness  $K_0$ .

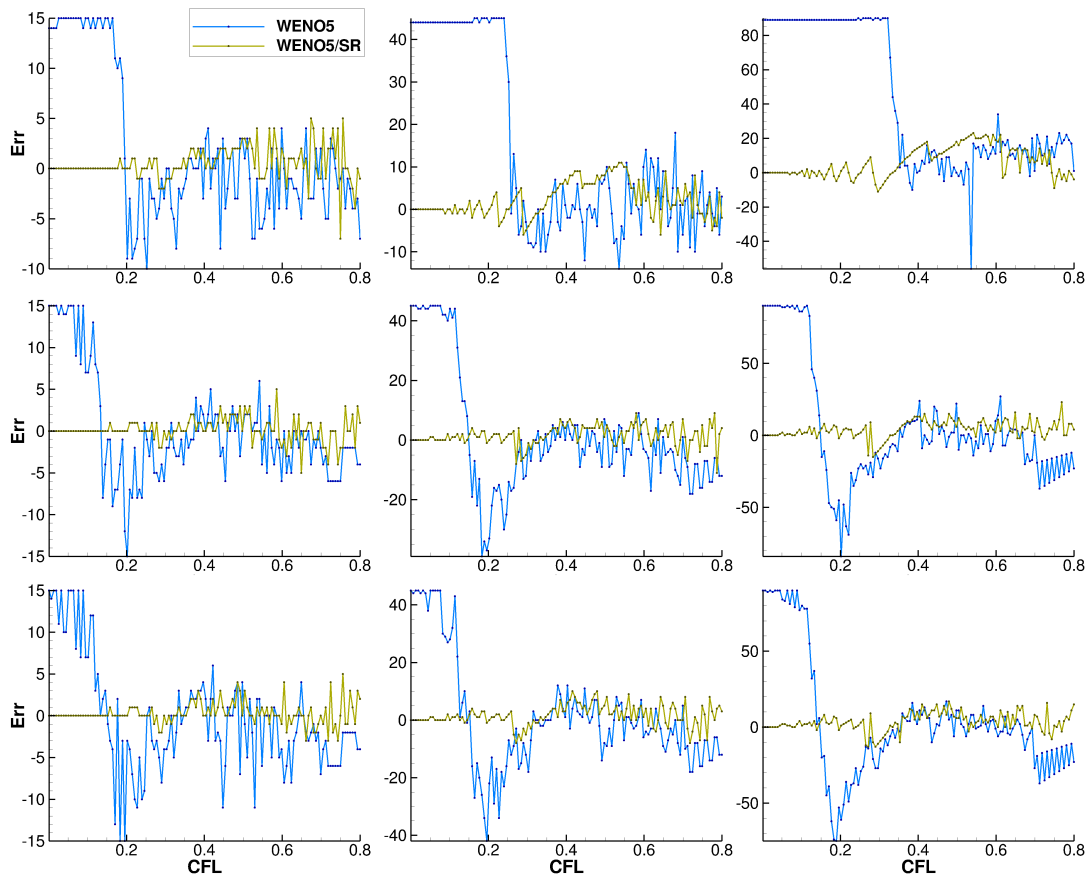


Figure 13: LeVeque and Yee linear advection and nonlinear stiff source term test case [24]: Number of grid points away from the reference shock solution (Err) as a function of the CFL number (128 discrete CFL values between (0.001, 8) with  $6.291338583 \times 10^{-3}$  equal increment) by WENO5 and WENO5/SR using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0, 100K_0, 1000K_0$  (top to bottom). See Fig. 9 for additional captions.

#### 4.1.3. Scheme Behavior as a Function of CFL, Grid Refinement and Stiffness of the Source Terms

The result from Figure 7 prompted us to perform a more systematic study on the spurious numerics for the test case. Figure 9 shows the effect of the time steps for seven CFL values that are under the CFL limit (left sub-figure), using 50 grid points and WENO5/SR. The right sub-figure shows the error in terms of the number of grid points away from the reference shock location (Err) for three stiffness coefficients  $K_0$ ,  $100K_0$  and  $1000K_0$  as the function of 128 discrete CFL values. The 128 discrete CFL values are ( $0.0001 \leq CFL \leq 0.803$ ) with  $6.316455696 \times 10^{-3}$  equal increment. Here, Err is round down to the nearest integer number. Note that the CFL limit for WENO5/SR and its filter counterparts are lower than 0.8 due to the explicit Euler reaction step. A negative "Err" value indicates the number of grid points behind the reference shock solution. For certain values of CFL, divergent solutions might occur that are outside the plotting area. See e.g., the red and green negative values of Err. All the CFL values for the inviscid simulations are based on the convection part of the PDEs. As the stiffness coefficient increases, it is more and more difficult to obtain the correct shock locations by WENO5/SR.

Figures 10 and 11 illustrate the error for 128 discrete CFL values for the three standard shock-capturing schemes (TVD, WENO5 and WENO7) and the three improved high order shock-capturing schemes (WENO5/SR, WENO5fi+split and WENO5fi/SR+split). The study is for three uniform grids 50, 150 and 300 (left to right columns in the plot) and the three stiffness coefficients  $K_0$ ,  $100K_0$  and  $1000K_0$  (top to bottom in the plot). Results indicated that even for  $CFL = 0.001$  using the original  $K_0$  stiffness, TVD, and WENO5, are not able to obtain the correct shock location using the three considered grid points and the three stiffness coefficients as indicated on the Err plot (Fig. 10). For WENO7 for the three grids with the original stiffness  $k_0$ , the correct shock speed can be obtained for most of the CFL values. As a matter of fact, for larger CFL, it performed better than WENO5/SR and its filter counterpart. In additions WENO5 produces less "Err" for larger CFL. This again indicates that the more accurate scheme results in a better chance of avoiding the wrong shock speed spurious numerics. As the stiffness increases, WENO7 no longer produces the correct shock speed by the considered three grids. On the contrary, for certain CFL values the improved high order shock-capturing methods for reacting flows, e.g., WENO5/SR, WENO5fi+split and WENO5fi/SR+split, are able to obtain the correct shock speed. These time steps (CFL values) that can avoid spurious numerics do not have to be very small, but they consist of disjoint segments for the time steps that are within the CFL limit. It appears that the special dissipation control exhibits more spurious behavior than WENO7 for the original  $K_0$  case. In addition, WENO5fi/SR+split performs better for the stiffer cases  $100K_0$  and  $1000K_0$  than the original  $K_0$ , whereas WENO5fi+split performs better than WENO5 for larger CFL.

The current study indicated that using the standard CFL condition for the homogeneous part of the PDEs (non-reacting part of the governing equations) does not guarantee a correct solution or the correct speed of propagation of discontinuities. A stiff ODEs solver with variable time step control in solving the reaction part of the operator using the fractional step approach allows the stiffness of the source terms to come into play. However, as indicated in [56, 57, 51, 38], spurious numerics due to the spatial discretization are more difficult to avoid because of the nonlinearity of the source terms. The search for further improvement of the aforementioned scheme continues. See further discussion on possible improvement on the source term treatment numerical strategy in the subsection after next.

#### 4.1.4. Scheme Behavior by RK3 and By a Single Scalar PDE Case

All of the results shown are by RK4 temporal discretization. Figure 12 shows that the RK4 and RK3 exhibit a similar trend but with slight variation in solution behavior for the 1D detonation problem. As an illustration, the above behavior of the studied schemes also occurs for a simple scalar case Eqn. 2 and Eqn. 3 studied by LeVeque and Yee [24] in 1990 for second-order schemes. Figure 13 shows a general trend of the scheme behavior by WENO5/SR. However, WENO5 behaves differently from the system test case. In this case all the nonlinearity and stiffness contained in the governing equation are due to the source term as the convection term in the LeVeque & Yee's scalar model PDE is linear. It appears that the nonlinearity due to the convection terms does not alter the general spurious behavior pattern.

#### 4.1.5. Effect of the $N_r$ parameter in Strang Splitting of the Reactive Equations [18]

All of the previous computations use  $N_r = 2$ . The effect of the  $N_r$  value on the accuracy of the solution is investigated in [18]. Figure 14, taken from [18], summarizes the comparison among the different values of  $N_r =$

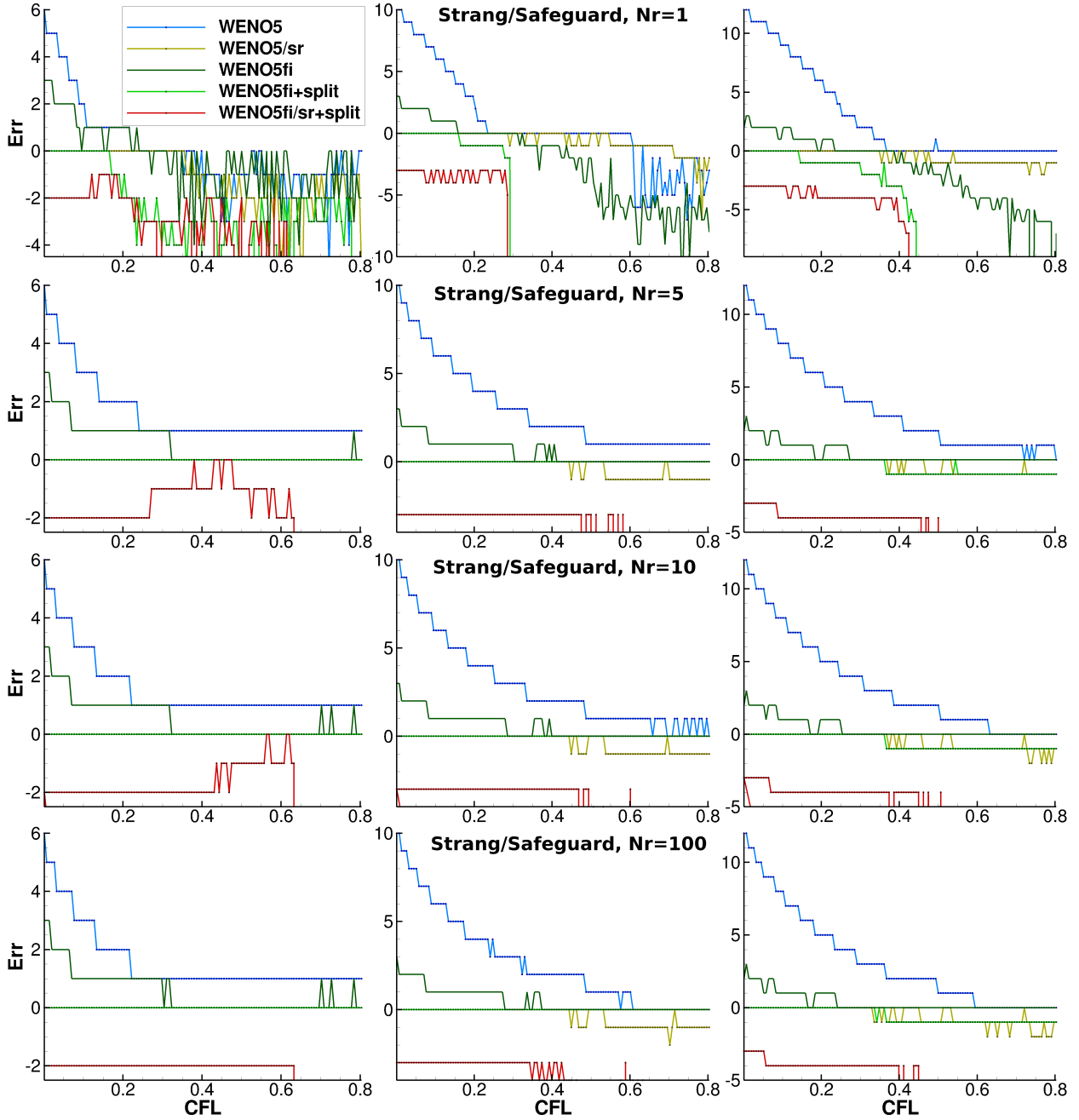


Figure 14:  $N_r = 1, 5, 10, 100$  study using Strang splitting by improved schemes for the 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ . Number of grid points away from the reference shock solution ( $Err$ ) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increments) using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ : WENO5, WENO5/SR, WENO5fi, WENO5fi+split and WENO5fi/SR+split. All of the computations use RK4.

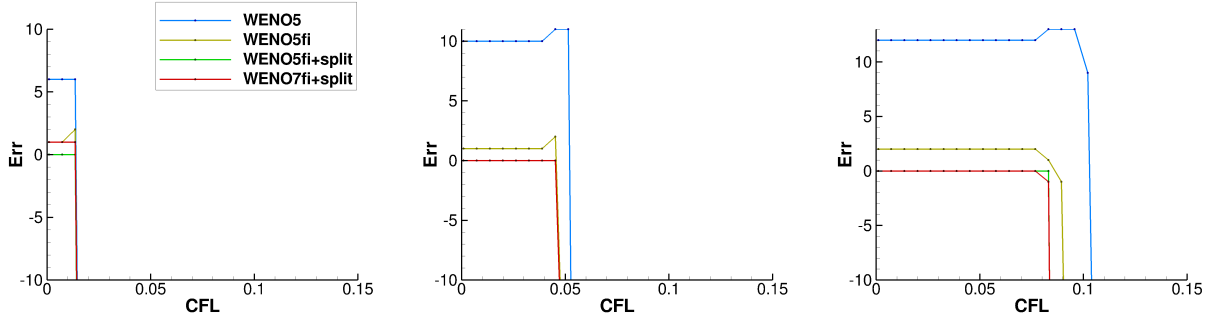


Figure 15: No Strang splitting results for the 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Number of grid point away from the reference shock solution (Err) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increment) for three low dissipative shock-capturing methods using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ .

1, 5, 10, 100 for case (a) (Strang/Safeguard) using  $K_0$  and 50, 150 and 300 grid points. The results indicate that a sufficient number of sub-reaction steps improves the overall accuracy and yields a reduction in spurious numerics. Further increase of  $N_r$  does not show a significant improvement. However, the results of the same study for case (b) (Strang/No-safeguard) indicate that a less spurious behavior for large  $N_r$  values (Figures not shown). In this case WENO5/SR and WENO5fi+split are able to obtain the correct solution using an increased CFL number as  $N_r$  is increases.

One method to further improve the spurious behavior is to use variable time step control. Preliminary studies indicate a significant reduction of spurious behavior in some cases when checking the positivity after each RK stage and refining the timestep by a factor of 2 in case of failing the positivity criteria.

#### 4.1.6. Solving Fully Coupled Reactive Equations vs. Strang Splitting of the Reactive Equations

All of the above computations (also the 2D case to be shown) are by solving the reactive equations via the Strang splitting procedure. Comparison of the solution behavior between solving the fully coupled reactive equations (without the Strang Splitting procedure) and the Strang splitting procedure is reported in [18]. Studies show that solving the fully coupled reactive equations is very unstable for standard shock-capturing schemes as well as for their high order filter counterparts. Using a very small CFL for  $K_0$ , and the same three grids and CFL range, a similar wrong propagation speed of discontinuities is observed by standard shock-capturing schemes for all considered CFL (with the exception one grid point error for WENO7fi+split using a 50 grid). However, WENO5fi+split and WENO7fi+split are able to obtain the correct shock speed using the same small CFL. For stiffness coefficients  $100K_0$  and  $1000K_0$  using the same three grids, no stable solutions are obtained except in the case of  $100K_0$  and 300 grid points using  $\text{CFL} = 6.316455696 \times 10^{-3}$  (a wrong speed solution is obtained). See Fig. 15 for the  $K_0$  result. To further examine the difference between the two procedures in solving the reactive equations, we compare the fully coupled solution procedure with the Strang splitting procedure using a 10,000 grid. For fine enough grid points, both procedures produce the same result.

#### 4.1.7. Effect of Employing a Cut Off Safeguard Procedure

All of the results presented employ a cut off safeguard procedure if densities are outside the permissible range. Figure 16 shows the same computation without the cut off safeguard procedure using the Strang splitting. The procedure is also very unstable. In addition, if the Zhang & Shu positivity-preserving WENO5 [58] using the Lax-Freidrichs flux without a cut off safeguard procedure is employed, again a similar behavior as their standard WENO5 counterparts. See [18] for some discussion. (One possible scheme improvement is to use the positivity preserving version of the studied schemes [12]. This will be a subject of the future investigation.) For  $K_0$ , and the same three grids and CFL range, a similar wrong propagation speed of discontinuities is observed by WENO5 for small CFL. However, WENO5/SR and WENO5fi+split are able to obtain the correct shock speed using the same small CFL. WENO5fi/SR+split is not able to obtain the correct shock speed for even the smallest considered CFL value ( $\text{CFL} = 6.316455696 \times 10^{-3}$ ). One of the possible causes might be due to the incompatibility of the combined

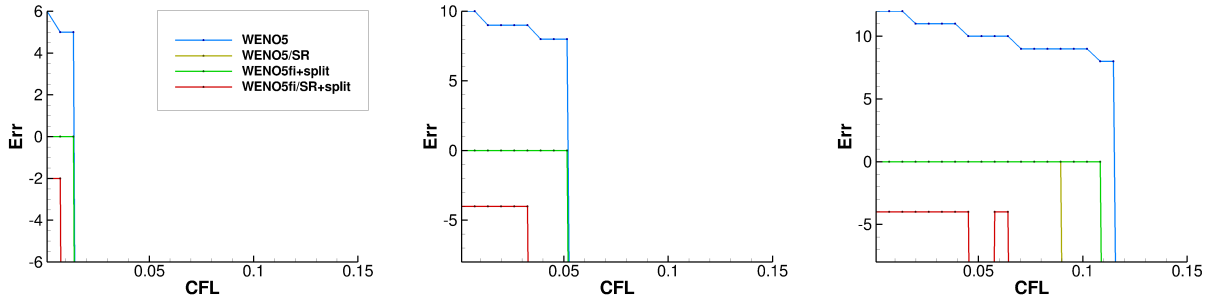


Figure 16: No cutoff safeguard procedure and Strang splitting results for the 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Number of grid point away from the reference shock solution (Err) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increment) for three low dissipative shock-capturing methods using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ .

Strang splitting using  $N_r = 2$ , and the nonlinear filter procedure. For stiffness coefficients  $100K_0$  and  $1000K_0$  using the same three grids, no stable solutions are obtained except in the case of  $100K_0$  and 300 grid points using  $\text{CFL} = 6.316455696 \times 10^{-3}$  (a wrong speed solution is obtained). See Fig. 16 for the  $K_0$  result. The solution behavior of solving the fully coupled reactive equations is similar to using the Strang splitting without the cut off safeguard procedure. Studies in [18] also indicate that there is no visible difference in solution behavior in using the cut off safeguard procedure or not when solving the fully coupled reactive equations.

#### 4.1.8. Positivity-Preserving High Order Methods [19]

The newly developed positivity-preserving flux limiters for general high-order schemes of [12] keep the original scheme unchanged and detect critical numerical fluxes that may lead to negative density and pressure, and then imposes a cut-off flux limitation to satisfy a positivity-preserving condition. The [12] method appears to be a better strategy than the simple safeguard procedure considered above.

Comparative study among [12] and [58] positivity-preserving schemes and the standard WENO counterparts in conjunction with the Strang splitting without the safeguard procedure has been performed. The results indicate that in case of using positivity-preserving methods it is possible to perform a computation with a slightly larger CFL than in the case of using the standard WENO counterparts (figure not shown). Figure 17 indicates the Hu et al. positivity-preserving scheme [12] (top row) exhibits a similar behavior as the Zhang & Shu positivity scheme [58] (bottom row) using the same  $N_r = 10$  and WENO5/LLF as the base scheme. Here “LLF” stands for the local Lax-Friedrichs flux. For more information, see our companion paper [19].

#### 4.1.9. Are Pointwise Evaluation of the Source Term and Roe’s Average State Appropriate?

On all of the above numerical computations, the pointwise evaluation of the source term was used. However, the studies by Lafon & Yee [20, 21] and Griffiths et al. [8] indicated that pointwise evaluation of the source term appears to be the least stable. One approach suggested in Lafon & Yee and Griffiths et al. is to use non-pointwise evaluation of the source term that is more compatible with the convection difference operator. The non-pointwise evaluation of the source term might improve numerical stability and minimize the wrong speed of propagation. In addition, there are studies in the literature showing that using the standard Roe’s average state for reacting/multi-phase flows can create spurious oscillations near the discontinuities. See for example Jenny, Müller & Thomann [14] and related later articles. Further investigation along these directions is planned. The current investigation is to confirm part of the spurious behavior in the studies by Lafon & Yee and Griffiths et al. for system cases.

## 4.2. 2D Detonation Waves

This example is taken from ([1]). The chemical reaction is modeled by the Heaviside form with the parameters

$$\gamma = 1.4, \quad q_0 = 0.5196 \times 10^{10}, \quad K_0 = 0.5825 \times 10^{10}, \quad T_{ign} = 0.1155 \times 10^{10}$$

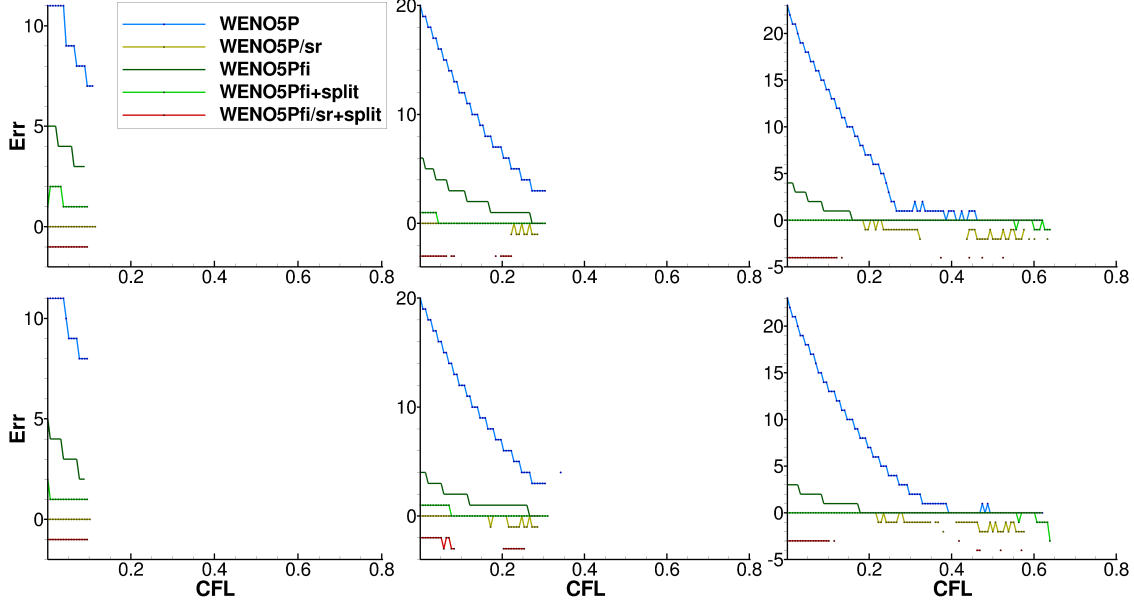


Figure 17: Strang splitting no safeguard schemes based on Hu et al. positivity-preserving method (top) and Zhang & Shu positivity-preserving method (bottom) for the 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ . Number of grid points away from the reference shock solution ( $Err$ ) as a function of the CFL number (128 discrete CFL values with  $6.316455696 \times 10^{-3}$  equal increments) using 50, 150, 300 uniform grid points (across) and for stiffness  $K_0$ : WENO5, WENO5/SR, WENO5fi, WENO5fi+split and WENO5fi/SR+split. All of the computations use RK3, Strang splitting with  $N_r = 10$ .

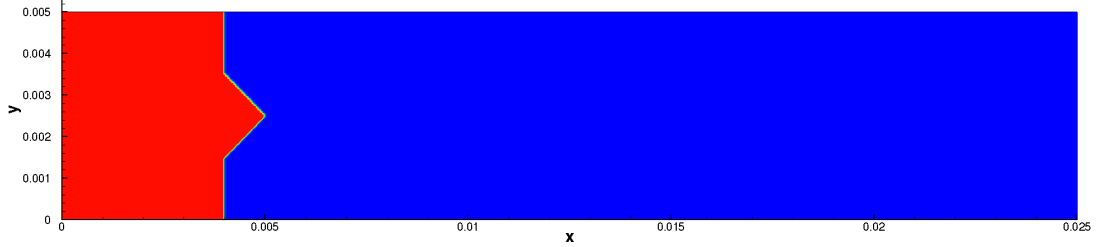


Figure 18: Schematic of the 2D detonation test case initial data.

in CGS units. Consider a two-dimensional channel of width 0.005 with solid walls at the upper and lower boundaries. The computational domain is  $[0, 0.025] \times [0, 0.005]$ . The initial conditions are

$$(\rho, u, v, p, z) = \begin{cases} (\rho_b, u_b, 0, p_b, 0), & \text{if } x \leq \xi(y), \\ (\rho_u, u_u, 0, p_u, 1), & \text{if } x > \xi(y), \end{cases} \quad (36)$$

where

$$\xi(y) = \begin{cases} 0.004 & |y - 0.0025| \geq 0.001, \\ 0.005 - |y - 0.0025| & |y - 0.0025| < 0.001, \end{cases} \quad (37)$$

and  $u_u = 0$ ,  $\rho_u = 1.201 \times 10^{-3}$ ,  $p_u = 8.321 \times 10^5$  and  $u_b = 8.162 \times 10^4$ . Values of  $p_b$  and  $\rho_b$  are defined by Eq. (30) and (31). In this case  $u_b$  is greater than defined by Eq. (33). Figure 18 shows the schematic of the 2D detonation problem.

#### 4.2.1. Initial Study of Scheme Behavior

One important feature of this solution is the appearance of triple points, which travel in the transverse direction and reflect from the upper and lower walls. A discussion of the mechanisms driving this solution is given in [16]. Again,



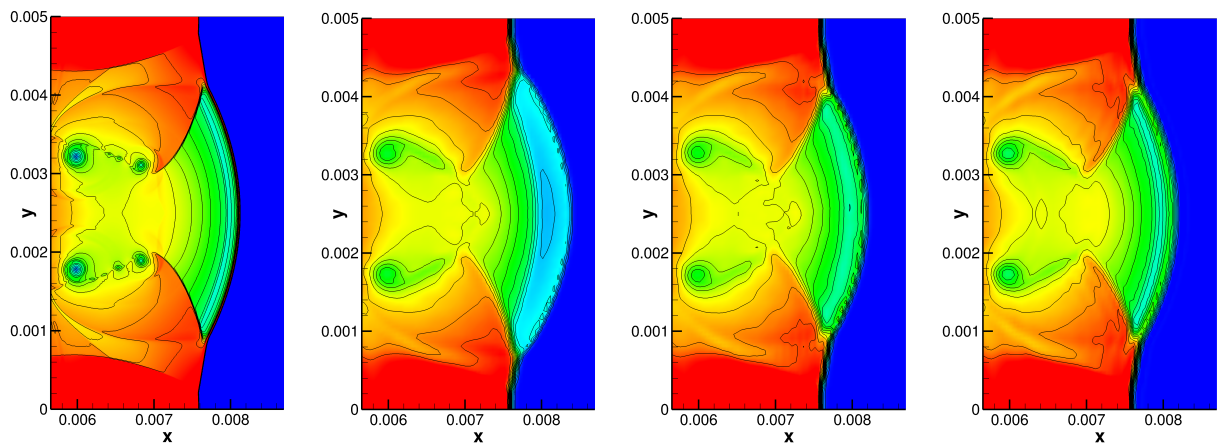


Figure 19: 2D detonation problem at  $t = 0.3 \times 10^{-7}$  and  $K_0 = 0.5825 \times 10^{10}$ : Density computed by different methods. From left to right: reference solution by the standard WENO5 method using  $4000 \times 800$  uniform grid points, WENO5, WENO5/SR and WENO5fi+split using  $500 \times 100$  uniform grid points with  $CFL = 0.05$

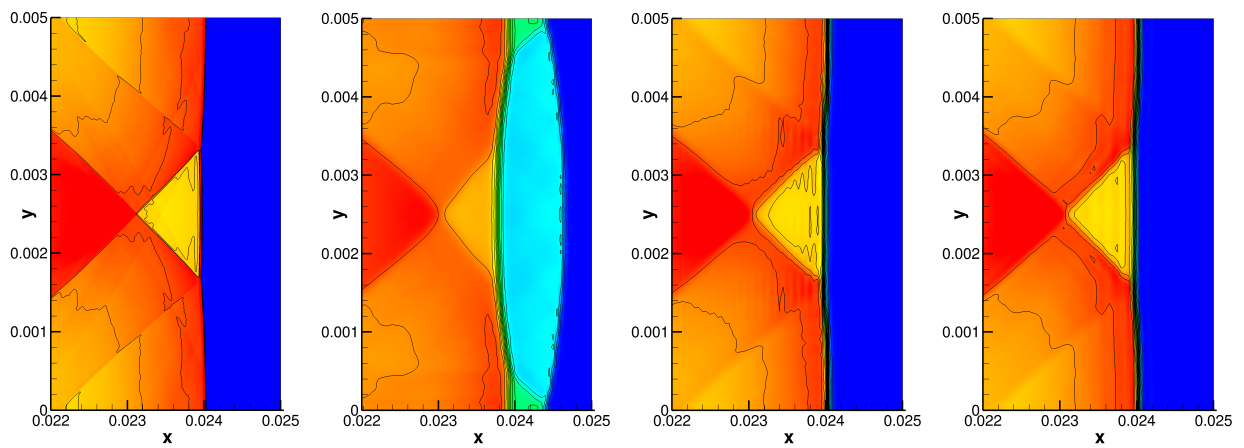


Figure 20: 2D detonation problem at  $t = 1.7 \times 10^{-7}$  and  $K_0 = 0.5825 \times 10^{10}$ : Density computed by different methods. From left to right: reference solution by the standard WENO5 method using  $4000 \times 800$  uniform grid points, WENO5, WENO5/SR and WENO5fi+split using  $500 \times 100$  uniform grid points with  $CFL = 0.05$ .

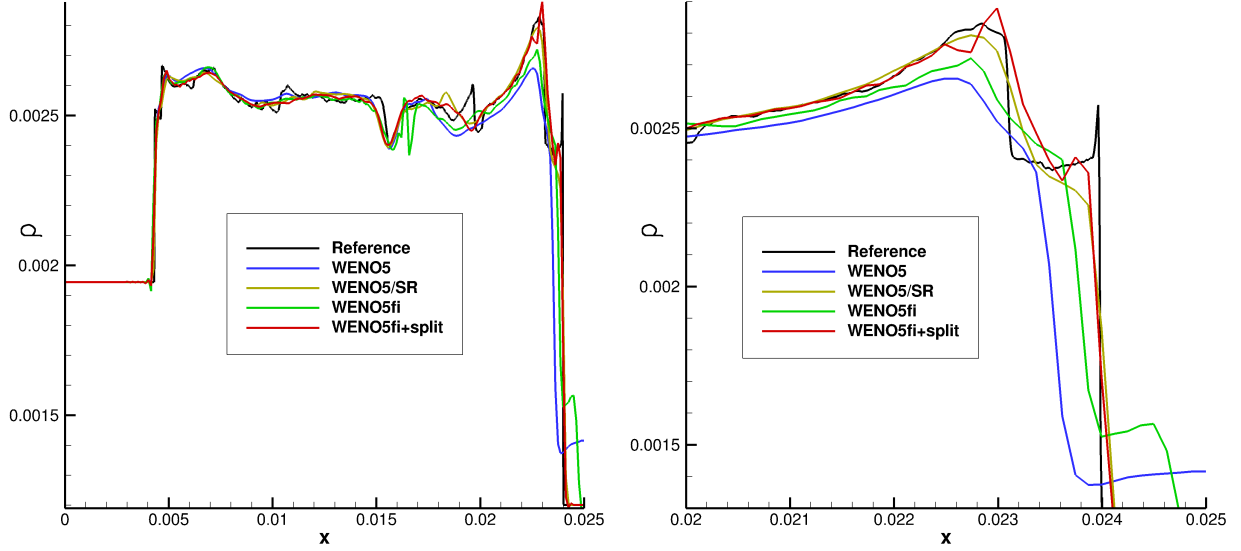


Figure 21: 1D cross-section of density at  $t = 1.7 \times 10^{-7}$  by four high order shock-capturing methods for the 2D detonation problem using  $200 \times 40$  uniform grid points,  $CFL = 0.05$  and  $K_0 = 0.5825 \times 10^{10}$ . The left figure zoomed in the vicinity of the discontinuity.

a pointwise evaluation of the source is employed for the 2D test case. Figures 19 and 20 show the density comparison among the standard WENO5 scheme, WENO5/SR and WENO5fi+split using  $500 \times 100$  uniform grid points at two different times for stiffness  $K_0 = 0.5825 \times 10^{10}$ . Figure 21 shows the density comparison among the standard WENO5 scheme, WENO5/SR, WENO5fi and WENO5fi+split using  $200 \times 40$  and  $500 \times 100$  uniform grid points. The reference solutions are computed by standard WENO5 with  $4000 \times 800$  grid points. Again, WENO5/SR and WENO5fi+split are able to obtain the correct shock speed with similar accuracy. WENO5fi gives a slightly oscillatory solution near  $x = 0.004$ . WENO5 and WENO5/SR produce no oscillations at the same location. Further improvement of the flow sensor of the filter scheme is needed in order to remove the spurious oscillations. Furthermore, for the  $500 \times 100$  grid, WENO5fi also obtained the correct shock speed. For  $CFL = 0.05$ , however, WENO5fi/SR+split is not able to obtain the correct shock speed for the stiff coefficient  $K_0$ .

#### 4.2.2. Scheme Behavior as a Function of CFL, Grid Refinement and Stiffness of the Source Terms

Figure 22 illustrates the error (number of grid points away from the reference shock location) for 128 discrete CFL values by the three high order shock-capturing schemes WENO5/SR, WENO5fi+split and WENO5fi/SR+split. The 128 discrete CFL values are ( $0.01 \leq CFL \leq 0.8$ ) with  $6.22047244094488 \times 10^{-3}$  equal increment. For this 2D case, to reduce computational cost, the smallest CFL is 0.01 instead of 0.001 in the 1D case. Figure 22 shows the error (Err) for two uniform grids  $200 \times 40$  and  $500 \times 100$  (left to right) and three stiffness coefficient  $K_0, 100K_0, 1000K_0$  (top to bottom). (Note that for the 2D case  $K_0 = 0.5825 \times 10^{10}$ .) Again, as can be seen in this figure, a similar spurious solution behavior as in the 1D detonation case carries over to the 2D detonation case. However, for this 2D case, WENO5fi+split performs better than WENO5fi/SR+split (the reverse of the 1D case). Overall, WENO5/SR and WENO5fi+split perform better than the other methods.

#### 4.3. Scheme Performance and Extreme Grid Refinement

Here, the relative CPU time performance by WENO5/SR, WENO5fi+split and WENO5fi/SR+split using the same computer and within the ADPDIS3D code by the pointwise evaluation of the source term is included. Fig. 23 shows the 1D and 2D detonation problem using 50 uniform grid for  $CFL = 0.05$  and RK4 time discretization. In all cases WENO5fi+split and WENO5fi/SR+split consume less CPU time than WENO5 and WENO5/SR, respectively. (Note that the larger the number indicated on the table implies less CPU.) Figure 24 shows the extreme refinement computation using 10,000 grid points for the 1D test case with  $CFL = 0.05$ . It appears that for this particular CFL, WENO5/SR is very close to the reference solution but with slight oscillation. WENO5fi/SR+split behaves similarly to

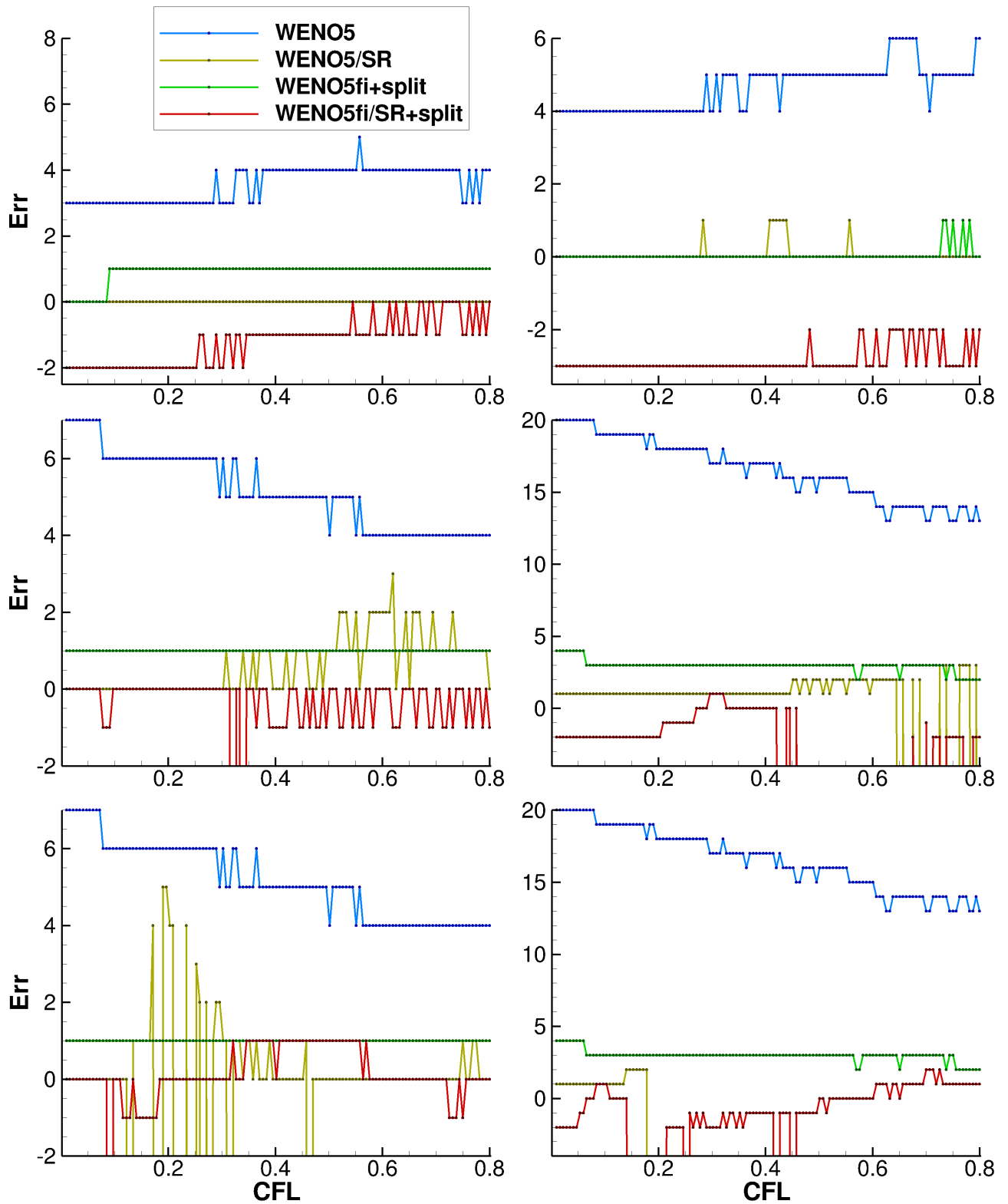


Figure 22: 2D detonation problem at  $t = 1.7 \times 10^{-7}$  and  $K_0 = 0.5825 \times 10^{10}$ : Number of grid point away from the reference shock solution as a function of the CFL number (128 discrete CFL values with  $6.22047244094488 \times 10^{-3}$  equal increment) for three low dissipative shock-capturing methods using  $200 \times 40$  and  $500 \times 100$  uniform grid points (across) and for stiffness  $K_0$ ,  $100K_0$ ,  $1000K_0$  (top to bottom). See Fig. 9 for additional captions.

# Scheme Performance

## 1D Detonation Problem (Grid 50, CFL = 0.05, RK4)

	WENO5	WENO5/SR	WENO5fi+split	WENO5fi/SR+split
CPU eff, iterations/sec	1600	1570	2050	1940
Discontinuity location error (grid points)	5	0	1	0

## 2D Detonation Problem (Grid 200x40, CFL = 0.05, RK4)

	WENO5	WENO5/SR	WENO5fi+split	WENO5fi/SR+split
CPU eff, iterations/sec	4.6	4.3	10.0	7.8
Discontinuity location max error (grid points)	21	0	0	0

Figure 23: Sample of scheme performance of WENO5, WENO5/SR, WENO5fi+split and WENO5fi/SR+split for  $CFL = 0.05$ . 50 grid points are used for the 1D case, and  $200 \times 40$  grid points are used for the 2D case with RK4 as the temporal discretization. The CPU times comparison here is based on 8 processor computations.

WENO5/SR except with an increase in small oscillations. However, WENO5fi+split and WENO7fi+ split are not able to obtain the correct shock location. This is another counter-intuitive spurious behavior of the considered schemes.

### 5. Concluding Remarks

In [45] we concluded that the filter version of the WENO5 in conjunction with the Ducros et al. splitting (WENO5fi+split) is able to obtain the correct propagation speed of discontinuities for two detonation problems. The results show that WENO5/SR and WENO5fi+split are able to obtain the correct shock speed with similar accuracy, whereas this is not the case for WENO5 & WENO5fi using the same coarse grids. Using its original form [54] without further modification, the accuracy of WENO5fi+split was found to be nearly as good as WENO5/SR. That conclusion was for one single  $CFL = 0.05$  and the original  $K_0$  stiffness. In addition, the studies in [45] focus only on solving the reactive system using the Strang splitting. The present more in-depth study also concentrates on solving the reactive system using the Strang splitting. All of the results include a cut off safeguard if densities are outside the permissible range. Spurious behavior of the same schemes by solving the fully coupled reactive system (without the Strang splitting) is reported in our companion study [18]. The main reason for the present focus study on the Strang splitting is due to the fact that it is widely used in combustion and reactive flow simulations. In addition, the high order new subcell resolution method utilizes the Strang splitting procedure and it is natural to compare among methods using the same procedure in solving the governing equations. The simple cut off safeguard procedure is also commonly used by practitioners in computational physics and engineering simulations.

With the present more extensive study the summary of the scheme behavior reported in [45] needs to be quantified. The behavior of these high order shock-capturing schemes is more complicated and does not fall in the standard non-reacting flow numerical solution behavior and practices. Aside from the accuracy of the scheme, the manner in which the spreading of discontinuities is contained plays a major role in obtaining the correct shock location. Choosing the right combination of time step and grid spacing also plays an equal role. Several counter-intuitive spurious behaviors are observed as discussed in the numerical result sections. For certain instances, smaller CFLs (not extremely small but

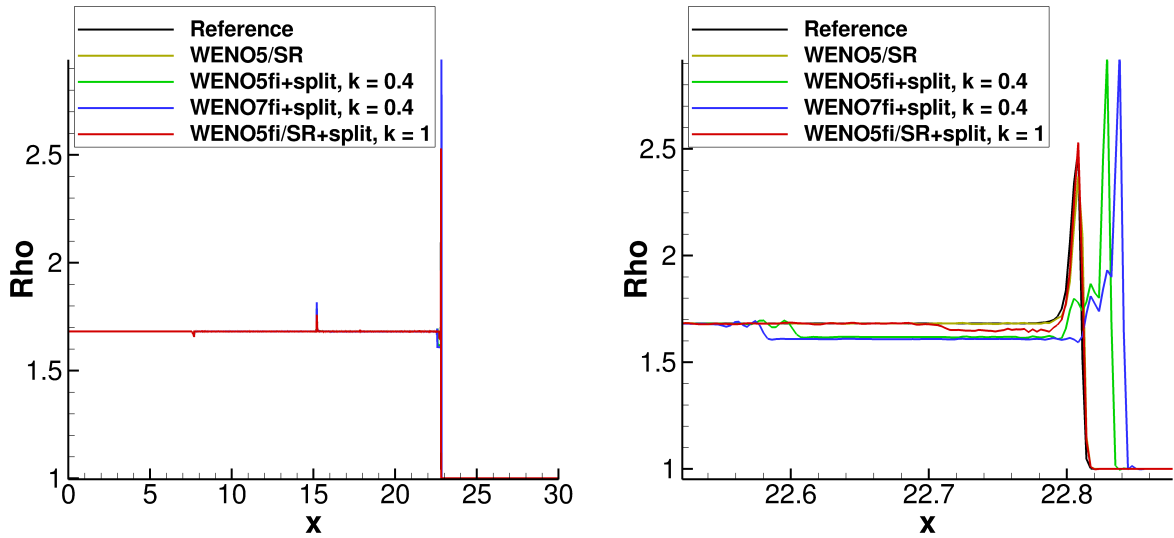


Figure 24: 1D C-J detonation problem, Arrhenius case at  $t = 1.8$ : Behavior of WENO5/SR, WENO5fi+split and WENO5fi/SR+split under extreme grid refinement with  $CFL = 0.05$  and 10,000 grid points. The value "k" is the  $\kappa$  value to control the amount of numerical dissipation indicated in the formula for the filter numerical fluxes [54].

practical for computation) exhibit more spurious behavior. Traditionally, for non-separable finite difference methods for non-reacting flow computations, a bigger CFL would give more accurate solutions for non-reacting problems, e.g., the MacCormack method. For problems with nonlinear stiff source terms using the Strang splitting and the cut off safe guide procedure, in some instances, it is the larger CFL which exhibits less spurious behavior. The results imply that the traditional concept of CFL guideline needs to be revised for this type of numerical procedure. Unlike the von Neumann analysis for constant coefficient model PDEs containing zero source terms, the linearized stability region for nonhomogeneous PDEs can consist of disjoint intervals, instead of a single continuous interval. The implication is that in practical computations where the exact values of these intervals are not known, one can easily land in regions that exhibit spurious solutions. One might suspect that our CFL guideline of using the homogeneous part of the governing equation is to blame. However, for very small CFL, the stiffness due to the reaction term has been accounted for.

In spite of the counter intuitive results, overall, the more accurate the numerical method, especially the less dissipative scheme in conjunction with the containment of spreading the discontinuity, the better the performance for very coarse grids (based on fixed grid spacing studies). It performs better than most of the previously suggested improved methods reported in the literature for problems containing stiff source terms and discontinuities. The subcell resolution method and its nonlinear filter counterparts delay the onset of wrong speed of propagation for stiffer coefficients on the same two stiff detonation test cases more than the methods reported in the literature. This study also indicated that since this type of scheme is designed for coarse grids and moderate stiff source terms, it has additional spurious behavior as the grid is refined and the stiffness is further increased. This finding might shed some light on the reported difficulties in numerical combustion and problems with stiff nonlinear (homogeneous) source terms and discontinuities in general.

In order to get a first hand examination of the behavior for practical problems, simplified EAST experiment setup simulations for a 13 species nonequilibrium flow were conducted. Due to the CPU intensive nature of the flow, less in-depth numerical investigations than for the two detonation test cases were conducted. Results indicate that the numerical method and grid dependence of the shear and shock locations are related to the stiffness of the source terms. The reason is that for non-reacting flows, numerical method and grid dependent solutions do not affect the location of the discontinuities, but rather change the degree of the smearing of the discontinuities. The implication of this exercise is to illustrate the danger of practical numerical simulation for problems containing stiff source terms where there is no reliable means of assessing the accuracy of the computed result other than by extreme grid refinement as good and reliable experimental data are not available. This extreme grid refinement approach is beyond the capability of the

current super computer for most practical simulations.

Several thoughts on the causes of the observed spurious behavior that are topics of future research are: (a) the spurious oscillations in the vicinity of discontinuities might be due to the use of Roe's average states [14], (b) the use of a stiff ODE solver with adaptive error control might alleviate some of the spurious numerics due to the reaction operator (however, it might present complications in the subcell resolution approach), and (c) as discussed in the 1D test case section, the non-pointwise evaluation of the source term that is more compatible with the convection difference operator might play a major role in minimizing spurious numerics. Studies by Lafon & Yee [20, 21] and Griffiths et al. [8] indicated that pointwise evaluation of the source term appears to be the least stable for higher than first-order numerical methods. All three of the above will be subjects of the future investigation with emphasis on (c) for higher than second-order methods.

## 6. Acknowledgments

The support of the DOE/SciDAC SAP grant DE-AI02-06ER25796 and the collaboration with B. Sjögren and A. Lani in developing the infrastructure of the nonequilibrium capability in the ADPDIS3D code used for this study are acknowledged. Insightful discussions throughout the course of this work with B. Sjögren and NASA Ames colleagues M. Panesi, A. Wray and D. Prabhu are gratefully acknowledged. The work was performed by the second author as a postdoc fellow at the Center for Turbulence Research, Stanford University. The financial support from the NASA Fundamental Aeronautics (Hypersonic) Program for the first author is gratefully acknowledged. The research of C.-W. Shu and W. Wang is partially supported by NASA grant NNX12AJ62A.

## References

- [1] W. Bao, S. Jin, The random projection method for hyperbolic conservation laws with stiff reaction terms, *J. Comp. Phys.* 163 (2000) 216–248.
- [2] M. Ben-Artzi, The generalized Riemann problem for reactive flows, *J. Comput. Phys.* 81 (1989) 70–101.
- [3] A. Berkenbosch, E. Kaasschieter, R. Klein, Detonation capturing for stiff combustion chemistry, *Combust. Theory Model.* 2 (1998) 313–348.
- [4] B. Bihari, D. Schwendeman, Multiresolution schemes for the reactive euler equations, *J. Comput. Phys.* 154 (1999) 197–230.
- [5] A. Bourlioux, A. Majda, V. Roytburd, Theoretical and numerical structure for unstable one-dimensional detonations, *SIAM J. Appl. Math.* 51 (1991) 303–343.
- [6] P. Colella, A. Majda, V. Roytburd, Theoretical and numerical structure for numerical reacting waves, *SIAM J. Sci. Stat. Comput.* 7 (1986) 1059–1080.
- [7] F. Ducros, F. Laporte, T. Soulères, V. Guinot, P. Moinat, B. Caruelle, High-order fluxes for conservative skew-symmetric-like schemes in structured meshes: Application to compressible flows, *J. Comp. Phys.* 161 (2000) 114–139.
- [8] D. Griffiths, A. Stuart, H. Yee, Numerical wave propagation in an advection equation with a nonlinear source term, *SIAM J. Numer. Anal.* 29 (1992) 1244–1260.
- [9] A. Harten, ENO schemes with subcell resolution, *J. Comp. Phys.* 83 (1989) 148–184.
- [10] C. Helzel, R. LeVeque, G. Warneke, A modified fractional step method for the accurate approximation of detonation waves, *SIAM J. Sci. Stat. Comp.* 22 (1999) 1489–1510.
- [11] A. Honein, Numerical Aspects of Compressible Turbulence Simulations, Ph.D. thesis, Stanford University, 2004.
- [12] X. Hu, N. Adams, C.W. Shu, Positivity-preserving flux limiters for high-order conservation schemes, *Journal of Computational Physics* (2012). ArXiv:1203.1540v4, in revision.
- [13] R. Jeltsch, P. Klingenstein, Error estimators for the position of discontinuities in hyperbolic conservation laws with source term which are solved using operator splitting, *Comput. Vis. Sci.* 1 (1999) 231–249.
- [14] P. Jenny, B. Muller, H. Thomann, Correction of conservative euler solvers for gas mixtures, *J. Comput. Phys.* 132 (1997) 91–107.
- [15] G.S. Jiang, C.W. Shu, Efficient implementation of weighted ENO schemes, *J. Comput. Phys.* 126 (1996) 202–228.
- [16] K. Kailasanath, E. Oran, J. Boris, T. Young, Determination of detonation cell size and the role of transverse waves in two-dimensional detonations, *Combust. Flame* 61 (1985) 199–209.
- [17] D. Kotov, H. Yee, M. Panesi, A. Wray, D. Prabhu, 1D and 2D simulation of the NASA EAST Experiments, Technical Report, Stanford University, 2012.
- [18] D. Kotov, H. Yee, W. Wang, C.W. Shu, On spurious numerics in solving reactive equations, in: *Proceedings of the ASTRONUM-2012, The Big Island, Hawaii.*
- [19] D.V. Kotov, H.C. Yee, B. Sjögren, Comparative study of high-order positivity-preserving WENO schemes, Technical Report, 2012. Stanford, California: Center for Turbulence Research.
- [20] A. Lafon, H. Yee, Dynamical approach study of spurious steady-state numerical solutions for nonlinear differential equations, Part III: The effects of nonlinear source terms in reaction-convection equations, *Intern. J. Comput. Fluid Dyn.* 6 (1996) 1–36.
- [21] A. Lafon, H. Yee, Dynamical approach study of spurious steady-state numerical solutions of nonlinear differential equations, Part IV: Stability vs. numerical treatment of nonlinear source terms, *Intern. J. Comput. Fluid Dyn.* 6 (1996) 89–123.
- [22] A. Lani, B. Sjögren, H.C. Yee, W.D. Henshaw, Variable high-ordermultiblock overlapping grid methods for mixed steady and unsteady multiscale viscous flows, part ii: Hypersonic nonequilibrium flows, *Commun. Comput. Phys.* 13 (2013) 583–602.

- [23] R. LeVeque, K.M. Shyue, One-dimensional front tracking based on high resolution wave propagation methods, *SIAM J. Sci. Comput.* 16 (1995) 348–377.
- [24] R. LeVeque, H.C. Yee, A study of numerical methods for hyperbolic conservation laws with stiff source terms, *J. Comp. Phys.* 86 (1990) 187–210.
- [25] E. McCorkle, H. Hassan, Study of radiation in electric arc shock tubes, in: 10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Chicago, Illinois.
- [26] D. Nguyen, F. Gibou, R. Fedkiw, A fully conservative ghost fluid method & stiff detonation waves, in: Proceedings of the 12th International Detonation Symposium, S. Diego, CA.
- [27] P. Olsson, Summation by parts, projections, and stability. I, *Math. Comp.* 64 (1995) 1035–1065.
- [28] P. Olsson, J. Oliger, Energy and Maximum Norm Estimates for Nonlinear Conservation Laws, Technical Report 94.01, RIACS, 1994.
- [29] M. Panesi, T. Magin, A. Bourdon, A. Bultel, O. Chazot, Study of electronically excited state populations of atoms and molecules predicted by means of a collisional-radiative model for the fire ii flight experiment, *Journal of Thermophysics and Heat Transfer* 25 (2011) 361–374.
- [30] R. Pember, Numerical methods for hyperbolic conservation laws with stiff relaxation, I. spurious solutions, *SIAM J. Appl. Math.* 53 (1993) 1293–1330.
- [31] P. Roe, Approximate riemann solvers, parameter vectors, and difference schemes, *J. Comput. Phys.* 43 (1981) 357–372.
- [32] C.W. Shu, Total-variation-diminishing time discretizations, *SIAM J. Sci. Stat. Comput.* 9 (1988) 105–121.
- [33] C.W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, II, *J. Comp. Phys.* 83 (1989) 32–78.
- [34] B. Sjögren, H. Yee, On tenth-order central spatial schemes, in: Proceedings of the Turbulence and Shear Flow Phenomena 5 (TSFP-5), Munich, Germany.
- [35] B. Sjögren, H.C. Yee, ultiresolution wavelet based adaptive numerical dissipation control for shock-turbulence computation, *J. Sci. Comput.* 20 (2004) 211–255.
- [36] B. Sjögren, H.C. Yee, On skew-symmetric splitting and entropy conservation schemes for the euler equations, in: Proc. of the 8th Euro. Conf. on Numerical Mathematics & Advanced Applications (ENUMATH 2009), Uppsala University, Uppsala, Sweden.
- [37] G. Strang, On the construction and comparison of difference schemes, *SIAM J. Numer. Anal.* 5 (1968) 506–517.
- [38] A. Stuart, A. Humphries, *Dynamical Systems and Numerical Analysis*, Cambridge Monographs on Applied and Computational Mathematics, 1998.
- [39] V. Ton, Improved shock-capturing methods for multicomponent and reacting flows, *J. Comput. Phys.* 128 (1996) 237–253.
- [40] L. Tosatto, L. Vigeveno, Numerical solution of under-resolved detonations, *J. Comp. Phys.* 227 (2008) 2317–2343.
- [41] M. Vinokur, H. Yee, Extension of efficient low dissipative high-order schemes for 3D curvilinear moving grids, *Frontiers of Computational Fluid dynamics*, World Scientific (2002) 129–164.
- [42] W. Wang, C. Shu, H.C. Yee, B. Sjögren, High order finite difference methods with subcell resolution for advection equations with stiff source terms, *J. Comput. Phys.* 231 (2012) 190–214.
- [43] W. Wang, H.C. Yee, B. Sjögren, T. Magin, C. Shu, Construction of low dissipative high-order well-balanced filter schemes for nonequilibrium flows, *J. Comput. Phys.* 230 (2011) 4316–4335.
- [44] H. Yee, G. Klopfer, J.L. Montagne, High-resolution shock-capturing schemes for inviscid and viscous hypersonic flows, *J. Comput. Phys.* 80 (1990) 31–61.
- [45] H. Yee, D. Kotov, B. Sjogreen, Numerical dissipation and wrong propagation speed of discontinuities for stiff source terms, in: Proceedings of the ASTRONUM-2011, Valencia, Spain, pp. June 13–17, 2011.
- [46] H. Yee, B. Sjögren, Designing adaptive low dissipative high order schemes for long-time integrations, in: D.D..B. Geurts (Ed.), *Turbulent Flow Computation*, Kluwer Academic, 2002.
- [47] H. Yee, B. Sjögren, Development of low dissipative high order filter schemes for multiscale navier-stokes/MHD systems, *J. Comput. Phys.* 225 (2007) 910–934.
- [48] H. Yee, P. Sweby, D. Griffiths, Dynamical approach study of spurious steady-state numerical solutions for nonlinear differential equations, Part I: The dynamics of time discretizations and its implications for algorithm development in computational fluid dynamics, *J. Comput. Phys.* 97 (1991) 249–310.
- [49] H. Yee, M. Vinokur, M. Djomehri, Entropy splitting and numerical dissipation, *J. Comput. Phys.* 162 (2000) 33–81.
- [50] H.C. Yee, A class of high-resolution explicit and implicit shock-capturing methods, VKI lecture series 1989-04, 1989.
- [51] H.C. Yee, Building blocks for reliable complex nonlinear numerical simulations, in: D.D..B. Geurts (Ed.), *Turbulent Flow Computation*, Kluwer Academic, 2002.
- [52] H.C. Yee, N. Sandham, M. Djomehri, Low dissipative high order shock-capturing methods using characteristic-based filters, *J. Comput. Phys.* 150 (1999) 199–238.
- [53] H.C. Yee, B. Sjögren, Adaptive filtering and limiting in compact high order methods for multiscale gas dynamics and MHD systems, *Computers & Fluids* 37 (2008) 593–619.
- [54] H.C. Yee, B. Sjögren, High order filter methods for wide range of compressible flow speeds, in: Proc. of ICOSAHOM 09 (International Conference on Spectral and High Order Methods), Trondheim, Norway.
- [55] H.C. Yee, B. Sjögren, A. Hadjadj, Comparative study of high order schemes for les of temporal-evolving mixing layers, in: Proc. of ASTRONUM-2010, San Diego, Calif.
- [56] H.C. Yee, P. Sweby, Dynamics of numerics & spurious behaviors in CFD computations, in: Keynote paper, 7th ISCFD Conf., Beijing, China. RIACS Technical Report 97.06, June 1997.
- [57] H.C. Yee, J. Torczynski, S. Morton, M. Visbal, P. Sweby, On spurious behavior of CFD simulations, *Int. J. Num. Meth. Fluids* 30 (1999) 675–711.
- [58] X. Zhang, C.W. Shu, Positivity-preserving high order finite difference weno schemes for compressible euler equations, *J. Comput. Phys.* 231 (2012) 2245–2258.