A HIGH ORDER ACCURATE BOUND-PRESERVING COMPACT FINITE DIFFERENCE SCHEME FOR TWO-DIMENSIONAL INCOMPRESSIBLE FLOW

HAO LI AND XIANGXIONG ZHANG *

Abstract. For solving two-dimensional incompressible flow in the vorticity form by the fourth-order compact finite difference scheme and explicit strong stability preserving (SSP) temporal discretizations, we show that the simple bound-preserving limiter in [5] can enforce the strict bounds of the vorticity, if the velocity field satisfies a discrete divergence free constraint. For reducing oscillations, a modified TVB limiter adapted from [2] is constructed without affecting the bound-preserving property. This bound-preserving finite difference method can be used for any passive convection equation with a divergence free velocity field.

Key words. Finite difference, monotonicity, bound-preserving, discrete maximum principle, passive convection, incompressible flow, total variation bounded limiter.

AMS subject classifications. 65M06, 65M12

1. Introduction.

In this paper, we are interested in constructing high order compact finite difference schemes solving the following two-dimensional time-dependent incompressible Euler equation in vorticity and stream-function formulation

$$\omega_t + (u\omega)_x + (v\omega)_y = 0, \tag{1.1a}$$

$$\psi = \Delta\omega, \tag{1.1b}$$

$$\langle u, v\rangle = \langle -\psi_y, \psi_x\rangle, \tag{1.1c}$$

with periodic boundary conditions and suitable initial conditions. In the above formulation, ω is the vorticity, ψ is the stream function, $\langle u, v\rangle$ is the velocity and Re is the Reynolds number.

For simplicity, we only focus on the incompressible Euler equation (1.1). With explicit time discretizations, the extension of high order accurate bound-preserving compact finite difference scheme to Navier-Stokes equation

$$\omega_t + (u\omega)_x + (v\omega)_y = \frac{1}{Re}\Delta\omega \tag{1.2}$$

would be straightforward following the approach in [5].

The equation (1.1c) implies the incompressbility condition

$$u_x + v_y = 0. \tag{1.3}$$

Due to (1.3), (1.1a) is equivalent to

$$\omega_t + u\omega_x + v\omega_y = 0 \tag{1.4}$$

for which the initial value problem satisfies a bound-preserving property:

$$\min_{x,y}\omega(x,y,0) = m \le \omega(x,y,t) \le M = \max_{x,y}\omega(x,y,0).$$

*Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067 (li2497@purdue.edu, zhan1966@purdue.edu). Research is supported by NSF DMS-1913120.

If solving (1.4) directly, it is usually easier to construct a bound-preserving scheme. For the sake of conservation, it is desired to solve the conservative form equation (1.1a). The divergence free constraint (1.3) is one of the main difficulties in solving incompressible flows. In order to enforce the bound-preserving property for (1.1a) without losing accuracy, the incompressibility condition must be properly used since the bound-preserving property may not hold for (1.1a) without (1.3), see [9, 8, 10].

Even though the bound-preserving property and the global conservation imply certain nonlinear stability, in practice a bound-preserving high order accurate compact finite difference scheme can still produce excessive oscillations for a pure convection problem. Thus an additional limiter for reducing oscillations is often needed, e.g., the total variation bounded (TVB) limiter discussed in [2]. One of the main focuses of this paper is to design suitable TVB type limiters, without losing bound-preserving property. Notice that the TVB limiter for a compact finite difference scheme is designed in a quite different way from those for discontinuous Galerkin method, thus it is nontrivial to have a bound-preserving TVB limiter for the compact finite difference schemes.

The paper is organized as follows. Section 2 is a review of the compact finite difference method and a simple bound-preserving limiter for scalar convection-diffusion equations. In Section 3, we show that the compact finite difference scheme can be rendered bound-preserving if the velocity field satisfies a discrete divergence free condition. We discuss the bound-preserving property of a TVB limiter in Section 4. Numerical tests are shown in Section 5. Concluding remarks are given in Section 6.

2. Review of compact finite difference method.

In this section we review the compact finite difference method and a bound-preserving limiter in [5].

2.1. A fourth-order accurate compact finite difference scheme.

Consider a smooth function $f(x)$ on the interval $[0,1]$. Let $x_i = \frac{i}{N}$ $(i = 1, \cdots, N)$ be the uniform grid points on the interval $[0,1]$. A fourth-order accurate compact finite difference approximation to derivatives on the interval $[0,1]$ is given as:

$$
\begin{aligned}
\frac{1}{6}(f'_{i+1} + 4f'_i + f'_{i-1}) &= \frac{f_{i+1} - f_{i-1}}{2\Delta x} + \mathcal{O}(\Delta x^4),\\
\frac{1}{12}(f''_{i+1} + 4f''_i + f''_{i-1}) &= \frac{f_{i+1} - 2f_{i-1} + f_{i-1}}{\Delta x^2} + \mathcal{O}(\Delta x^4),
\end{aligned}
\tag{2.1}
$$

where f_i, $f'_i$and f''_i are point values of a function $f(x)$, its derivative $f'(x)$ and its second order derivative $f''(x)$ at uniform grid points x_i $(i = 1, \cdots, N)$ respectively.

Let $\mathbf{f}$ be a column vector with numbers $f_1, f_2, \cdots, f_N$ as entries. Let W_1, W_2, D_x and D_{xx} denote four linear operators as follows:

$$
W_1\mathbf{f} = \frac{1}{6}\begin{pmatrix} 4 & 1 & & & 1 \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ 1 & & & 1 & 4 \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix}, D_x\mathbf{f} = \frac{1}{2}\begin{pmatrix} 0 & 1 & & & -1 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ 1 & & & -1 & 0 \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix},
\tag{2.2}
$$

(2.3)
$$W_2\mathbf{f} = \frac{1}{12}\begin{pmatrix} 10 & 1 & & & 1 \\ 1 & 10 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 10 & 1 \\ 1 & & & 1 & 10 \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix}, D_{xx}\mathbf{f} = \begin{pmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & & & 1 & -2 \end{pmatrix}\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix}.$$

If $f(x)$ is periodic with with period 1, the fourth-order compact finite difference approximation (2.1) to the first order derivative and second order derivative can be denoted as

$$W_1\mathbf{f}' = \frac{1}{\Delta x}D_x\mathbf{f}, \quad W_2\mathbf{f}'' = \frac{1}{\Delta x^2}D_{xx}\mathbf{f},$$

which can be explicitly written as

$$\mathbf{f}' = \frac{1}{\Delta x}W_1^{-1}D_x\mathbf{f}, \quad \mathbf{f}'' = \frac{1}{\Delta x^2}W_2^{-1}D_{xx}\mathbf{f},$$

where W_1^{-1} and W_2^{-1} are the inverse operators. For convenience, by abusing notations we let $W_1^{-1}f_i$ denote the i-th entry of the vector $W_1^{-1}\mathbf{f}$.

2.2. High order time discretizations. For time discretizations, we use the strong stability preserving (SSP) Runge-Kutta and multistep methods, which are convex combinations of formal forward Euler steps. Thus we only need to discuss the bound-preserving for one forward Euler step since convex combination can preserve the bounds.

For the numerical tests in this paper, we use a third order explicit SSP Runge–Kutta method SSPRK(3,3), see [3], which is widely known as the Shu-Osher method, with SSP coefficient $C = 1$ and effective SSP coefficient $C_{eff} = \frac{1}{3}$. For solving $u_t = F(u)$, it is given by

$$\begin{aligned} u^{(1)} &= u^n, \\ u^{(2)} &= u^{(1)} + dtF(u^{(1)}), \\ u^{(3)} &= \frac{3}{4}u^{(1)} + \frac{1}{4}(u^{(2)} + F(u^{(2)})), \\ u^{n+1} &= \frac{1}{3}u^{(1)} + \frac{2}{3}(u^{(3)} + F(u^{(3)})). \end{aligned}$$

2.3. A three-point stencil bound-preserving limiter. In this subsection, we review the three-point stencil bound-preserving limiter in [5]. Given a sequence of periodic point values u_i $(i = 1, \cdots, N)$, $u_0 := u_N$, $u_{N+1} := u_1$ and constant $a \geq 2$, assume all local weighted averages are in the range $[m, M]$:

$$m \leq \frac{1}{a+2}(u_{i-1} + au_i + u_{i+1}) \leq M, \quad i = 1, \cdots, N, \quad a \geq 2.$$

We separate the point values $\{u_i, i = 1, \cdots, N\}$ into two classes of subsets consisting of consecutive point values. In the following discussion, a *set* refers to a set of consecutive point values $u_l, u_{l+1}, u_{l+2}, \cdots, u_{m-1}, u_m$. For any set $S = \{u_l, u_{l+1}, \cdots, u_{m-1}, u_m\}$, we call the first point value u_l and the last point value u_m as *boundary points*, and call the other point values $u_{l+1}, \cdots, u_{m-1}$ as *interior points*. A set of class I is defined as a set satisfying the following:

1. It contains at least four point values.
2. Both *boundary points* are in $[m, M]$ and all *interior points* are out of range.
3. It contains both undershoot and overshoot points.

Notice that in a set of class I, at least one undershoot point is next to an overshoot point. For given point values $u_i, i = 1, \cdots, N$, suppose all the sets of class I are $S_1 = \{u_{m_1}, u_{m_1+1}, \cdots, u_{n_1}\}$, $S_2 = \{u_{m_2}, \cdots, u_{n_2}\}$, $\cdots$, $S_K = \{u_{m_K}, \cdots, u_{n_K}\}$, where $m_1 < m_2 < \cdots < u_{m_K}$.

A set of class II consists of point values between S_i and S_{i+1} and two boundary points u_{n_i} and $u_{m_{i+1}}$. Namely they are $T_0 = \{u_1, u_2, \cdots, u_{m_1}\}$, $T_1 = \{u_{n_1}, \cdots, u_{m_2}\}$, $T_2 = \{u_{n_2}, \cdots, u_{m_3}\}, \cdots, T_K = \{u_{n_K}, \cdots, u_N\}$. For periodic data u_i, we can combine T_K and T_0 to define $T_K = \{u_{n_K}, \cdots, u_N, u_1, \cdots, u_{m_1}\}$.

In the sets of class I, the undershoot and the overshoot are neighbors. In the sets of class II, the undershoot and the overshoot are separated, i.e., an overshoot is not next to any undershoot. As a matter of fact, in the numerical tests, the sets of class I are hardly encountered. Here we include them in the discussion for the sake of completeness. When there are no sets of class I, all point values form a single set of class II.

Algorithm 2.1 A bound-preserving limiter for periodic data u_i satisfying $\bar{u}_i \in [m, M]$

Require: the input u_i satisfies $\bar{u}_i = \frac{1}{a+2}(u_{i-1} + a u_i + u_{i+1}) \in [m, M]$, $a \geq 2$. Let u_0, u_{N+1} denote u_N, u_1 respectively.

Ensure: the output satisfies $v_i \in [m, M], i = 1, \cdots, N$ and $\sum_{i=1}^{N} v_i = \sum_{i=1}^{N} u_i$.

Step 0: First set $v_i = u_i$, $i = 1, \cdots, N$. Let v_0, v_{N+1} denote v_N, v_1 respectively.

Step I: Find all the sets of class I $S_1, \cdots, S_K$ (all local saw-tooth profiles) and all the sets of class II $T_1, \cdots, T_K$.

Step II: For each T_j $(j = 1, \cdots, K)$,

for all index i in T_j **do**

  if $u_i < m$ **then**

    $v_{i-1} \leftarrow v_{i-1} - \frac{(u_{i-1}-m)_+}{(u_{i-1}-m)_+ + (u_{i+1}-m)_+}(m - u_i)_+$

    $v_{i+1} \leftarrow v_{i+1} - \frac{(u_{i+1}-m)_+}{(u_{i-1}-m)_+ + (u_{i+1}-m)_+}(m - u_i)_+$

    $v_i \leftarrow m$

  end if

  if $u_i > M$ **then**

    $v_{i-1} \leftarrow v_{i-1} + \frac{(M-u_{i-1})_+}{(M-u_{i-1})_+ + (M-u_{i+1})_+}(u_i - M)_+$

    $v_{i+1} \leftarrow v_{i+1} + \frac{(M-u_{i+1})_+}{(M-u_{i-1})_+ + (M-u_{i+1})_+}(u_i - M)_+$

    $v_i \leftarrow M$

  end if

end for

Step III: for each saw-tooth profile $S_j = \{u_{m_j}, \cdots, u_{n_j}\}$ $(j = 1, \cdots, K)$, let N_0 and N_1 be the numbers of undershoot and overshoot points in S_j respectively.

Set $U_j = \sum_{i=m_j}^{n_j} v_i$.

for $i = m_j + 1, \cdots, n_j - 1$ **do**

  if $u_i > M$ **then**

    $v_i \leftarrow M$.

  end if

  if $u_i < m$ **then**

    $v_i \leftarrow m$.

  end if

end for

Set $V_j = N_1 M + N_0 m + v_{m_j} + v_{n_j}$.

Set $A_j = v_{m_j} + v_{n_j} + N_1 M - (N_1 + 2)m$, $B_j = (N_0 + 2)M - v_{m_j} - v_{n_j} - N_0 m$.

if $V_j - U_j > 0$ **then**

  for $i = m_j, \cdots, n_j$ **do**

    $v_i \leftarrow v_i - \frac{v_i - m}{A_j}(V_j - U_j)$

  end for

else

  for $i = m_j, \cdots, n_j$ **do**

    $v_i \leftarrow v_i + \frac{M - v_i}{B_j}(U_j - V_j)$

  end for

end if

The algorithm 2.1 can enforce $\bar{u}_i \in [m, M]$ without losing conservation [5]:

THEOREM 1. *Assume periodic data $u_i (i = 1, \cdots, N)$ satisfies $\bar{u}_i = \frac{1}{a+2}(u_{i-1} + a u_i + u_{i+1}) \in [m, M]$ for some fixed $a \geq 2$ and all $i = 1, \cdots, N$ with $u_0 := u_N$ and*

$u_{N+1} := u_1$, *then the output of Algorithm 2.1 satisfies* $\sum_{i=1}^{N} v_i = \sum_{i=1}^{N} u_i$ *and* $v_i \in [m, M]$, $\forall i$.

For the two-dimensional case, the same limiter can be used in a dimension by dimension fashion to enforce $u_{ij} \in [m, M]$.

3. A bound-preserving scheme for the two-dimensional incompressible flow. In this section we first show the fourth-order compact finite difference with forward Euler time discretization satisfies the weak monotonicity [5], thus it is bound-preserving with a naturally constructed discrete divergence-free velocity field.

For simplicity, we only consider a periodic boundary condition on a square $[0,1]\times[0,1]$. Let $(x_i, y_j) = (\frac{i}{N_x}, \frac{j}{N_y})$ $(i = 1, \cdots, N_x, j = 1, \cdots, N_y)$ be the uniform grid points on the domain $[0,1]\times[0,1]$. All notation in this paper is consistent with those in [5].

3.1. Weak monotonicity and bound-preserving. Let $\lambda_1 = \frac{\Delta t}{\Delta x}$ and $\lambda_2 = \frac{\Delta t}{\Delta y}$, the fourth-order compact finite difference scheme with the forward Euler method for (1.1a) can be given as

$$\omega_{ij}^{n+1} = \omega_{ij}^{n} - \lambda_1 [W_{1x}^{-1} D_x(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij} - \lambda_2 [W_{1y}^{-1} D_y(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij}. \tag{3.1}$$

With the same notation as in [5], the weighted average in two dimensions can be denoted as

$$\bar{\omega} = W_{1x} W_{1y} \omega. \tag{3.2}$$

Then the scheme (3.1) is equivalent to

$$\begin{aligned} \bar{\omega}_{ij}^{n+1} &= \bar{\omega}_{ij}^{n} - \lambda_1 [W_{1y} D_x(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij} - \lambda_2 [W_{1x} D_y(\mathbf{v}^n \circ \boldsymbol{\omega}^n)]_{ij} \\ &= \frac{1}{36}\begin{pmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{pmatrix} : \Omega^n - \frac{\lambda_1}{12}\begin{pmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -1 & 0 & 1 \end{pmatrix} : (U^n \circ \Omega^n) - \frac{\lambda_2}{12}\begin{pmatrix} 1 & 4 & 1 \\ 0 & 0 & 0 \\ -1 & -4 & -1 \end{pmatrix} : (V^n \circ \Omega^n), \end{aligned} \tag{3.3}$$

where $\circ$ denotes the matrix Hadamard product, and

$$U = \begin{pmatrix} u_{i-1,j+1} & u_{i,j+1} & u_{i+1,j+1} \\ u_{i-1,j} & u_{i,j} & u_{i+1,j} \\ u_{i-1,j-1} & u_{i,j-1} & u_{i+1,j-1} \end{pmatrix}, V = \begin{pmatrix} v_{i-1,j+1} & v_{i,j+1} & v_{i+1,j+1} \\ v_{i-1,j} & v_{i,j} & v_{i+1,j} \\ v_{i-1,j-1} & v_{i,j-1} & v_{i+1,j-1} \end{pmatrix},$$

$$\Omega = \begin{pmatrix} \omega_{i-1,j+1} & \omega_{i,j+1} & \omega_{i+1,j+1} \\ \omega_{i-1,j} & \omega_{i,j} & \omega_{i+1,j} \\ \omega_{i-1,j-1} & \omega_{i,j-1} & \omega_{i+1,j-1} \end{pmatrix}.$$

It is straightforward to verify the *weak monotonicity*, i.e., $\bar{\omega}_{ij}^{n+1}$ is a monotonically increasing function with respect to all point values ω_{ij}^n involved in (3.3) under the CFL condition

$$\frac{\Delta t}{\Delta x} \max_{ij} |u_{ij}^n| + \frac{\Delta t}{\Delta y} \max_{ij} |v_{ij}^n| \leq \frac{1}{3}.$$

However, the monotonicity is sufficient for bound-preserving $\bar{\omega}_{ij}^{n+1} \in [m, M]$, only if the following consistency condition holds:

$$\omega_{ij}^n \equiv m \Rightarrow \bar{\omega}_{ij}^{n+1} = m, \quad \omega_{ij}^n \equiv M \Rightarrow \bar{\omega}_{ij}^{n+1} = M. \tag{3.4}$$

Plugging $\omega_{ij}^n \equiv m$ in (3.3), we get

$$\bar{\omega}_{ij}^{n+1} = m\left(1 - \lambda_1 \left(W_{1y} D_x \mathbf{u}^n\right)_{ij} - \lambda_2 \left(W_{1x} D_y \mathbf{v}^n\right)_{ij}\right).$$

Thus the consistency (3.4) holds only if the velocity $\langle \mathbf{u}^n, \mathbf{v}^n \rangle$ satisfies:

$$\frac{1}{\Delta x} D_x W_{1y} \mathbf{u}^n + \frac{1}{\Delta x} D_y W_{1x} \mathbf{v}^n = 0. \tag{3.5}$$

Therefore we have the following bound-preserving result:

THEOREM 2. *If the velocity $\langle \mathbf{u}^n, \mathbf{v}^n \rangle$ satisfies the discrete divergence free constraint (3.5) and $\omega_{ij}^n \in [m, M]$, then under the CFL constraint*

$$\frac{\Delta t}{\Delta x} \max_{ij} |u_{ij}^n| + \frac{\Delta t}{\Delta y} \max_{ij} |v_{ij}^n| \le \frac{1}{3},$$

the scheme (3.3) satisfies $\bar{\omega}_{ij}^{n+1} \in [m, M]$.

3.2. A discrete divergence free velocity field. In the following discussion, we may discard the superscript n for convenience assuming everything discussed is at time step n.

Note that (3.5) is a discrete divergence free constraint and we can construct a fourth-order accurate velocity field satisfying (3.5). Given ω_{ij}, we first compute ψ_{ij} by a fourth-order compact finite difference scheme for the Poisson equation (1.1b). The detail of the Poisson solvers including the fast Poisson solver is given in the appendices.

With the fourth-order compact finite difference we have

$$-\frac{1}{\Delta y} D_y \mathbf{\Psi} = W_{1y} \mathbf{u}, \quad \frac{1}{\Delta x} D_x \mathbf{\Psi} = W_{1x} \mathbf{v}, \tag{3.6}$$

where

$$\mathbf{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1,N_y} \\ \psi_{21} & \psi_{22} & \cdots & \psi_{2,N_y} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N_x-1,1} & \psi_{N_x-1,2} & \cdots & \psi_{N_x-1,N_y} \\ \psi_{N_x,1} & \psi_{N_x,2} & \cdots & \psi_{N_x,N_y} \end{pmatrix}_{N_x \times N_y}.$$

Since the two finite difference operators D_x and D_y commute, it is straightforward to verify that the velocity field computed by (3.6) satisfies (3.5).

3.3. A fourth-order accurate bound-preserving scheme. For the Euler equations (1.1), the following implementation of the fourth-order compact finite difference with forward Euler time discretization scheme can preserve the bounds:

1. Given $\omega_{ij}^n \in [m, M]$, solve the Poisson equation (1.1b) by the fourth-order accurate compact finite difference scheme to obtain point values of the stream function ψ_{ij}.
2. Construct $\mathbf{u}$ and $\mathbf{v}$ by (3.6).
3. Obtain $\bar{\omega}_{ij}^{n+1} \in [m, M]$ by scheme (3.3).
4. Apply the limiting procedure in Section 2.3 to obtain $\omega_{ij}^{n+1} \in [m, M]$.

For high order SSP time discretizations, we should use the same implementation above for each time stage or time step.

For the Navier-Stokes equations (1.2), with $\mu_1 = \frac{\Delta t}{\Delta x^2}$ and $\mu_2 = \frac{\Delta t}{\Delta y^2}$, the scheme can be written as

$$\begin{aligned} \omega_{ij}^{n+1} =& \omega_{ij}^{n} - \lambda_1 [W_{1x}^{-1} D_x(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij} - \lambda_2 [W_{1y}^{-1} D_y(\mathbf{v}^n \circ \boldsymbol{\omega}^n)]_{ij} \\ &+ \frac{\mu_1}{Re} W_{2x}^{-1} D_{xx} \omega_{ij}^n + \frac{\mu_2}{Re} W_{2y}^{-1} D_{yy} \omega_{ij}^n, \end{aligned} \tag{3.7}$$

In a manner similar to (3.2), we define

$$\tilde{\omega} := W_{2x} W_{2y} \omega, \tag{3.8}$$

with $W_1 := W_{1x} W_{1y}$ and $W_2 := W_{2x} W_{2y}$. Due to definition (3.2) and the fact operators W_1 and W_2 commute, i.e. $W_1 W_2 = W_2 W_1$, we have

$$\tilde{\bar{\omega}} = W_2 W_1 \omega = W_1 W_2 \omega = \bar{\tilde{\omega}}.$$

Then scheme (3.7) is equivalent to

$$\begin{aligned} \tilde{\bar{\omega}}_{ij}^{n+1} =& \tilde{\bar{\omega}}_{ij}^{n} - \frac{\lambda_1}{12} [W_2 W_{1y} Dx(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij} - \frac{\lambda_2}{12} [W_2 W_{1x} Dy(\mathbf{u}^n \circ \boldsymbol{\omega}^n)]_{ij} \\ &+ \frac{\mu_1}{Re} W_1 W_{2y} D_{xx} \omega_{ij}^n + \frac{\mu_2}{Re} W_1 W_{2x} D_{yy} \omega_{ij}^n. \end{aligned} \tag{3.9}$$

Following the discussion in Section 3.1 and the discussion for the two-dimensional convection-diffusion in [5], we have the following result:

THEOREM 3. *If the velocity $\langle \mathbf{u}^n, \mathbf{v}^n \rangle$ satisfies the constraint (3.5) and $\omega_{ij}^n \in [m, M]$, then under the CFL constraint*

$$\frac{\Delta t}{\Delta x} \max_{ij} |u_{ij}^n| + \frac{\Delta t}{\Delta y} \max_{ij} |v_{ij}^n| \leq \frac{1}{6}, \quad \frac{\Delta t}{Re \Delta x^2} + \frac{\Delta t}{Re \Delta y^2} \leq \frac{5}{24},$$

the scheme (3.9) satisfies $\tilde{\bar{\omega}}_{ij}^{n+1} \in [m, M]$.

Given $\tilde{\bar{\omega}}_{ij}$, we can recover point values ω_{ij} by obtaining first $\tilde{\omega}_{ij} = W_1^{-1} \tilde{\bar{\omega}}_{ij}$ then $\omega_{ij} = W_2^{-1} \tilde{\omega}_{ij}$. Given point values ω_{ij} satisfying $\tilde{\bar{\omega}}_{ij} \in [m, M]$ for any i and j, we can use the limiter in Algorithm 2.1 in a dimension by dimension fashion several times to enforce $\omega_{ij} \in [m, M]$:

1. Given $\tilde{\bar{\omega}}_{ij} \in [m, M]$, compute $\tilde{\omega}_{ij} = W_1^{-1} \tilde{\bar{\omega}}_{ij}$ and apply the limiting Algorithm 2.1 with $a = 4$ in both x-direction and y-direction to ensure $\tilde{\omega}_{ij} \in [m, M]$.
2. Given $\bar{\omega}_{ij} \in [m, M]$, compute $\omega_{ij} = W_2^{-1} \tilde{\omega}_{ij}$ and apply the limiting algorithm Algorithm 2.1 with $a = 10$ in both x-direction and y-direction to ensure $\omega_{ij} \in [m, M]$.

4. A TVB limiter for the two-dimensional incompressible flow. To have nonlinear stability and eliminate oscillations for shocks, a TVBM (total variation bounded in the means) limiter was introduced for the compact finite difference scheme solving scalar convection equations in [2]. In this section, we will modify this limiter for the incompressible flow so that it does not affect the bound-preserving property. Thus we can use both the TVB limiter and the bound-preserving limiter in Algorithm 2.1 to preserve bounds while reducing oscillations. For simplicity, we only consider the numerical scheme for the incompressible Euler equations (1.1). In this section, we may discard the superscript n if a variable is defined at time step n.

4.1. The TVB limiter. The scheme (3.3) can be written in a conservative form:

$$\bar{\omega}_{ij}^{n+1} = \bar{\omega}_{ij}^{n} - \lambda_1[(\hat{u\omega})_{i+\frac{1}{2},j}^{n} - (\hat{u\omega})_{i-\frac{1}{2},j}^{n}] - \lambda_2[(\hat{v\omega})_{i,j+\frac{1}{2}}^{n} - (\hat{v\omega})_{i,j-\frac{1}{2}}^{n}], \tag{4.1}$$

involving a numerical flux $(\hat{u\omega})_{i+\frac{1}{2},j}^{n}$ and $(\hat{v\omega})_{i,j+\frac{1}{2}}^{n}$ as local functions of u_{kl}^n, v_{kl}^n and ω_{kl}^n. The numerical flux is defined as

$$\begin{aligned}(\hat{u\omega})_{i+\frac{1}{2},j} &= \frac{1}{2}\left([W_{1y}(\mathbf{u}\circ\boldsymbol{\omega})]_{ij} + [W_{1y}(\mathbf{u}\circ\boldsymbol{\omega})]_{i+1,j}\right),\\ (\hat{v\omega})_{i,j+\frac{1}{2}} &= \frac{1}{2}\left([W_{1x}(\mathbf{v}\circ\boldsymbol{\omega})]_{ij} + [W_{1x}(\mathbf{v}\circ\boldsymbol{\omega})]_{i,j+1}\right).\end{aligned} \tag{4.2}$$

Similarly we denote

$$\begin{aligned}\hat{u}_{i+\frac{1}{2},j} &= \frac{1}{2}\Big((W_{1y}\mathbf{u})_{ij} + (W_{1y}\mathbf{u})_{i+1,j}\Big),\\ \hat{v}_{i,j+\frac{1}{2}} &= \frac{1}{2}\Big((W_{1x}\mathbf{v})_{ij} + (W_{1x}\mathbf{v})_{i,j+1}\Big).\end{aligned} \tag{4.3}$$

The limiting is defined in a dimension by dimension manner. For the flux splitting, it is done as in one-dimension. Consider a splitting of u satisfying

$$u^+ \geq 0, \quad u^- \leq 0. \tag{4.4}$$

The simplest such splitting is the Lax-Friedrichs splitting

$$u^{\pm} = \frac{1}{2}(u \pm \alpha), \quad \alpha = \max_{(x,y)\in\Omega} |u(x,y)|.$$

Then we have

$$u = u^+ + u^-, \quad u\omega = u^+\omega + u^-\omega,$$

and we write the flux $(\hat{u\omega})_{i+\frac{1}{2},j}$ and $\hat{u}_{i+\frac{1}{2},j}$ as

$$(\hat{u\omega})_{i+\frac{1}{2},j} = (\hat{u\omega})_{i+\frac{1}{2},j}^{+} + (\hat{u\omega})_{i+\frac{1}{2},j}^{-}, \quad \hat{u}_{i+\frac{1}{2},j} = \hat{u}_{i+\frac{1}{2},j}^{+} + \hat{u}_{i+\frac{1}{2},j}^{-}$$

where $(\hat{u\omega})_{i+\frac{1}{2},j}^{\pm}$ and $\hat{u}_{i+\frac{1}{2},j}^{\pm}$ are obtained by adding superscripts $\pm$ to u_{ij} in (4.2) and (4.3) respectively, i.e.

$$\begin{aligned}(\hat{u\omega})_{i+\frac{1}{2},j}^{\pm} =& \frac{1}{2}\left([W_{1y}(\mathbf{u}^{\pm}\circ\boldsymbol{\omega})]_{ij} + [W_{1y}(\mathbf{u}^{\pm}\circ\boldsymbol{\omega})]_{i+1,j}\right),\\ \hat{u}_{i+\frac{1}{2},j}^{\pm} =& \frac{1}{2}\Big(\big(W_{1y}\mathbf{u}^{\pm}\big)_{ij} + \big(W_{1y}\mathbf{u}^{\pm}\big)_{i+1,j}\Big),\end{aligned}$$

where $\mathbf{u}^{\pm} = (u_{ij}^{\pm})$. With a dummy index j referring y value, we first take the differences between the high-order numerical flux and the first-order upwind flux

$$d(\hat{u\omega})_{i+\frac{1}{2},j}^{+} = (\hat{u\omega})_{i+\frac{1}{2},j}^{+} - u_{i+\frac{1}{2},j}^{+}\bar{\omega}_{ij}, \quad d(\hat{u\omega})_{i+\frac{1}{2},j}^{-} = u_{i+\frac{1}{2},j}^{-}\bar{\omega}_{i+1,j} - (\hat{u\omega})_{i+\frac{1}{2},j}^{-}. \tag{4.5}$$

Limit them by

$$\begin{aligned}d(\hat{u\omega})_{i+\frac{1}{2},j}^{+(m)} &= m\left(d(\hat{u\omega})_{i+\frac{1}{2},j}^{+},\, u_{i+\frac{1}{2},j}^{+}\Delta_+^x\bar{\omega}_{ij},\, u_{i-\frac{1}{2},j}^{+}\Delta_+^x\bar{\omega}_{i-1,j}\right),\\ d(\hat{u\omega})_{i+\frac{1}{2},j}^{-(m)} &= m\left(d(\hat{u\omega})_{i+\frac{1}{2},j}^{-},\, u_{i+\frac{1}{2},j}^{-}\Delta_+^x\bar{\omega}_{ij},\, u_{i+\frac{3}{2},j}^{-}\Delta_+^x\bar{\omega}_{i+1,j}\right),\end{aligned} \tag{4.6}$$

where $\Delta_+^x v_{ij} \equiv v_{i+1,j} - v_{ij}$ is the forward difference operator in the x direction, and m is the standard *minmod* function

$$m(a_1, \ldots, a_k) = \begin{cases} s \min_{1 \le i \le k} |a_i|, & \text{if } sign(a_1) = \cdots = sign(a_k) = s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

As mentioned in [2], the limiting defined in (4.6) maintains the formal accuracy of the compact schemes in smooth regions of the solution with the assumption

$$\bar{\omega}_{ij} = (W_{1x} W_{1y} \omega)_{ij} = \omega_{ij} + \mathcal{O}\left(\Delta x^2\right) \text{ for } \omega \in C^2. \tag{4.8}$$

Under the assumption (4.8), by simple Taylor expansion,

$$\begin{aligned} d(\hat{u\omega})^{\pm}_{i+\frac{1}{2},j} &= \frac{1}{2} u^{\pm}_{i+\frac{1}{2},j} \omega_{x,ij} \Delta x + \mathcal{O}\left(\Delta x^2\right), \\ u^{\pm}_{k+\frac{1}{2},j} \Delta^x_+ \bar{\omega}_{kj} &= u^{\pm}_{i+\frac{1}{2},j} \omega_{x,ij} \Delta x + \mathcal{O}\left(\Delta x^2\right), \quad k = i-1, i, i+1. \end{aligned} \tag{4.9}$$

Hence in smooth regions away from critical points of ω, for sufficiently small Δx, the minmod function (4.7) will pick the first argument, yielding

$$d(\hat{u\omega})^{\pm(m)}_{i+\frac{1}{2},j} = d(\hat{u\omega})^{\pm}_{i+\frac{1}{2},j}.$$

Since the accuracy may degenerate to first-order at critical points, as a remedy, the modified *minmod* function [7, 1] is introduced

$$\tilde{m}(a_1, \ldots, a_k) = \begin{cases} a_1, & \text{if } |a_1| \le P \Delta x^2, \\ m(a_1, \ldots, a_k), & \text{otherwise,} \end{cases} \tag{4.10}$$

where P is a positive constant independent of Δx and m is the standard *minmod* function (4.7). See more detailed discussion in [2].

Then we obtain the limited numerical fluxes as

$$(\hat{u\omega})^{+(m)}_{i+\frac{1}{2},j} = u^{+}_{i+\frac{1}{2},j} \bar{\omega}_{ij} + d(\hat{u\omega})^{+(m)}_{i+\frac{1}{2},j}, \quad (\hat{u\omega})^{-(m)}_{i+\frac{1}{2},j} = u^{-}_{i+\frac{1}{2},j} \bar{\omega}_{i+1,j} - d(\hat{u\omega})^{-(m)}_{i+\frac{1}{2},j}. \tag{4.11}$$

and

$$(\hat{u\omega})^{(m)}_{i+\frac{1}{2},j} = (\hat{u\omega})^{+(m)}_{i+\frac{1}{2},j} + (\hat{u\omega})^{-(m)}_{i+\frac{1}{2},j} \tag{4.12}$$

The flux in the y-direction can be defined analogously.

The following result was proven in [2]:

LEMMA 4.1. *For any n and Δt such that $0 \le n\Delta t \le T$, scheme (4.1) with flux (4.12) satisfies a maximum principle in the means:*

$$\max_{i,j} \left|\bar{\omega}^{n+1}_{ij}\right| \le \max_{i,j} \left|\bar{\omega}^{n}_{ij}\right|$$

under the CFL condition

$$\left[\max\left(u^+\right) + \max\left(-u^-\right)\right] \frac{\Delta t}{\Delta x} + \left[\max\left(v^+\right) + \max\left(-v^-\right)\right] \frac{\Delta t}{\Delta y} \le \frac{1}{2}$$

where the maximum is taken in $\min_{i,j} u^n_{ij} \le u \le \max_{i,j} u^n_{ij}$, $\min_{i,j} v^n_{ij} \le v \le \max_{i,j} v^n_{ij}$. ∎

4.2. The bound-preserving property of the nonlinear scheme with modified flux. The compact finite difference scheme with the TVB limiter in the last section is

$$\bar{\omega}_{ij}^{n+1} = \bar{\omega}_{ij}^{n} - \lambda_1 \left((\hat{u\omega})_{i+\frac{1}{2},j}^{(m)} - (\hat{u\omega})_{i-\frac{1}{2},j}^{(m)} \right) - \lambda_2 \left((\hat{v\omega})_{i,j+\frac{1}{2}}^{(m)} - (\hat{v\omega})_{i,j-\frac{1}{2}}^{(m)} \right), \tag{4.13}$$

where the numerical flux $(\hat{u\omega})_{i+\frac{1}{2},j}^{(m)}$, $(\hat{u\omega})_{i,j+\frac{1}{2}}^{(m)}$ is the modified flux approximating (4.2).

THEOREM 4. *If $\omega_{ij}^n \in [m, M]$, under the CFL condition*

$$\lambda_1 \max_{i,j} \left| u_{ij}^{(\pm)} \right| \leq \frac{1}{24}, \quad \lambda_2 \max_{i,j} \left| v_{ij}^{(\pm)} \right| \leq \frac{1}{24}, \tag{4.14}$$

the nonlinear scheme (4.13) satisfies

$$\bar{\omega}_{ij}^{n+1} \in [m, M] .$$

Proof. We have

$$\begin{aligned}
&\bar{\omega}_{ij}^{n+1} = \bar{\omega}_{ij}^{n} - \lambda_1 \left((\hat{u\omega})_{i+\frac{1}{2},j}^{(m)} - (\hat{u\omega})_{i-\frac{1}{2},j}^{(m)} \right) - \lambda_2 \left((\hat{v\omega})_{i,j+\frac{1}{2}}^{(m)} - (\hat{v\omega})_{i,j-\frac{1}{2}}^{(m)} \right) \\
=&\frac{1}{8} \Bigg((\bar{\omega}_{ij}^{n} - 8\lambda_1 (\hat{u\omega})_{i+\frac{1}{2},j}^{+(m)}) + (\bar{\omega}_{ij}^{n} - 8\lambda_1 (\hat{u\omega})_{i+\frac{1}{2},j}^{-(m)}) + (\bar{\omega}_{ij}^{n} + 8\lambda_1 (\hat{u\omega})_{i-\frac{1}{2},j}^{+(m)}) + (\bar{\omega}_{ij}^{n} + 8\lambda_1 (\hat{u\omega})_{i-\frac{1}{2},j}^{-(m)}) \\
&+ (\bar{\omega}_{ij}^{n} - 8\lambda_2 (\hat{v\omega})_{i,j+\frac{1}{2}}^{+(m)}) + (\bar{\omega}_{ij}^{n} - 8\lambda_2 (\hat{v\omega})_{i,j+\frac{1}{2}}^{-(m)}) + (\bar{\omega}_{ij}^{n} + 8\lambda_2 (\hat{v\omega})_{i,j-\frac{1}{2}}^{+(m)}) + (\bar{\omega}_{ij}^{n} + 8\lambda_2 (\hat{v\omega})_{i,j-\frac{1}{2}}^{-(m)}) \Bigg) .
\end{aligned} \tag{4.15}$$

Under the CFL condition (4.14), we will prove that the eight terms satisfy the following bounds

$$\begin{aligned}
\bar{\omega}_{ij}^{n} - 8\lambda_1 (\hat{u\omega})_{i+\frac{1}{2},j}^{+(m)} &\in \left[m - 8\lambda_1 \hat{u}_{i+\frac{1}{2},j}^{+} m, M - 8\lambda_1 \hat{u}_{i+\frac{1}{2},j}^{+} M \right], \\
\bar{\omega}_{ij}^{n} - 8\lambda_1 (\hat{u\omega})_{i+\frac{1}{2},j}^{-(m)} &\in \left[m - 8\lambda_1 \hat{u}_{i+\frac{1}{2},j}^{-} m, M - 8\lambda_1 \hat{u}_{i+\frac{1}{2},j}^{-} M \right], \\
\bar{\omega}_{ij}^{n} + 8\lambda_1 (\hat{u\omega})_{i-\frac{1}{2},j}^{+(m)} &\in \left[m + 8\lambda_1 \hat{u}_{i-\frac{1}{2},j}^{+} m, M + 8\lambda_1 \hat{u}_{i-\frac{1}{2},j}^{+} M \right], \\
\bar{\omega}_{ij}^{n} + 8\lambda_1 (\hat{u\omega})_{i-\frac{1}{2},j}^{-(m)} &\in \left[m + 8\lambda_1 \hat{u}_{i-\frac{1}{2},j}^{-} m, M + 8\lambda_1 \hat{u}_{i-\frac{1}{2},j}^{-} M \right], \\
\bar{\omega}_{ij}^{n} - 8\lambda_2 (\hat{v\omega})_{i,j+\frac{1}{2}}^{+(m)} &\in \left[m - 8\lambda_2 \hat{v}_{i,j+\frac{1}{2}}^{+} m, M - 8\lambda_2 \hat{v}_{i,j+\frac{1}{2}}^{+} M \right], \\
\bar{\omega}_{ij}^{n} - 8\lambda_2 (\hat{v\omega})_{i,j+\frac{1}{2}}^{-(m)} &\in \left[m - 8\lambda_2 \hat{v}_{i,j+\frac{1}{2}}^{-} m, M - 8\lambda_2 \hat{v}_{i,j+\frac{1}{2}}^{-} M \right], \\
\bar{\omega}_{ij}^{n} + 8\lambda_2 (\hat{v\omega})_{i,j-\frac{1}{2}}^{+(m)} &\in \left[m + 8\lambda_2 \hat{v}_{i,j-\frac{1}{2}}^{+} m, M + 8\lambda_2 \hat{v}_{i,j-\frac{1}{2}}^{+} M \right], \\
\bar{\omega}_{ij}^{n} + 8\lambda_2 (\hat{v\omega})_{i,j-\frac{1}{2}}^{-(m)} &\in \left[m + 8\lambda_2 \hat{v}_{i,j-\frac{1}{2}}^{-} m, M + 8\lambda_2 \hat{v}_{i,j-\frac{1}{2}}^{-} M \right].
\end{aligned} \tag{4.16}$$

For (4.16), by taking the sum of the lower bounds and upper bounds and multiplying them by $\frac{1}{8}$, we obtain

$$\bar{\omega}_{ij}^{n+1} \in [m - mO_{ij}, M - MO_{ij}] , \tag{4.17}$$

with

$$
\begin{aligned}
O_{ij} =&\lambda_1(\hat{u}_{i+\frac{1}{2},j} - \hat{u}_{i-\frac{1}{2},j}) - \lambda_2(\hat{u}_{i,j+\frac{1}{2}} - \hat{u}_{i,j-\frac{1}{2}})\\
=&\frac{\lambda_1}{2}\left((W_{1y}\mathbf{u})_{i+1,j} - (W_{1y}\mathbf{u})_{i-1,j})\right) + \frac{\lambda_2}{2}\left((W_{1y}\mathbf{v})_{i,j+1} - (W_{1y}\mathbf{v})_{i,j-1})\right)\\
=&\frac{\Delta t}{2}(D_x W_{1y}\mathbf{u} + D_y W_{1x}\mathbf{v}) = 0.
\end{aligned} \tag{4.18}
$$

Therefore, we conclude $\bar{\omega}_{ij}^{n+1} \in [m, M]$.

We only discuss the first two term in (4.16) since the proof for the rest is similar. By the definition of the modified *minmod* function (4.10) and (4.11), we have

$$
\begin{aligned}
&(\hat{u\omega})^{+(m)}_{i+\frac{1}{2},j} \in \left[\min\{(\hat{u\omega})^{+}_{i+\frac{1}{2},j}, u^{+}_{i+\frac{1}{2},j}\bar{\omega}_{ij}\}, \max\{(\hat{u\omega})^{+}_{i+\frac{1}{2},j}, u^{+}_{i+\frac{1}{2},j}\bar{\omega}_{ij}\}\right],\\
&(\hat{u\omega})^{-(m)}_{i+\frac{1}{2},j} \in \left[\min\{(\hat{u\omega})^{-}_{i+\frac{1}{2},j}, u^{-}_{i+\frac{1}{2},j}\bar{\omega}_{i+1,j}\}, \max\{(\hat{u\omega})^{-}_{i+\frac{1}{2},j}, u^{-}_{i+\frac{1}{2},j}\bar{\omega}_{i+1,j}\}\right].
\end{aligned} \tag{4.19}
$$

We notice that under CFL condition (4.14),

$$
\bar{\omega}^n_{ij} - 8\lambda_1(\hat{u\omega})^{+}_{i+\frac{1}{2},j}, \quad \bar{\omega}^n_{ij} - 8\lambda_1 u^{+}_{i+\frac{1}{2},j}\bar{\omega}^n_{ij}, \quad \bar{\omega}^n_{ij} - 8\lambda_1(\hat{u\omega})^{-}_{i+\frac{1}{2},j} \tag{4.20}
$$

are all monotonically increasing functions with respect to variables ω^n_{kj}, $k = i-1, i, i+1$. Due to the flux splitting (4.4),

$$
\bar{\omega}^n_{ij} - 8\lambda_1 u^{-}_{i+\frac{1}{2},j}\bar{\omega}^n_{i+1,j} \tag{4.21}
$$

is also a monotonically increasing function with respect to variables ω^n_{kj}, $k = i-1, i, i+1, i+2$. Therefore, with the assumption $\omega^n_{ij} \in [m, M]$, we obtain

$$
\begin{aligned}
\bar{\omega}^n_{ij} - 8\lambda_1(\hat{u\omega})^{+}_{i+\frac{1}{2},j}, \quad \bar{\omega}^n_{ij} - 8\lambda_1 u^{+}_{i+\frac{1}{2},j}\bar{\omega}^n_{ij} \in \left[m - 8\lambda_1\hat{u}^{+}_{i+\frac{1}{2},j}m, M - 8\lambda_1\hat{u}^{+}_{i+\frac{1}{2},j}M\right],\\
\bar{\omega}^n_{ij} - 8\lambda_1(\hat{u\omega})^{-}_{i+\frac{1}{2},j}, \quad \bar{\omega}^n_{ij} - 8\lambda_1 u^{-}_{i+\frac{1}{2},j}\bar{\omega}^n_{i+1,j} \in \left[m - 8\lambda_1\hat{u}^{-}_{i+\frac{1}{2},j}m, M - 8\lambda_1\hat{u}^{-}_{i+\frac{1}{2},j}M\right],
\end{aligned} \tag{4.22}
$$

with (4.19) , which implies the first two terms of (4.16). ∎

REMARK 1. *We remark here the above proof is independent of the second and third arguments of the minmod function* (4.10)*. Therefore, the proof hold for other limiters with different second and third arguments in the same minmod function* (4.10)*.*

REMARK 2. *The TVB limiter in this paper is designed to modify the convection flux only thus it also applies to the Navier-Stokes equation. Moreover, under suitable CFL condition, the full scheme with TVB limiter can still preserve $\tilde{\bar{\omega}}^{n+1}_{ij} \in [m, M]$ with $\omega^n_{ij} \in [m, M]$.*

4.3. An alternative TVB limiter. Another TVB limiter can be defined by replacing (4.6) with

$$
\begin{aligned}
d(\hat{u\omega})^{+(m)}_{i+\frac{1}{2},j} &= m\left(d(\hat{u\omega})^{+}_{i+\frac{1}{2},j},\, \Delta^{+}_{x}(u^{+}_{i+\frac{1}{2},j}\bar{\omega}_{ij}),\, \Delta^{+}_{x}(u^{+}_{i-\frac{1}{2},j}\bar{\omega}_{i-1,j})\right),\\
d(\hat{u\omega})^{-(m)}_{i+\frac{1}{2},j} &= m\left(d(\hat{u\omega})^{-}_{i+\frac{1}{2},j},\, \Delta^{+}_{x}(u^{-}_{i-\frac{1}{2},j}\bar{\omega}_{i,j}),\, \Delta^{+}_{x}(u^{-}_{i+\frac{1}{2},j}\bar{\omega}_{i+1,j})\right).
\end{aligned} \tag{4.23}
$$

All the other procedures in the limiter are exactly the same as in Section 4.1. The limiter does not affect the bound-preserving property due to the arguments in Remark 1.

5. Numerical Tests. In this subsection, we test the fourth-order compact finite difference scheme with both the bound-preserving and the TVB limiter for the two-dimensional incompressible flow.

In the numerical test, we refer to the bound-preserving limiter as BP, the TVB limiter in Section 4.1 as TVB1, and the TVB limiter in section 4.3 as TVB2. The parameter in the minmod function used in TVB limiters is denoted as P. In all the following numerical tests, we use SSPRK(3,3) as mentioned in section 2.2.

5.1. Accuracy Test. For the Euler Equation (1.1) with periodic boundary condition and initial data $\omega(x, y, 0) = -2\sin(2x)\sin(y)$ on the domain $[0, 2\pi]\times[0, 2\pi]$, the exact solution is $\omega(x, y, t) = -2\sin(2x)\sin(y)$. We test the accuracy of the proposed scheme on this solution. The errors for $P = 300$ are given in Table 1, and we observe the designed order of accuracy for this special steady state solution.

Table 1: Incompressible Euler equations. Fourth-order compact FD for vorticity, $t = 0.5$. With BP and TVB1 limiters, P =300.

$N \times N$	L^2 error	order	L^∞ error	order
32×32	3.16E-3	-	1.00E-3	-
64×64	1.86E-4	4.09	5.90E-5	4.09
128×128	1.14E-5	4.02	3.63E-6	4.02
256×256	7.13E-7	4.01	2.67E-7	4.00

5.2. Double Shear Layer Problem. We test the scheme for the double shear layer problem on the domain $[0, 2\pi] \times [0, 2\pi]$ with a periodic boundary condition. The initial condition is

$$\omega(x, y, 0) = \begin{cases} \delta cos(x) - \frac{1}{\rho} sech^2((y - \frac{\pi}{2})/\rho),\ y \le \pi \\ \delta cos(x) + \frac{1}{\rho} sech^2((\frac{3\pi}{2} - y)/\rho), y > \pi \end{cases}$$

with $\delta = 0.05$ and $\rho = \pi/15$. The vorticity ω at time $T = 6$ and $T = 8$ are shown in Figure 1, Figure 2 and Figure 3. With both the bound-preserving limiter and TVB limiter, the numerical solutions are ensured to be in the range $[-\delta - \frac{1}{\rho}, \delta + \frac{1}{\rho}]$. The TVB limiter can also reduce oscillations.

5.3. Vortex Patch Problem.. We test the limiters for the vortex patch problem in the domain $[0, 2\pi] \times [0, 2\pi]$ with a periodic boundary condition. The initial condition is

$$\omega(x, y, 0) = \begin{cases} -1, & (x, y) \in [\frac{\pi}{2}, \frac{3\pi}{2}] \times [\frac{\pi}{4}, \frac{3\pi}{4}]; \\ 1, & (x, y) \in [\frac{\pi}{2}, \frac{3\pi}{2}] \times [\frac{5\pi}{4}, \frac{7\pi}{4}]; \\ 0, & \text{otherwise.} \end{cases}$$

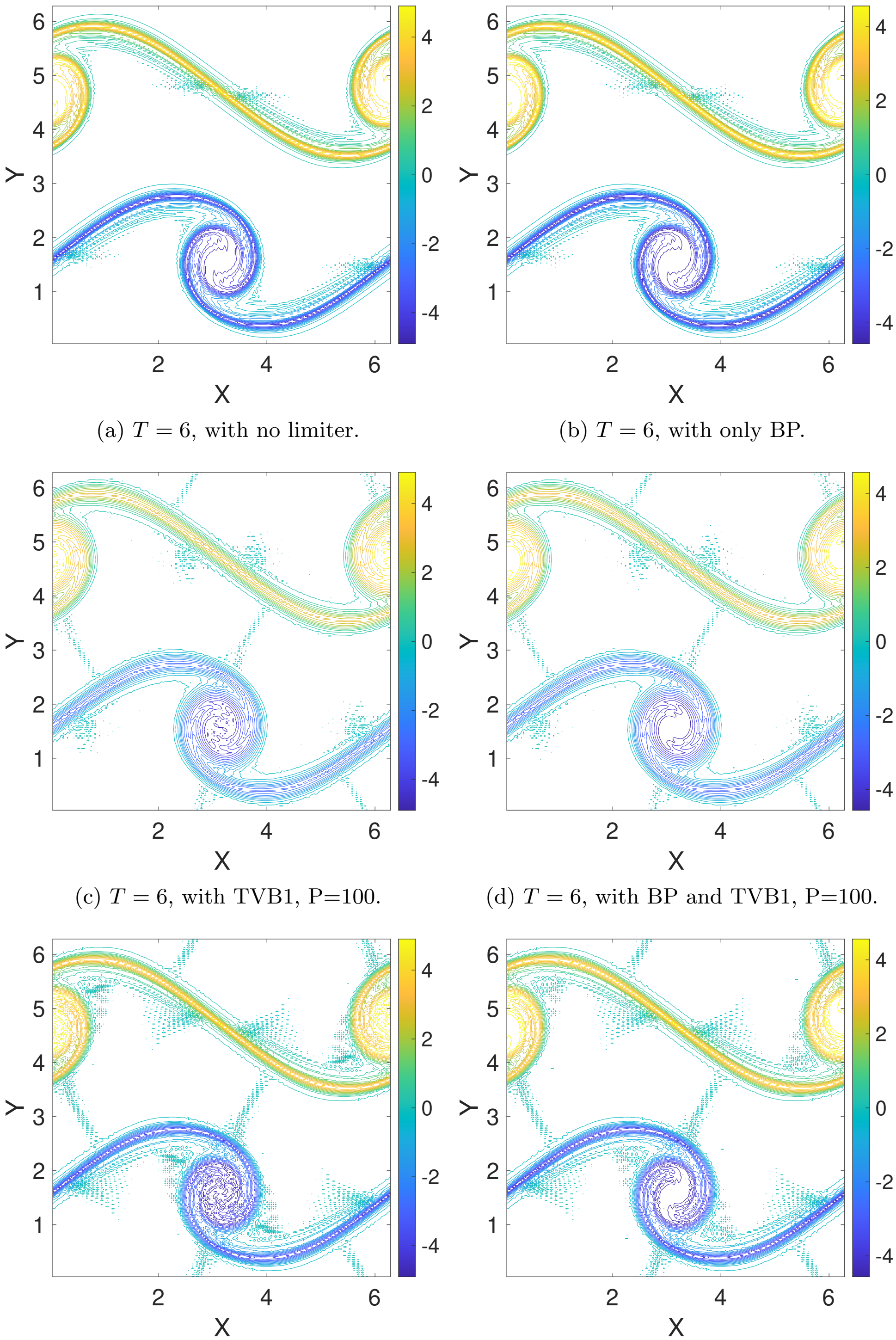
Numerical solutions for incompressible Euler are shown in Figure 4, Figure 5, Figure 6 and Figure 7. We can observe that the solutions generated by the compact finite difference scheme with only the bound-preserving limiter are still highly oscillatory for the Euler equation without the TVB limiter.

Notice that the oscillations in Figure 4 suggest that the artificial viscosity induced by the bound-preserving limiter is quite low.

6. Concluding Remarks. We have proven that a simple limiter can preserve bounds for the fourth-order compact finite difference method solving the two dimensional incompressible Euler equation, with a discrete divergence-free velocity field.

We also prove that the TVB limiter modified from [2] does not affect the bound-preserving property of $\bar{\omega}$. With both the TVB limiter and the bound-preserving limiter, the numerical solutions of high order compact finite difference scheme can be rendered non-oscillatory and strictly bound-preserving.

For the sixth-order and eighth-order compact finite difference method for convection problem with weak monotonicty in [5], the divergence-free velocity can be constructed accordingly, which gives a higher order bound-preserving scheme for the incompressible flow by applying Algorithm 2.1 for several times. The TVB limiting procedure in Section 4.1 can also be defined with a similar result as Theorem 4.



(a) $T = 6$, with no limiter.

(b) $T = 6$, with only BP.

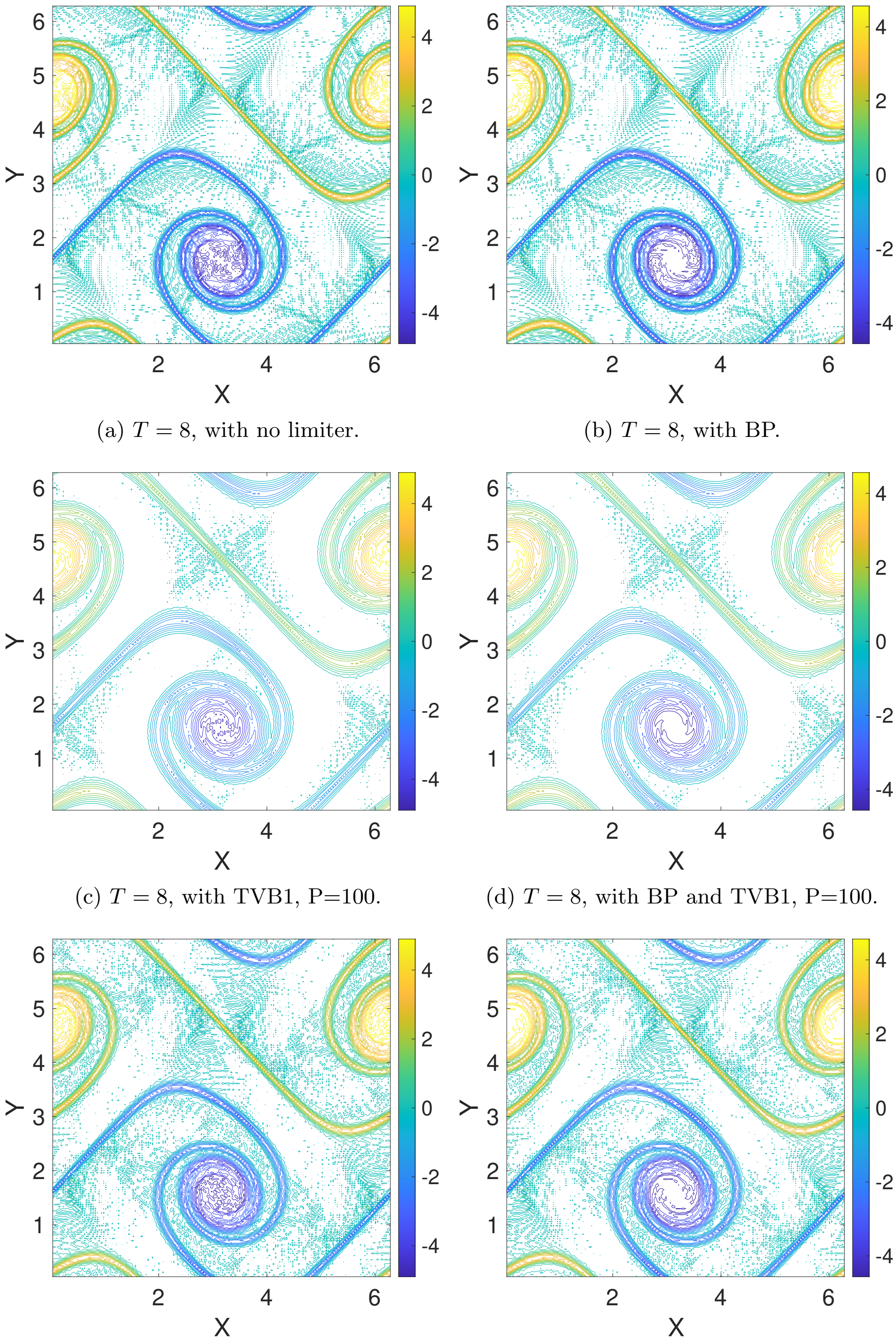
(c) $T = 6$, with TVB1, P=100.

(d) $T = 6$, with BP and TVB1, P=100.

(e) $T = 6$, with TVB1, P=300.

(f) $T = 6$, with BP and TVB1, P=300.

Fig. 1: Double shear layer problem. Fourth-order compact finite difference with SSP Runge–Kutta method on a 160×160 mesh solving the incompressible Euler equation (1.1) at $T = 6$. The time step is $\Delta t = \frac{1}{24 \max_x |\mathbf{u}_0|} \Delta x$.



(a) $T = 8$, with no limiter.

(b) $T = 8$, with BP.

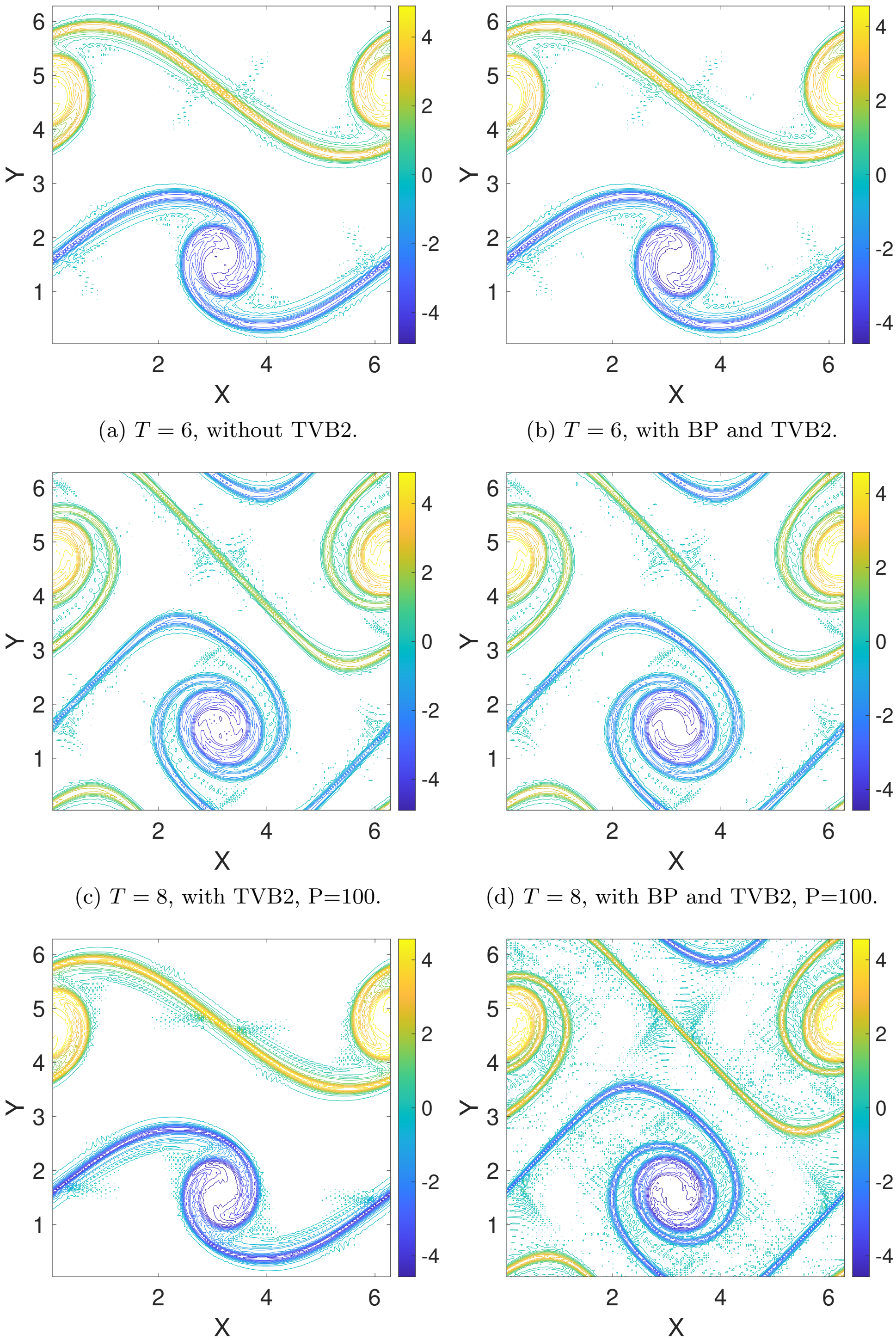
(c) $T = 8$, with TVB1, P=100.

(d) $T = 8$, with BP and TVB1, P=100.

(e) $T = 8$, with TVB1, P=300.

(f) $T = 8$, with BP and TVB1, P=300.

Fig. 2: Double shear layer problem. Fourth-order compact finite difference with SSP Runge–Kutta method on a 160×160 mesh solving the incompressible Euler equation (1.1) at $T = 8$. The time step is $\Delta t = \frac{1}{24 \max_x |\mathbf{u}_0|} \Delta x$.



(a) $T = 6$, without TVB2.

(b) $T = 6$, with BP and TVB2.

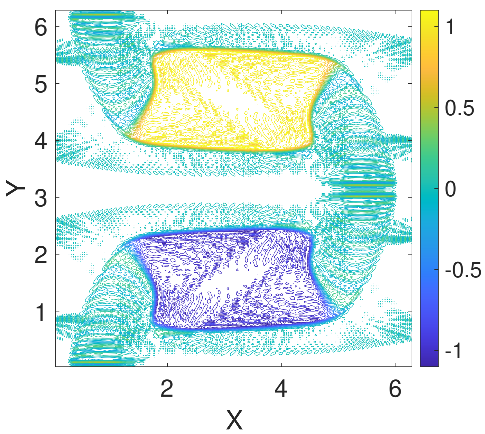
(c) $T = 8$, with TVB2, P=100.

(d) $T = 8$, with BP and TVB2, P=100.

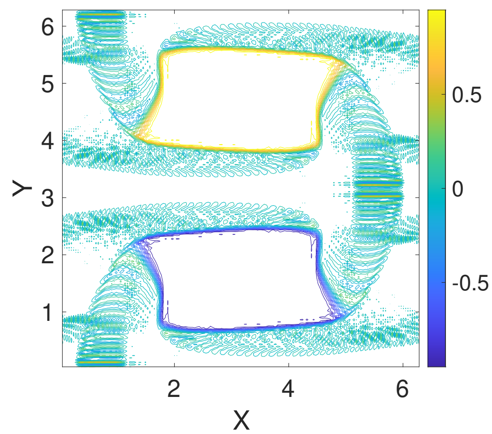
(e) $T = 6$, with BP and TVB2, P=300.

(f) $T = 8$, with BP and TVB2, P=300.

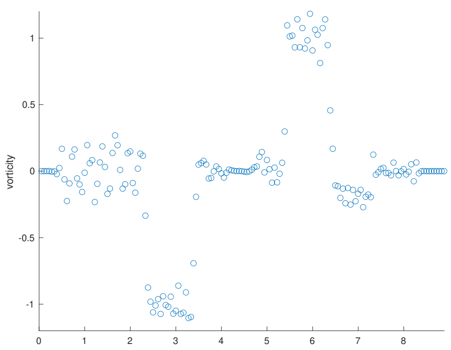
Fig. 3: Double shear layer problem. Fourth-order compact finite difference with SSP Runge–Kutta method on a 160×160 mesh solving the incompressible Euler equation (1.1) at $T = 6$ and $T = 8$. The time step is $\Delta t = \frac{1}{24 \max_x |\mathbf{u}_0|} \Delta x$.



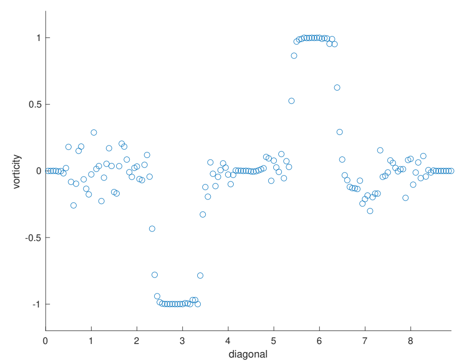
(a) $T = 5$, with no limiter.



(b) $T = 5$, with BP.

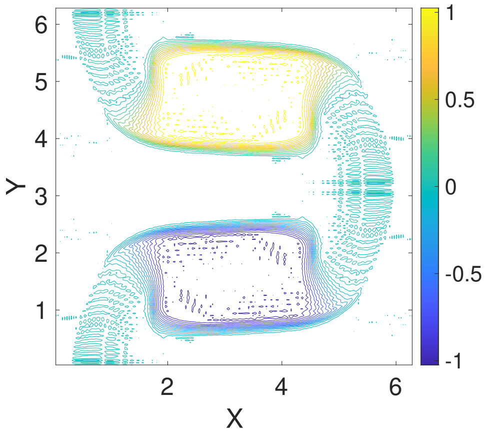


(c) $T = 5$, with no limiter.

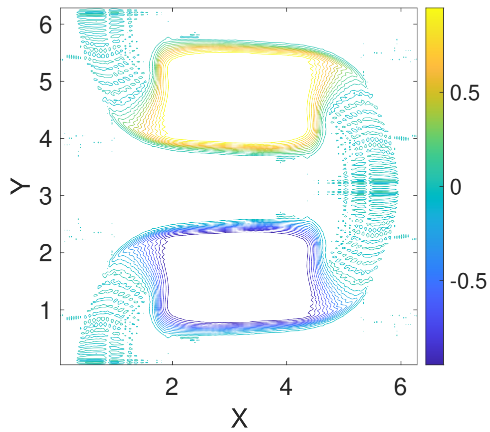


(d) $T = 5$, with BP.

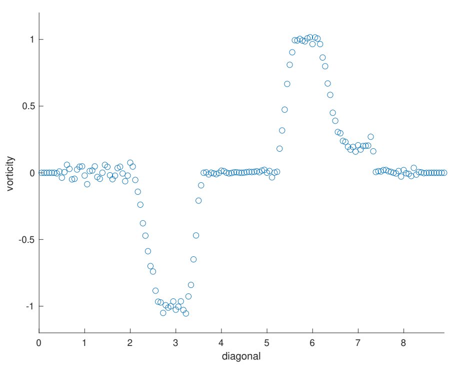
Fig. 4: A fourth-order accurate compact finite difference scheme for the incompressible Euler equation at $T = 5$ on a 160×160 mesh. The time step is $\Delta t = \frac{1}{24 \max |\mathbf{u}_0|} \Delta x$. The second row is the cut along the diagonal of the two-dimensional array.



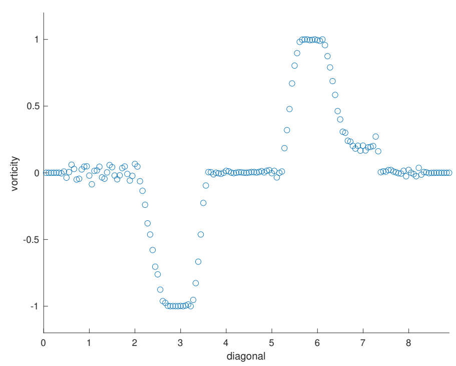
(a) $T = 5$, with TVB1, P=10.



(b) $T = 5$, with BP and TVB1, P=10.

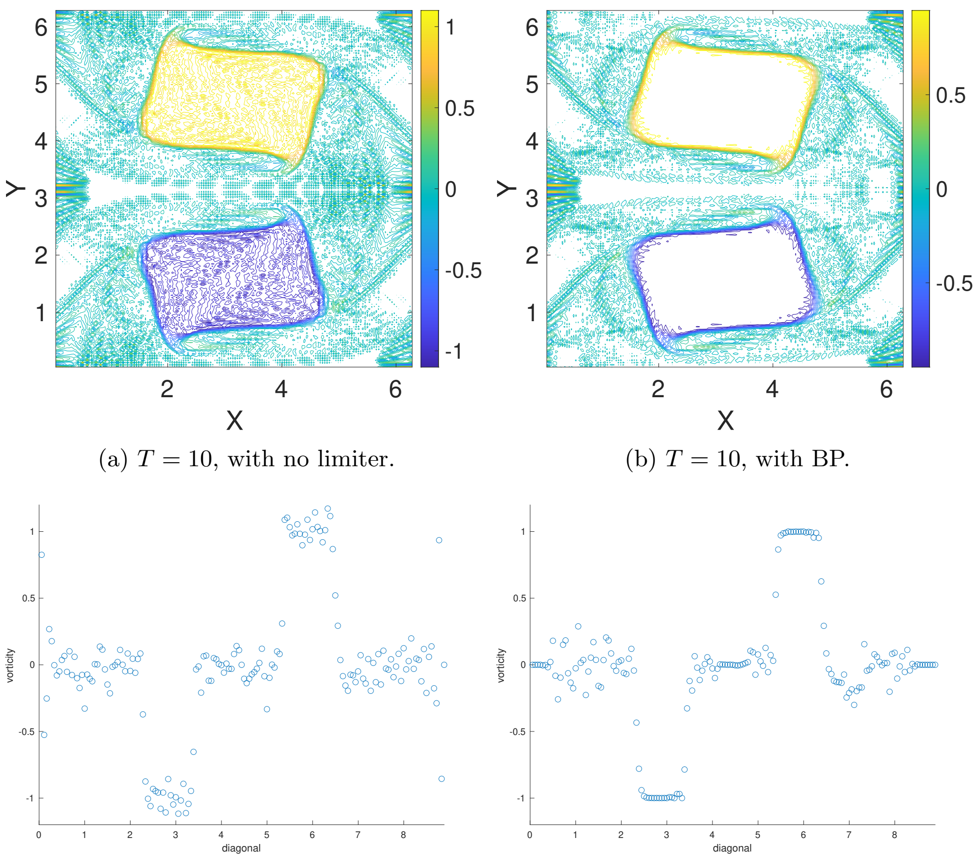


(c) $T = 5$, with TVB1, P $= 10$.



(d) $T = 5$, with BP and TVB1, P $= 10$.

Fig. 5: A fourth-order accurate compact finite difference scheme for the incompressible Euler equation at $T = 5$ on a 160×160 mesh. The time step is $\Delta t = \frac{1}{24 \max |\mathbf{u}_0|} \Delta x$. The second row is the cut along the diagonal of the two-dimensional array.



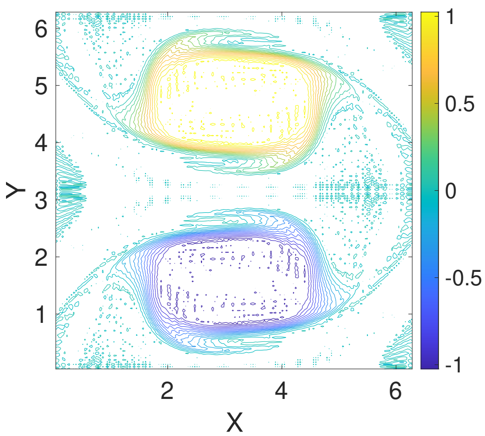
(a) $T = 10$, with no limiter.

(b) $T = 10$, with BP.

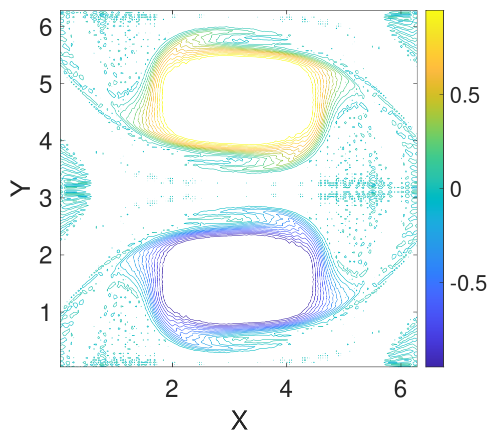
(c) $T = 10$, with no limiter.

(d) $T = 10$, with BP.

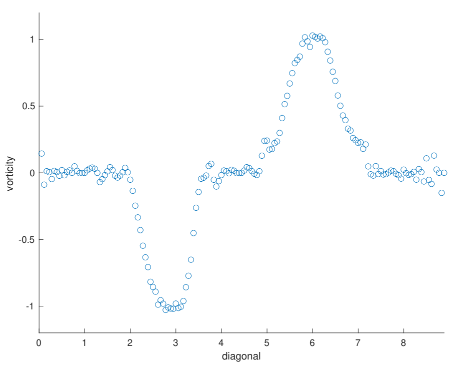
Fig. 6: A fourth-order accurate compact finite difference scheme for the incompressible Euler equation at $T = 5$ on a 160×160 mesh. The time step is $\Delta t = \frac{1}{24 \max |\mathbf{u}_0|} \Delta x$. The second row is the cut along the diagonal of the two-dimensional array.



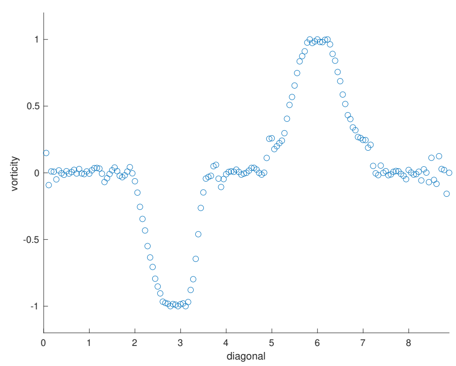
(a) $T = 10$, with TVB1, P=10.



(b) $T = 10$, with BP and TVB1, P=10.



(c) $T = 10$, with TVB1, P = 10.



(d) $T = 10$, with BP and TVB1, P = 10.

Fig. 7: A fourth-order accurate compact finite difference scheme for the incompressible Euler equation at $T = 10$ on a 160×160 mesh. The time step is $\Delta t = \frac{1}{12 \max |\mathbf{u}_0|} \Delta x$. The second row is the cut along the diagonal of the two-dimensional array.

Appendix A: Comparison With The Nine-point Discrete Laplacian. Consider solving the two-dimensional Poisson equations $u_{xx}+u_{yy}=f$ with either homogeneous Dirichlet boundary conditions or periodic boundary conditions on a rectangular domain. Let $\mathbf{u}$ be a $N_x \times N_y$ matrix with entries $u_{i,j}$ denoting the numerical solutions at a uniform grid $(x_i, y_j) = (\frac{i}{Nx}, \frac{j}{Ny})$. Let $\mathbf{f}$ be a $N_x \times N_y$ matrix with entries $f_{i,j} = f(x_i, y_j)$. The fourth order compact finite difference method in Section 2 for $u_{xx}+u_{yy}=f$ can be written as:

$$\frac{1}{\Delta x^2} W_{2x}^{-1} D_{xx}\mathbf{u} + \frac{1}{\Delta y^2} W_{2y}^{-1} D_{yy}\mathbf{u} = f(\mathbf{u}). \tag{6.1}$$

For convenience, we introduce two matrices,

$$U = \begin{pmatrix} u_{i-1,j+1} & u_{i,j+1} & u_{i+1,j+1} \\ u_{i-1,j} & u_{i,j} & u_{i+1,j} \\ u_{i-1,j-1} & u_{i,j-1} & u_{i+1,j-1} \end{pmatrix}, \quad F = \begin{pmatrix} f_{i-1,j+1} & f_{i,j+1} & f_{i+1,j+1} \\ f_{i-1,j} & f_{i,j} & f_{i+1,j} \\ f_{i-1,j-1} & f_{i,j-1} & f_{i+1,j-1} \end{pmatrix}.$$

Notice that the scheme (6.1) is equivalent to

$$\frac{1}{\Delta x^2} W_{2y} D_{xx}\mathbf{u} + \frac{1}{\Delta y^2} W_{2x} D_{yy}\mathbf{u} = W_{2x} W_{2y} f(\mathbf{u}),$$

which can be written as

$$\frac{1}{12\Delta x^2}\begin{pmatrix} 1 & -2 & 1 \\ 10 & -20 & 10 \\ 1 & -2 & 1 \end{pmatrix} : U + \frac{1}{12\Delta y^2}\begin{pmatrix} 1 & 10 & 1 \\ -2 & -20 & -2 \\ 1 & 10 & 1 \end{pmatrix} : U = \frac{1}{144}\begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : F, \tag{6.2}$$

where : denotes the sum of all entrywise products in two matrices of the same size.

In particular, if $\Delta x = \Delta y = h$, the scheme above reduces to

$$\frac{1}{6h^2}\begin{pmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{pmatrix} : U = \frac{1}{144}\begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : F.$$

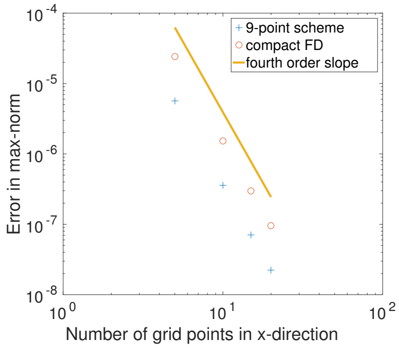
Recall that the classical nine-point discrete Laplacian [4] for the Poisson equation can be written as

$$\frac{1}{12\Delta x^2}\begin{pmatrix} 1 & -2 & 1 \\ 10 & -20 & 10 \\ 1 & -2 & 1 \end{pmatrix} : U + \frac{1}{12\Delta y^2}\begin{pmatrix} 1 & 10 & 1 \\ -2 & -20 & -2 \\ 1 & 10 & 1 \end{pmatrix} : U = \frac{1}{12}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 8 & 1 \\ 0 & 1 & 0 \end{pmatrix} : F, \tag{6.3}$$

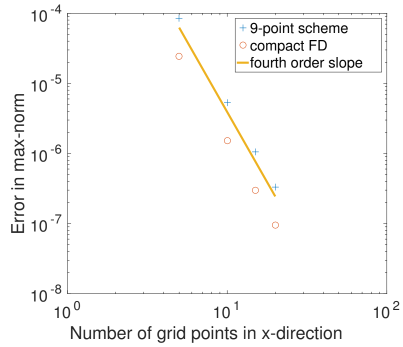
which reduces to the following under the assumption $\Delta x = \Delta y = h$,

$$\frac{1}{6h^2}\begin{pmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{pmatrix} : U = \frac{1}{12}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 8 & 1 \\ 0 & 1 & 0 \end{pmatrix} : F.$$

Both schemes (6.2) and (6.3) are fourth order accurate and they have the same stencil in the left hand side. As to which scheme produces smaller errors, it seems to be problem dependent, see Figure 8. Figure 8 shows the errors of two schemes (6.2) and (6.3) using uniform grids with $\Delta x = \frac{2}{3}\Delta y$ for solving the Poisson equation $u_{xx}+u_{yy} = f$ on a rectangle $[0,1]\times[0,2]$ with Dirichlet boundary conditions. For solution 1, we have $u(x,y) = \sin(\pi x)\sin(\pi y) + 2x$, for solution 2, we have $u(x,y) = \sin(\pi x)\sin(\pi y) + 4x^4y^4$.



(a) Solution 1.



(b) Solution 2.

Fig. 8: Error comparison.

Appendix B: M-matrices And A Discrete Maximum Principle. Consider solving the heat equation $u_t = u_{xx} + u_{yy}$ with a periodic boundary condition. It is well known that a discrete maximum principle is satisfied under certain time step constraints if the spatial discretization is the nine-point discrete Laplacian or the compact scheme (6.1) with backward Euler and Crank-Nicolson time discretizations. For simplicity, we only consider the compact scheme (6.1) and the discussion for the nine-point discrete Laplacian is similar. Assume $\Delta x = \Delta y = h$. For backward Euler, the scheme can be written as

$$\frac{1}{144}\begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : (U^{n+1} - U^n) = \frac{\Delta t}{6h^2}\begin{pmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{pmatrix} : U^{n+1},$$

thus

$$\frac{1}{144}\begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : U^{n+1} - \frac{\Delta t}{6h^2}\begin{pmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{pmatrix} : U^{n+1} = \frac{1}{144}\begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : U^n.$$

Let A and B denote the matrices corresponding to the operator in the left hand side and right hand side above respectively, then the scheme can be written as

$$A\mathbf{u}^{n+1} = B\mathbf{u}^n,$$

and A is a M-Matrix (diagonally dominant, positive diagonal entries and non-positive off diagonal entries) under the following constraint which allows very large time steps:

$$\frac{\Delta t}{h^2} \geq \frac{5}{48}.$$

The inverses of M-Matrices have non-negative entries, e.g., see [6]. Thus A^{-1} has non-negative entries. Moreover, it is straightforward to check that $A\mathbf{e} = \mathbf{e}$ where $\mathbf{e} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}^T$. Thus $A^{-1}\mathbf{e} = \mathbf{e}$, which implies the sum of each row of A^{-1} is 1 thus each row of A^{-1} multiplying any vector V is a convex combination of entries of V. It is also obvious that each entry of B is non-negative and the sum of each row of B is 1. Therefore, $\mathbf{u}^{n+1} = A^{-1}B\mathbf{u}^n$ satisfies a discrete maximum principle:

$$\min_{i,j} u^n_{i,j} \leq u^{n+1}_{i,j} \leq \max_{i,j} u^n_{i,j}.$$

For the second order accurate Crank-Nicolson time discretization, the scheme can be written as

$$\frac{1}{144}\begin{pmatrix}1 & 10 & 1\\ 10 & 100 & 10\\ 1 & 10 & 1\end{pmatrix} : (U^{n+1} - U^n) = \frac{\Delta t}{6h^2}\begin{pmatrix}1 & 4 & 1\\ 4 & -20 & 4\\ 1 & 4 & 1\end{pmatrix} : \frac{U^{n+1} + U^n}{2},$$

thus

$$\left[\frac{1}{144}\begin{pmatrix}1 & 10 & 1\\ 10 & 100 & 10\\ 1 & 10 & 1\end{pmatrix} - \frac{\Delta t}{12h^2}\begin{pmatrix}1 & 4 & 1\\ 4 & -20 & 4\\ 1 & 4 & 1\end{pmatrix}\right] : U^{n+1} =$$

$$\left[\frac{1}{144}\begin{pmatrix}1 & 10 & 1\\ 10 & 100 & 10\\ 1 & 10 & 1\end{pmatrix} + \frac{\Delta t}{12h^2}\begin{pmatrix}1 & 4 & 1\\ 4 & -20 & 4\\ 1 & 4 & 1\end{pmatrix}\right] : U^n.$$

Let the matrix-vector form of the scheme above be $A\mathbf{u}^{n+1} = B\mathbf{u}^n$. Then for A to be an M-Matrix, we only need $\frac{\Delta t}{h^2} \geq \frac{5}{24}$. However, for B to have non-negative entries, we need $\frac{\Delta t}{h^2} \leq \frac{5}{12}$. Thus the Crank-Nicolson method can ensure a discrete maximum principle if the time step satisfies,

$$\frac{5}{24}h^2 \leq \Delta t \leq \frac{5}{12}h^2.$$

Appendix C: Fast Poisson Solvers.

Dirichlet boundary conditions. Consider solving the Poisson equation $u_{xx} + u_{yy} = f(x, y)$ on a rectangular domain $[0, L_x] \times [0, L_y]$ with homogeneous Dirichlet boundary conditions. Assume we use the grid $x_i = i\Delta x$, $i = 0, \cdots, N_x + 1$ with uniform spacing $\Delta x = \frac{L_x}{N_x+1}$ for the x-variable and $y_j = j\Delta y$, $j = 0, \cdots, N_y + 1$ with uniform spacing $\Delta y = \frac{L_y}{N_y+1}$ for y-variable. Let $\mathbf{u}$ be a $N_x \times N_y$ matrix such that its (i, j) entry $u_{i,j}$ is the numerical solution at interior grid points (x_i, y_j). Let $\mathbf{F}$ be a $(N_x + 2) \times (N_y + 2)$ matrix with entries $f(x_i, y_j)$ for $i = 0, \cdots, N_x + 1$ and $j = 0, \cdots, N_y + 1$.

To obtain the matrix representation of the operator in (6.2) and (6.3), we consider two operators:

- Kronecker product of two matrices: if A is $m \times n$ and B is $p \times q$, then $A \otimes B$ is $mp \times nq$ give by

$$A \otimes B = \begin{pmatrix}a_{11}B & \cdots & a_{1n}B\\ \vdots & \vdots & \vdots\\ a_{m1}B & \cdots & a_{mn}B\end{pmatrix}.$$

- For a $m \times n$ matrix X, $\mathrm{vec}(X)$ denotes a column vector of size mn made of the columns of X stacked atop one another from left to right.

The following properties will be used:

1. $(A \otimes B)(C \otimes D) = AC \otimes BD$.

2. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
3. $(B^T \otimes A)\operatorname{vec}(X) = \operatorname{vec}(AXB)$.

We define two tridiagonal square matrices of size $N_x \times N_x$:

$$D_{xx} = \begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}, W_{2x} = \frac{1}{12}\begin{pmatrix} 10 & 1 & & & & \\ 1 & 10 & 1 & & & \\ & 1 & 10 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 10 & 1 \\ & & & & 1 & 10 \end{pmatrix}.$$

Let $\overline{W}_{2x}$ denote a $N_x \times (N_x+2)$ tridiagonal matrix of the following form:

$$\overline{W}_{2x} = \frac{1}{12}\begin{pmatrix} 1 & 10 & 1 & & & \\ & 1 & 10 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 10 & 1 \end{pmatrix}. \tag{6.4}$$

The matrices D_{yy}, W_{2y} and $\overline{W}_{2y}$ are similarly defined.

Then the scheme (6.2) can be written in a matrix-vector form:

$$\frac{1}{\Delta x^2} D_{xx}\mathbf{u}W_{2y}^T + \frac{1}{\Delta y^2} W_{2x}\mathbf{u}D_{yy}^T = \overline{W}_{2x}\mathbf{F}\overline{W}_{2y}^T,$$

or equivalently,

$$\left(W_{2y} \otimes \frac{1}{\Delta x^2} D_{xx} + \frac{1}{\Delta y^2} D_{yy} \otimes W_{2x}\right)\operatorname{vec}(\mathbf{u}) = (\overline{W}_{2x} \otimes \overline{W}_{2y})\operatorname{vec}(\mathbf{F}). \tag{6.5}$$

Let $\mathbf{h}_x = [h_1, h_2, \cdots, h_{N_x}]^T$ with $h_i = \frac{i}{N_x+1}$, and $\sin(m\pi\mathbf{h}_x)$ denote a column vector of size N_x with its i-th entry being $\sin(m\pi h_i)$. Then $\sin(m\pi\mathbf{h}_x)$ are the eigenvectors of D_{xx} and W_{2x} with the associated eigenvalues being $2\cos(\frac{m\pi}{N_x+1}) - 2$ and $\frac{5}{6} + \frac{1}{6}\cos(\frac{m\pi}{N_x+1})$ respectively for $m = 1, \cdots, N_x$. Let

$$S_x = [\sin(\pi\mathbf{h}_x), \sin(2\pi\mathbf{h}_x), \cdots, \sin(N_x\pi\mathbf{h}_x)]$$

be the $N_x \times N_x$ eigenvector matrix, then S_x is a symmetric matrix. Let Λ_{1x} denote a diagonal matrix with diagonal entries $2\cos(\frac{m\pi}{N_x+1}) - 2$ and Λ_{2x} denote a diagonal matrix with diagonal entries $\frac{5}{6} + \frac{1}{6}\cos(\frac{m\pi}{N_x+1})$, then we have $D_{xx} = S_x\Lambda_{1x}S_x^{-1}$ and $W_{2x} = S_x\Lambda_{2x}S_x^{-1}$, thus

$$W_{2y} \otimes D_{xx} = (S_y\Lambda_{2y}S_y^{-1}) \otimes (S_x\Lambda_{1x}S_x^{-1}) = (S_y \otimes S_x)(\Lambda_{2y} \otimes \Lambda_{1x})(S_y^{-1} \otimes S_x^{-1}).$$

The scheme can be written as

$$(S_y \otimes S_x)(\frac{1}{\Delta x^2}\Lambda_{2y} \otimes \Lambda_{1x} + \frac{1}{\Delta y^2}\Lambda_{1y} \otimes \Lambda_{2x})(S_y^{-1} \otimes S_x^{-1})\operatorname{vec}(\mathbf{u}) = (\overline{W}_{2y} \otimes \overline{W}_{2x})\operatorname{vec}(\mathbf{F}).$$

Let Λ be a $N_x \times N_y$ matrix with $\Lambda_{i,j}$ being equal to

$$\frac{1}{3\Delta x^2}\left(\cos(\frac{i\pi}{N_x+1}) - 1\right)\left(\cos(\frac{m\pi}{N_y+1}) + 5\right) + \frac{1}{3\Delta y^2}\left(\cos(\frac{m\pi}{N_x+1}) + 5\right)\left(\cos(\frac{j\pi}{N_y+1}) - 1\right),$$

then vec(Λ) are precisely the diagonal entries of the diagonal matrix $\frac{1}{\Delta x^2}\Lambda_{2y}\otimes\Lambda_{1x}+\frac{1}{\Delta y^2}\Lambda_{1y}\otimes\Lambda_{2x}$, then the scheme above is equivalent to

$$S_x(\Lambda\circ(S_x^{-1}\mathbf{u}S_y^{-1}))S_y=\overline{W}_{2x}\mathbf{F}\overline{W}_{2y}^T,$$

where the symmetry of S matrices is used. The solution is given by

$$\mathbf{u}=S_x\{[S_x^{-1}(\overline{W}_{2x}\mathbf{F}\overline{W}_{2y}^T)S_y^{-1}]./\Lambda\}S_y, \tag{6.6}$$

where ./ denotes the entrywise division for two matrices of the same size.

Since the multiplication of the matrices S and S^{-1} can be implemented by the *Discrete Sine Transform*, (6.6) gives a fast Poisson solver.

For nonhomogeneous Dirichlet boundary conditions, the fourth order accurate compact finite difference scheme can also be written in the form of (6.5):

$$\left(W_{2y}\otimes\frac{1}{\Delta x^2}D_{xx}+\frac{1}{\Delta y^2}D_{yy}\otimes W_{2x}\right)\mathrm{vec}(\mathbf{u})=\mathrm{vec}(\tilde{\mathbf{F}}), \tag{6.7}$$

where $\tilde{\mathbf{F}}$ consists of both $\mathbf{F}$ and the Dirichlet boundary conditions. Thus the scheme can still be efficiently implemented by the *Discrete Sine Transform*.

Periodic boundary conditions. For periodic boundary conditions on a rectangular domain, we should consider the uniform grid $x_i=i\Delta x$, $i=1,\cdots,N_x$ with $\Delta x=\frac{L_x}{N_x}$ and $y_j=j\Delta y$, $j=1,\cdots,N_y$ with uniform spacing $\Delta y=\frac{L_y}{N_y}$, then the fourth order accurate compact finite difference scheme can still be written in the form of (6.5) with the D_{xx}, D_{yy}, W_{2x} and W_{2y} matrices being redefined as circulant matrices:

$$D_{xx}=\begin{pmatrix}-2 & 1 & & & & 1\\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1\\ 1 & & & & 1 & -2\end{pmatrix}, W_{2x}=\frac{1}{12}\begin{pmatrix}10 & 1 & & & & 1\\ 1 & 10 & 1 & & & \\ & 1 & 10 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 10 & 1\\ 1 & & & & 1 & 10\end{pmatrix}.$$

The Discrete Fourier Matrix is the eigenvector matrix for any circulant matrices, and the corresponding eigenvalues are for D_{xx} and W_{2x} are $2\cos(\frac{m2\pi}{N_x})-2$ and $\frac{1}{6}\cos(\frac{m2\pi}{N_x})+\frac{5}{6}$ for $m=0,1,2,\cdots,Nx-1$. The matrix $W_{2y}\otimes\frac{1}{\Delta x^2}D_{xx}+\frac{1}{\Delta y^2}D_{yy}\otimes W_{2x}$ is singular because its first eigenvalue $\Lambda_{1,1}$ is zero. Nonetheless, the scheme can still be implemented by solving (6.6) with Fast Fourier Transform. For the zero eigenvalue, we can simply reset the division by eigenvalue zero to zero. Since the eigenvector for eigenvalue zero is $\mathbf{e}=\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}^T$, and the columns of the Discrete Fourier Matrix are orthogonal to one another, resetting the division by eigenvalue zero to zero simply means that we obtain a numerical solution satisfying $\sum_i\sum_j u_{i,j}=0$. And this is also the least square solution to the singular linear system.

Neumann boundary conditions. For Dirichlet and periodic boundary conditions, we can invert the matrix coefficient matrix in (6.5) using eigenvectors of much smaller matrices W_{2x} and D_{xx} due to the fact that $W_{2x}-\frac{1}{12}D_{xx}$ is the identity matrix Id. Here we discuss how to achieve a fourth order accurate boundary approximation for Neumann boundary conditions by keeping $W_{2x} - \frac{1}{12}D_{xx} = Id$. We first consider a one-dimensional problem with homogeneous Neumann boundary conditions:

$$u''(x) = f(x), x \in [0, L_x],$$
$$u'(0) = u'(L_x) = 0.$$

Assume we use the uniform grid $x_i = i\Delta x$, $i = 0, \cdots, N_x + 1$ with $\Delta x = \frac{L_x}{N_x+1}$. The two boundary point values u_0 and u_{N_x+1} can be expressed in terms of interior point values through boundary conditions. For approximating the boundary conditions, we can apply the fourth order one-sided difference at $x = 0$:

$$u'(0) \approx \frac{-25u(0) + 48u(\Delta x) - 36u(2\Delta x) + 16u(3\Delta x) - 3u(4\Delta x)}{12\Delta x}$$

which implies the finite difference approximation:

$$u_0 = \frac{48u_1 - 36u_2 + 16u_3 - 3u_4}{25}.$$

Define two column vectors:

$$\mathbf{u} = [u_1, u_2, \cdots, u_{N_x}]^T, \quad \mathbf{f} = [f(x_0), f(x_1), \cdots, f(x_{N_x}), f(x_{N_x+1})]^T,$$

then a fourth order accurate compact finite difference scheme can be written as

$$\frac{1}{\Delta x^2}\overline{D}_{xx} I_x \mathbf{u} = \overline{W}_{2x}\mathbf{f},$$

where $\overline{W}_{2x}$ is the same as in (6.4), and $\overline{D}_{xx}$ is a matrix of size $N_x \times (N_x + 2)$ and I_x is a matrix of size $(N_x + 2) \times N_x$:

$$\overline{D}_{xx} = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{pmatrix}, I_x = \begin{pmatrix} \frac{48}{25} & -\frac{36}{25} & \frac{16}{25} & -\frac{3}{25} & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & -\frac{3}{25} & \frac{16}{25} & -\frac{36}{25} & \frac{48}{25} \end{pmatrix}.$$

Now consider solving the Poisson equation $u_{xx}+u_{yy} = f(x, y)$ on a rectangular domain $[0, L_x]\times[0, L_y]$ with homogeneous Neumann boundary conditions. Assume we use the grid $x_i = i\Delta x$, $i = 0, \cdots, N_x + 1$ with $\Delta x = \frac{L_x}{N_x+1}$ and $y_j = j\Delta y$, $j = 0, \cdots, N_y + 1$ with uniform spacing $\Delta y = \frac{L_y}{N_y+1}$. Let $\mathbf{u}$ be a $N_x \times N_y$ matrix such that $u_{i,j}$ is the numerical solution at (x_i, y_j) and $\mathbf{F}$ be a $(N_x + 2) \times (N_y + 2)$ matrix with entries $f(x_i, y_j)$ ($i = 0, \cdots, N_x+1$, $j = 0, \cdots, N_y+1$). Then a fourth order accurate compact finite difference scheme can be written as

$$\frac{1}{\Delta x^2}\overline{D}_{xx} I_x \mathbf{u} I_y^T \overline{W}_{2y}^T + \frac{1}{\Delta y^2}\overline{W}_{2x} I_x \mathbf{u} I_y^T \overline{D}_{yy}^T = \overline{W}_{2x}\mathbf{F}\overline{W}_{2y}^T.$$

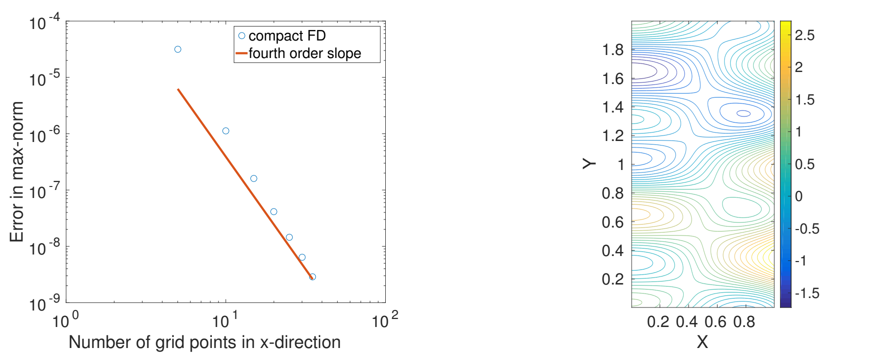
Let $D_{xx} = \overline{D}_{xx} I_x$ and $W_{2x} = \overline{W}_{2x} I_x$, then the scheme can be written as (6.5).

Notice that $W_{2x} - \frac{1}{12}D_{xx} = (\overline{W}_{2x} - \frac{1}{12}\overline{D}_{xx})I_x$ is still the identity matrix thus W_{2x} and D_{xx} still have the same eigenvectors. Let S be the eigenvector matrix and Λ_1 and Λ_2 be diagonal matrices with eigenvalues, then the scheme can still be implemented as (6.6). The eigenvectors S and the eigenvalues can be obtained by computing eigenvalue problems for two small matrices D_{xx} of size $N_x \times N_x$ and D_{yy} of size $N_y \times N_y$. If such a Poisson problem needs to be solved in each time step in a time-dependent problem such as the incompressible flow equations, then this is an efficient Poisson solver because S and Λ can be computed before time evolution without considering eigenvalue problems for any matrix of size $N_x N_y \times N_x N_y$.

For nonhomogeneous Neumann boundary conditions, the point values of u along the boundary should be expressed in terms of interior ones as follows:

1. First obtain the point values except the two cell ends (i.e., corner points of the rectangular domain) for each of the four boundary line segments. For instance, if the left boundary condition is $\frac{\partial u}{\partial x}(0, y) = g(y)$, then we obtain
$$u_{0,j} = \frac{48u_{1,j} - 36u_{2,j} + 16u_{3,j} - 3u_{4,j} + 12\Delta x g(y_j)}{25}, \quad j = 1, \cdots, N_y.$$
2. Second, obtain the approximation at four corners using the point values along the boundary. For instance, if the bottom boundary condition is $\frac{\partial u}{\partial y}(x, 0) = h(x)$, then
$$u_{0,0} = \frac{48u_{1,0} - 36u_{2,0} + 16u_{3,0} - 3u_{4,0} + 12\Delta y h(0)}{25}$$

The scheme can still be written as (6.7) with $\tilde{\mathbf{F}}$ consisting of $\mathbf{F}$ and the nonhomogeneous boundary conditions. Notice that the matrix in (6.7) is singular thus we need to reset the division by eigenvalue zero to zero, which however no longer means that the obtained solution satisfies $\sum_i \sum_j u_{i,j} = 0$ since the eigenvectors are not necessarily orthogonal to one another. See Figure 9 for the accuracy test of the fourth order compact finite difference scheme using uniform grids with $\Delta x = \frac{3}{2}\Delta y$ for solving the Poisson equation $u_{xx} + u_{yy} = f$ on a rectangle $[0, 1] \times [0, 2]$ with nonhomogeneous Neumann boundary conditions. The exact solution is $u(x, y) = \cos(\pi x)\cos(3\pi y) + \sin(\pi y) + x^4$. Since the solutions to Neumann boundary conditions are unique up to any constant, when computing errors, we need to add a constant $\frac{1}{N_x}\frac{1}{N_y}\sum_{i,j}[u(x_i, y_j) - u_{i,j}]$ to each entry of $\mathbf{u}$.



(a) Convergence rate. (b) The contour of the solution.

Fig. 9: Accuracy test for Neumann boundary condition.

Statements and Declarations.

Funding. The research is supported by NSF DMS-1913120.

Conflict of interest. On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

[1] B. Cockburn and C.-W. Shu, *TVB runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework*, Mathematics of Computation, 52 (1989), pp. 411–435.

[2] B. Cockburn and C.-W. Shu, *Nonlinearly stable compact schemes for shock calculations*, SIAM Journal on Numerical Analysis, 31 (1994), pp. 607–627.

[3] S. Gottlieb, D. I. Ketcheson, and C.-W. Shu, *Strong stability preserving Runge-Kutta and multistep time discretizations*, World Scientific, 2011.

[4] R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*, SIAM, 2007.

[5] H. Li, S. Xie, and X. Zhang, *A high order accurate bound-preserving compact finite difference scheme for scalar convection diffusion equations*, SIAM Journal on Numerical Analysis, 56 (2018), pp. 3308–3345.

[6] T. Qin and C.-W. Shu, *Implicit positivity-preserving high-order discontinuous galerkin methods for conservation laws*, SIAM Journal on Scientific Computing, 40 (2018), pp. A81–A107.

[7] C.-W. Shu, *TVB uniformly high-order schemes for conservation laws*, Mathematics of Computation, 49 (1987), pp. 105–121.

[8] X. Zhang, Y. Liu, and C.-W. Shu, *Maximum-principle-satisfying high order finite volume weighted essentially nonoscillatory schemes for convection-diffusion equations*, SIAM Journal on Scientific Computing, 34 (2012), pp. A627–A658.

[9] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, Journal of Computational Physics, 229 (2010), pp. 3091–3120.

[10] Y. Zhang, X. Zhang, and C.-W. Shu, *Maximum-principle-satisfying second order discontinuous Galerkin schemes for convection–diffusion equations on triangular meshes*, Journal of Computational Physics, 234 (2013), pp. 295–316.