Recent Advances in the Study of a Fourth-Order Compact Scheme for the One-Dimensional Biharmonic Equation

D. Fishelov · M. Ben-Artzi · J.-P. Croisille

Abstract It is well-known that non-periodic boundary conditions reduce considerably the overall accuracy of an approximating scheme. In previous papers the present authors have studied a fourth-order compact scheme for the one-dimensional biharmonic equation. It relies on Hermitian interpolation, using functional values and Hermitian derivatives on a three-point stencil. However, the fourth-order accuracy is reduced to a mere first-order near the boundary. In turn this leads to an "almost third-order" accuracy of the approximate solution. By a careful inspection of the matrix elements of the discrete operator, it is shown that the boundary does not affect the approximation, and a full ("optimal") fourth-order convergence is attained. A number of numerical examples corroborate this effect.

Keywords Discrete biharmonic operator · Nonhomogeneous boundary conditions · Fourth-order convergence · Hermite interpolation · Compact schemes

1 Introduction

In this paper we discuss convergence in the sup-norm of a compact approximation to the one-dimensional biharmonic equation. We consider a *boundary value problem*, so that the

D. Fishelov (✉)
Afeka - Tel-Aviv Academic College of Engineering, 218 Bnei-Efraim St., Tel-Aviv 69107, Israel
e-mail: daliaf@post.tau.ac.il

M. Ben-Artzi
Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
e-mail: mbartzi@math.huji.ac.il

J.-P. Croisille
Department of Mathematics, LMAM, UMR 7122, University of Paul Verlaine-Metz, Metz 57045, France
e-mail: croisil@poncelet.univ-metz.fr

values of the function and its derivative are given on the boundary. While the discrete scheme under consideration is fourth-order accurate, the truncation error deteriorates to first-order at the boundary, affecting presumably the rate of convergence of the global approximation. Indeed, in a previous work [10] we have obtained a "suboptimal" convergence rate of almost $O(h^3)$. We show here that, in fact, the convergence rate is optimal, namely $O(h^4)$.

The *Discrete Biharmonic Operator* (henceforth *DBO*) considered here is both compact and of high order accuracy. Such schemes have recently been gaining in popularity, as may be seen in the papers (a very partial list) [2, 5–9, 11, 14–16].

A convergence analysis was performed in [1, 12, 13] in cases where the accuracy of the scheme deteriorates near the boundary. In particular, in [12] and [13] a hyperbolic system of first order and a parabolic problem were analyzed in the case where extra boundary conditions were given in order to "close" the numerical scheme. It was shown in [12, 13] that if the accuracy of the extra boundary conditions is one less than that of the inner scheme, then the overall accuracy of the scheme is determined by the accuracy at inner points. In [1] it was proved for a parabolic equation that if the scheme is of order $O(h^\alpha)$ at inner points and of order $O(h^{\alpha-s})$ near the boundary, then if $s = 0, 1$ the accuracy of the scheme is $O(h^\alpha)$. However, if $s \geq 2$ then the overall accuracy the scheme is $O(h^{\alpha-s+3/2})$. In some sense our approach is an extension of the convergence analysis described in [1, 12, 13]. Here we treat a differential equation of order four. Since our scheme is of fourth order at interior points and of first order at near-boundary points, we have $\alpha = 4$ and $s = 3$. We show that the overall accuracy of the scheme does not deteriorate at all due to the lower-order approximation near the boundary.

Compact high-order schemes for the biharmonic equation can be traced back to Stephenson [16], who proposed such a scheme in two dimensions. The DBO studied here may be viewed as a one-dimensional analog of Stephenson's scheme.

In our approach, the DBO is obtained as a fourth-order derivative of an interpolating polynomial. This polynomial requires not only functional values at neighboring points, but also suitable approximate derivatives. It turns out that in order to maintain accuracy at high order, the approximate derivatives need to be evaluated as fourth-order accurate Hermitian approximations.

Here we investigate in detail the various mathematical features of the discrete approximation:

- The truncation error of the biharmonic operator.
- Optimal, fourth-order, convergence of the discrete solution to the continuous one.

In Sect. 2 we construct a compact fourth-order approximation to the biharmonic operator. In particular, the approximation to the first-order derivative, called the Hermitian derivative, is described. The operators δ_x^2, δ_x as well as the Hermitian derivative are studied in Sect. 3. Their matrix representations are given, as well as the fourth-order accuracy of the Hermitian derivative.

Section 4 is concerned with the study of the truncation error for the approximation of the fourth-order derivative, namely the discrete biharmonic operator. It is shown that it is of fourth order at interior points and of first order at near-boundary points.

Section 5 contains the optimal convergence of the discrete approximation to the exact solution of the one-dimensional biharmonic problem. The error (see Theorem 6) is shown to be of *fourth-order* in the discrete l_h^2 norm.

In Sect. 6 we show numerical results, which validate the fourth-order accuracy of the discrete solution of the one-dimensional biharmonic problem and some of its generalizations.

2 Derivation of Three-Point Compact Operators

We consider here the one-dimensional biharmonic equation on the interval $[a, b]$. For the simplicity of the presentation, we choose homogeneous boundary conditions. The one-dimensional biharmonic equation is

$$\begin{cases} u^{(4)}(x) = f(x), & a < x < b, \\ u(a) = 0, \quad u(b) = 0, \quad u'(a) = 0, \quad u'(b) = 0. \end{cases} \tag{1}$$

We look for a high-order compact approximation to (1). We lay out a uniform grid $a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$. Here $x_i = ih$ for $0 \le i \le N$ and $h = (b-a)/N$.

In what follows, we shall use the notion of grid functions. A *grid function* is a function defined on the discrete grid $\{x_i\}_{i=0}^N$. We denote grid functions with Fraktur letters such as $\mathfrak{u}, \mathfrak{v}$. We have

$$\mathfrak{u} = \big(\mathfrak{u}(x_0), \mathfrak{u}(x_1), \ldots, \mathfrak{u}(x_{N-1}), \mathfrak{u}(x_N)\big). \tag{2}$$

In addition, we denote by $u^* = (u(x_0), u(x_1), \ldots, u(x_{N-1}), u(x_N))$ the grid function, which consists of the values of $u(x)$ at grid points.

We denote by l_h^2 the functional space of grid functions. This space is equipped with a scalar product and an associated norm

$$(\mathfrak{u}, \mathfrak{v})_h = h \sum_{i=0}^{N} \mathfrak{u}(x_i)\mathfrak{v}(x_i), \qquad |\mathfrak{u}|_h = (\mathfrak{u}, \mathfrak{u})_h^{1/2}. \tag{3}$$

The subspace of grid functions, having zero boundary conditions at $x_0 = a$ and $x_N = b$, is denoted by $l_{h,0}^2$. For grid functions $\mathfrak{u}, \mathfrak{v} \in l_{h,0}^2$, we have

$$(\mathfrak{u}, \mathfrak{v})_h = h \sum_{i=1}^{N-1} \mathfrak{u}(x_i)\mathfrak{v}(x_i). \tag{4}$$

We also define the sup norm for a grid function $\mathfrak{u}$

$$|\mathfrak{u}|_\infty = \max_{0 \le i \le N} \big|\mathfrak{u}(x_i)\big|. \tag{5}$$

We define the difference operators δ_x, δ_x^2 on grid functions by

$$\delta_x \mathfrak{u}_i = \frac{\mathfrak{u}_{i+1} - \mathfrak{u}_{i-1}}{2h}, \quad 1 \le i \le N-1, \tag{6}$$

$$\delta_x^2 \mathfrak{u}_i = \frac{\mathfrak{u}_{i+1} - 2\mathfrak{u}_i + \mathfrak{u}_{i-1}}{h^2}, \quad 1 \le i \le N-1. \tag{7}$$

In these definitions the boundary values $\mathfrak{u}_0, \mathfrak{u}_N$ are assumed to be known.

Suppose that we are given data $u_{i-1}^*, u_i^*, u_{i+1}^*$ at the grid points x_{i-1}, x_i, x_{i+1}. In addition, we are given some approximations $u_{x,i-1}^*, u_{x,i+1}^*$ for $u'(x_{i-1}), u'(x_{i+1})$.

We seek a polynomial of degree 4

$$p(x) = u_i^* + a_1(x - x_i) + a_2(x - x_i)^2 + a_3(x - x_i)^3 + a_4(x - x_i)^4, \tag{8}$$

which interpolates the data $u_{i-1}^*, u_i^*, u_{i+1}^*, u_{x,i-1}^*, u_{x,i+1}^*$.

The coefficients a_1, a_2, a_3, a_4 of the polynomial are

$$
\begin{cases}
a_1 = \frac{3}{4h}(u^*_{i+1} - u^*_{i-1}) - (\frac{1}{4}u^*_{x,i+1} + \frac{1}{4}u^*_{x,i-1}), \\
a_2 = \frac{1}{h^2}(u^*_{i+1} + u^*_{i-1} - 2u^*_i) - \frac{1}{4h}(u^*_{x,i+1} - u^*_{x,i-1}) = \delta_x^2 u^*_i - \frac{1}{2}(\delta_x u^*_x)_i, \\
a_3 = -\frac{1}{4h^3}(u^*_{i+1} - u^*_{i-1}) + \frac{1}{4h^2}(u^*_{x,i+1} + u^*_{x,i-1}), \\
a_4 = -\frac{1}{2h^4}(u^*_{i+1} + u^*_{i-1} - 2u^*_i) + \frac{1}{4h^3}(u^*_{x,i+1} - u^*_{x,i-1}) = \frac{1}{2h^2}((\delta_x u^*_x)_i - \delta_x^2 u^*_i).
\end{cases}
\tag{9}
$$

The coefficients above require the data u^*_i and $u^*_{x,i}$. In the case where only the values of u^*_i are given, then $\{u^*_{x,i}\}_{i=1}^{N-1}$ have to be evaluated in terms of $\{u^*_i\}_{i=0}^N$. Looking at the first equation in (9), we see that a natural candidate for $u^*_{x,i}$ is

$$
u^*_{x,i} = a_1.
$$

This yields

$$
u^*_{x,i} = \frac{3}{4h}\left(u^*_{i+1} - u^*_{i-1}\right) - \left(\frac{1}{4}u^*_{x,i+1} + \frac{1}{4}u^*_{x,i-1}\right),
$$

or equivalently

$$
\frac{1}{6}u^*_{x,i} + \frac{2}{3}u^*_{x,i} + \frac{1}{6}u^*_{x,i+1} = \delta_x u^*_i.
\tag{10}
$$

This is by definition the Hermitian derivative. If we introduce the three-point operator σ_x on grid functions by

$$
\sigma_x \mathfrak{v}_i = \frac{1}{6}\mathfrak{v}_{i-1} + \frac{2}{3}\mathfrak{v}_i + \frac{1}{6}\mathfrak{v}_{i+1}, \quad 1 \leq i \leq N-1,
\tag{11}
$$

can rewrite (10) as

$$
\sigma_x u^*_{x,i} = \delta_x u^*_i, \quad 1 \leq i \leq N-1.
\tag{12}
$$

Observe that a knowledge of

$$
u^*_{x,0} = u'_0, \qquad u^*_{x,N} = u'_N,
\tag{13}
$$

is needed in order to solve (12). In addition, we will invoke the following relation

$$
\sigma_x = I + \frac{h^2}{6}\delta_x^2.
\tag{14}
$$

Natural approximations to $u''(x_i)$, $u'''(x_i)$, $u''''(x_i)$ are $2a_2, 6a_3, 24a_4$, respectively (see (9)). We use the notation $\tilde{\delta}_x^2 u^*, \delta_x^3 u^*, \delta_x^4 u^*$ for the following operators.

$$
\begin{cases}
\tilde{\delta}_x^2 u^*_i = 2a_2 = 2\delta_x^2 u^*_i - (\delta_x u^*_x)_i, \\
\delta_x^3 u^*_i = 6a_3 = \frac{1}{h^2}(u^*_{x,i+1} + u^*_{x,i-1} - 2u^*_{x,i}) = (\delta_x^2 u^*_x)_i, \\
\delta_x^4 u^*_i = 24a_4 = \frac{12}{h^2}((\delta_x u^*_x)_i - \delta_x^2 u^*_i).
\end{cases}
\tag{15}
$$

This suggests that $\delta_x^4 u_i^*$ is an approximation to the fourth-order derivative of u at x_i, namely,

$$\delta_x^4 u_i^* = \frac{12}{h^2}\left((\delta_x u_x^*)_i - \delta_x^2 u_i^*\right). \tag{16}$$

This approximation, called the discrete biharmonic approximation, is the one-dimensional analog of the Stephenson's scheme [16]. Note that, in the non-periodic setting, boundary values of u_x should be given in order to compute δ_x^4 at near boundary points x_1, x_{N-1}.

3 The Operators δ_x^2, δ_x and the Hermitian Derivative

3.1 Matrix Representation of the Hermitian Derivative

Let us provide now some matrix representations of the operators appearing in the Hermitian gradient. Let $U \in \mathbb{R}^{N-1}$ be the vector corresponding to the grid function $\mathfrak{u} \in l_{h,0}^2$,

$$U = [\mathfrak{u}_1, \ldots, \mathfrak{u}_{N-1}]^T. \tag{17}$$

The vector corresponding to the grid function $\delta_x \mathfrak{u}$ is

$$\frac{1}{2h} KU, \tag{18}$$

where the matrix $K = (K_{i,m})_{1 \leq i,m \leq N-1}$ is the skew-symmetric matrix

$$K = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ -1 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & \ldots & -1 & 0 & 1 \\ 0 & \ldots & 0 & -1 & 0 \end{bmatrix}. \tag{19}$$

The matrix which corresponds to σ_x is $P/6$, where P is the positive definite $(N-1) \times (N-1)$ matrix

$$P = \begin{bmatrix} 4 & 1 & 0 & \ldots & 0 \\ 1 & 4 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & \ldots & 1 & 4 & 1 \\ 0 & \ldots & 0 & 1 & 4 \end{bmatrix}. \tag{20}$$

Thus, (12) can be written as

$$\frac{1}{6} P U_x = \frac{1}{2h} KU, \tag{21}$$

where U, U_x are the vectors corresponding to $\mathfrak{u}, \mathfrak{u}_x$, respectively.

In the sequel, we shall also need the matrix representation of δ_x^2. The matrix T, which corresponds to $-h^2 \delta_x^2$, is the $(N-1) \times (N-1)$ symmetric matrix

$$
T = \begin{bmatrix}
2 & -1 & 0 & \ldots & 0 \\
-1 & 2 & -1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ldots & \vdots \\
0 & \ldots & -1 & 2 & -1 \\
0 & \ldots & 0 & -1 & 2
\end{bmatrix}. \tag{22}
$$

The matrix P is related to T by

$$
P = 6I - T. \tag{23}
$$

Therefore, the matrix which corresponds to the operator σ_x (restricted to $l_{h,0}^2$) is $P/6 = I - T/6$.

3.2 The Eigenvalues and the Eigenvectors of δ_x^2

To simplify the notations, we assume from now on that $[a, b] = [0, 1]$, thus $Nh = 1$.

In order to prove the fourth-order accuracy of the scheme, we shall need the eigenvalues and eigenvectors corresponding to δ_x^2, and thus to the matrix T. The eigenvalues of T are

$$
\lambda_j = 4 \sin^2 \left(\frac{j\pi}{2N} \right), \quad j = 1, \ldots, N-1 \tag{24}
$$

and the corresponding normalized eigenvectors are $Z^k = (Z_{1k}, \ldots, Z_{N-1,k})^T$ (with respect to the Euclidean norm in $\mathbb{R}^{N-1}$), where

$$
Z_{jk} = \left(\frac{2}{N} \right)^{1/2} \sin \frac{kj\pi}{N}, \quad 1 \leq k, j \leq N-1. \tag{25}
$$

We denote the column vectors as $Z^k \in \mathbb{R}^{N-1}$ and the row vectors as $Z_j \in \mathbb{R}^{N-1}$.

The matrix $Z = (Z_{jk})_{1 \leq j,k \leq N-1} \in \mathbb{M}_{N-1}(\mathbb{R})$ is an orthogonal positive-definite matrix. Thus,

$$
Z^2 = ZZ^T = I_{N-1}. \tag{26}
$$

It follows that the matrix T satisfies

$$
T = Z\Lambda Z^T, \tag{27}
$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{N-1})$. The normalized vectors (with respect to $(|\cdot|_h)$), which diagonalize the operator $-\delta_x^2$, are the grid functions z^k, which are defined by

$$
z_{jk} = Z_{jk}/h^{1/2}. \tag{28}
$$

Equivalently, they may be written as (noting that $Nh = 1$)

$$
z_{jk} = \sqrt{2} \sin \frac{kj\pi}{N}, \quad 1 \leq k, j \leq N-1. \tag{29}
$$

We have

$$
\begin{cases}
z_{jk} = \sqrt{2}\sin(j\frac{k\pi}{N}), & j = 1, \ldots, N-1, k = 1, \ldots, N-1 \\
z_{0k} = 0, \quad z_{Nk} = 0, \\
-\delta_x^2 z^k = \tilde{\lambda}_k z^k, & \tilde{\lambda}_k = \frac{4}{h^2}\sin^2(\frac{k\pi}{2N}), k = 1, \ldots, N-1.
\end{cases}
\tag{30}
$$

3.3 The Accuracy of the Hermitian Derivative

Now we state a lemma, proved in [3], which indicates the fourth-order accuracy of the Hermitian derivative.

Lemma 1 *Suppose that $u(x)$ is a smooth function on $[a, b]$ and let $\mathfrak{u} = u^*$. Then, the Hermitian derivative $\mathfrak{u}_x$, as obtained from the values $u(x_i)$, $0 \leq i \leq N$ by*

$$
(\sigma_x \mathfrak{u}_x)_i = (\delta_x u^*)_i, \quad 1 \leq i \leq N-1
\tag{31}
$$

and

$$
(\mathfrak{u}_x)_0 = (u')^*(x_0), \qquad (\mathfrak{u}_x)_N = (u')^*(x_N),
\tag{32}
$$

has a truncation error $\mathfrak{u}_x - (u')^*$ of order $O(h^4)$. More precisely,

$$
\left| \mathfrak{u}_x - (u')^* \right|_\infty \leq Ch^4 \left\| u^{(5)} \right\|_{L^\infty}.
\tag{33}
$$

4 The DBO and Its Truncation Error

As mentioned in Sect. 2 the approximation $\delta_x^4 u_i^*$, suggested in (16), may serve as approximation to $u^{(4)}(x_i)$. We refer to δ_x^4 as the *discrete biharmonic operator (DBO)*. We define

Definition 2 (Discrete biharmonic operator (DBO)) Let $\mathfrak{u} \in l_h^2$ be a given grid function. The *discrete biharmonic operator* is defined by

$$
\delta_x^4 \mathfrak{u}_i = \frac{12}{h^2} (\delta_x \mathfrak{u}_{x,i} - \delta_x^2 \mathfrak{u}_i), \quad 1 \leq i \leq N-1.
\tag{34}
$$

Here $\mathfrak{u}_x$ is the Hermitian derivative of $\mathfrak{u}$ satisfying (12) with given boundary values $\mathfrak{u}_{x,0}$ and $\mathfrak{u}_{x,N}$.

Using (16) and (10), the solution of (1) may be approximated by the scheme

$$
\begin{cases}
\text{(a)} & \delta_x^4 \mathfrak{u}_i = f(x_i) \quad 1 \leq i \leq N-1, \\
\text{(b)} & \frac{1}{6}\mathfrak{u}_{x,i-1} + \frac{2}{3}\mathfrak{u}_{x,i} + \frac{1}{6}\mathfrak{u}_{x,i+1} = \delta_x \mathfrak{u}_i, \quad 1 \leq i \leq N-1, \\
\text{(c)} & \mathfrak{u}_0 = 0, \quad \mathfrak{u}_N = 0, \quad \mathfrak{u}_{x,0} = 0, \quad \mathfrak{u}_{x,N} = 0.
\end{cases}
\tag{35}
$$

The scheme in (35) is the one-dimensional restriction of the scheme proposed by Stephenson in [16]. In the sequel, this scheme is referred to as the *one-dimensional Stephenson Scheme to the biharmonic equation*. Note that it approximates both u and u' at the grid points.

We first study in detail its truncation error. Let $u(x)$ be a smooth function on $[a, b]$, such that $u(a) = u(b) = 0$, $u'(a) = u'(b) = 0$. We denote by u^* its related grid function.

We begin by considering the action of $\sigma_x \delta_x^4$ at the interior point x_i.

$$\sigma_x \delta_x^4 u_i^* = \frac{1}{6} \delta_x^4 u_{i-1}^* + \frac{2}{3} \delta_x^4 u_i^* + \frac{1}{6} \delta_x^4 u_{i+1}^*, \quad 2 \le i \le N-2, \tag{36}$$

where σ_x is the Simpson operator defined in (11). The right-hand side can be expressed as

$$\frac{12}{h^2} \left(\left(\frac{1}{6} \delta_x u_{x,i-1}^* + \frac{2}{3} \delta_x u_{x,i}^* + \frac{1}{6} \delta_x u_{x,i+1}^* \right) - \left(\frac{1}{6} \delta_x^2 u_{i-1}^* + \frac{2}{3} \delta_x^2 u_i^* + \frac{1}{6} \delta_x^2 u_{i+1}^* \right) \right). \tag{37}$$

Using the definition of u_x^*, the first term in this expression is

$$\frac{1}{6} \delta_x u_{x,i-1}^* + \frac{2}{3} \delta_x u_{x,i}^* + \frac{1}{6} \delta_x u_{x,i+1}^* = \sigma_x \delta_x u_{x,i}^* = \delta_x \sigma_x u_{x,i}^* = \delta_x \delta_x u_i^*$$

$$= \frac{1}{4h^2} (u_{i+2}^* - 2u_i^* + u_{i-2}^*), \quad 2 \le i \le N-2. \tag{38}$$

The second term in (37) may be written as

$$\frac{1}{6} \delta_x^2 u_{i-1}^* + \frac{2}{3} \delta_x^2 u_i^* + \frac{1}{6} \delta_x^2 u_{i+1}^*$$

$$= \frac{1}{6h^2} (u_{i-2}^* + 2u_{i-1}^* - 6u_i^* + 2u_{i+1}^* + u_{i+2}^*), \quad 2 \le i \le N-2. \tag{39}$$

Therefore, inserting (38)–(39) in (36), we have

$$\sigma_x \delta_x^4 u_i^* = \frac{1}{h^4} \left(u_{i-2}^* - 4u_{i-1}^* + 6u_i^* - 4u_{i+1}^* + u_{i+2}^* \right) = \delta_x^2 \delta_x^2 u_i^*, \quad 2 \le i \le N-2. \tag{40}$$

Thus, in the absence of boundaries, there is a strong connection between δ_x^4 and $(\delta_x^2)^2$. Explicit estimates for $\sigma_x \delta_x^4 u_i^*$ at near boundary points x_1, x_{N-1} are given below (see (42)). It results from this representation that $\sigma_x \delta_x^4$ actually coincides with the operator $(\delta_x^2)^2$ at points $x_i, 2 \le i \le N-2$. Only at near boundary points, $i=1, i=N-1$, we have a "numerical boundary layer" effect. Let us now investigate the accuracy of the DBO.

The following proposition deals with the truncation error of the DBO.

Proposition 3 *Suppose that $u(x)$ is a smooth function on $[a, b]$. Assume, in addition, that $u(a) = u(b) = 0$, $u'(a) = u'(b) = 0$. Let $u_i^* = u(x_i)$, $(u^{(4)})^*(x_i) = u^{(4)}(x_i)$ be the grid functions corresponding, respectively, to $u, u^{(4)}$. Then the DBO δ_x^4 satisfies the following accuracy properties*:

•

$$\left| \sigma_x \delta_x^4 u_i^* - \sigma_x (u^{(4)})^*(x_i) \right| \le Ch^4 \|u^{(8)}\|_{L^\infty}, \quad 2 \le i \le N-2. \tag{41}$$

• *At near boundary points $i=1$ and $i=N-1$, the fourth order accuracy of (41) drops to first order,*

$$\left| \sigma_x \delta_x^4 u_1^* - \sigma_x (u^{(4)})^*(x_1) \right| \le Ch \|u^{(5)}\|_{L^\infty}, \tag{42}$$

with a similar estimate for $i = N-1$.

• *The error in the energy norm is given by*

$$\left| \delta_x^4 u^* - (u^{(4)})^* \right|_h \le Ch^{3/2} \left(\|u^{(5)}\|_{L^\infty} + \|u^{(8)}\|_{L^\infty} \right). \tag{43}$$

In the above estimates C is a generic constant, that does not depend on u.

Proof According to (40), we have

$$\sigma_x \delta_x^4 u_i^* = \left(\delta_x^2\right)^2 u_i^*, \quad i = 2, \ldots, N-2. \tag{44}$$

We now expand $(\delta_x^2)^2$ in Taylor series. We have

$$\delta_x^2 u_i = u''(x_i) + \frac{h^2}{12} u^{(4)}(x_i) + \frac{h^4}{360} u^{(6)}(\xi_i), \quad 1 \leq i \leq N-1, \tag{45}$$

where $\xi_i \in (x_i - h, x_i + h)$.

We note that for any smooth function $v(x)$, $x \in [a, b]$, we have for $a + h < x < b - h$

$$\frac{v(x+h) - 2v(x) + v(x-h)}{h^2} = v''(\eta), \tag{46}$$

where $\eta \in (x - h, x + h)$.

Applying δ_x^2 to (45) at the interior points x_i, $2 \leq i \leq N-2$, we obtain

$$\sigma_x \delta_x^4 u_i^* = (\delta_x^2)^2 u_i^*$$
$$= u^{(4)}(x_i) + \frac{h^2}{6} u^{(6)}(x_i) + p_i, \qquad |p_i| \leq C_1 h^4 \left\| u^{(8)} \right\|_{L^\infty}, \quad 2 \leq i \leq N-2. \tag{47}$$

On the other hand, $\sigma_x (u^{(4)})^*(x_i)$ may be expanded around x_i, $2 \leq i \leq N-2$, as follows.

$$\sigma_x \left(u^{(4)} \right)^*(x_i) = \left(I + \frac{h^2}{6} \delta_x^2 \right) \left(u^{(4)} \right)^*(x_i)$$
$$= u^{(4)}(x_i) + \frac{h^2}{6} u^{(6)}(x_i) + q_i, \qquad |q_i| \leq C_2 h^4 \left\| u^{(8)} \right\|_{L^\infty}, \quad 2 \leq i \leq N-2.$$

Therefore, subtracting this equation from (47), we obtain the estimate (41).

Consider now the near boundary point x_1. We set $\delta_x^4 u_0^* = (u^{(4)})^*(x_0)$ and then, using the definition of $\sigma_x \delta_x^4$, we have

$$\sigma_x \delta_x^4 u_1^* - \sigma_x \left(u^{(4)} \right)^*(x_1) = \left(\frac{2}{3} \delta_x^4 u_1^* + \frac{1}{6} \delta_x^4 u_2^* \right) - \left(\frac{2}{3} \left(u^{(4)} \right)(x_1) + \frac{1}{6} \left(u^{(4)} \right)(x_2) \right)$$
$$= \frac{2}{3} \left(\delta_x^4 u_1^* - \left(u^{(4)} \right)(x_1) \right) + \frac{1}{6} \left(\delta_x^4 u_2^* - \left(u^{(4)} \right)(x_2) \right). \tag{48}$$

First, we consider the terms evaluated at x_1. Recall that

$$\delta_x^4 u_1^* = \frac{12}{h^2} \left(\left(\delta_x u_x^* \right)_1 - \delta_x^2 u_1^* \right), \tag{49}$$

where u_x^* is the Hermitian derivative of u^*. Using the boundary values $u_0^* = u_{x,0}^* = 0$, we have, in view of (33),

$$\left(\delta_x u_x^* \right)_1 = \frac{u_{x,2}^*}{2h} = u''(x_1) + \frac{h^2}{6} u^{(4)}(x_1) + r_1, \qquad |r_1| \leq C h^3 \left\| u^{(5)} \right\|_{L^\infty}, \tag{50}$$

and

$$\delta_x^2 u_1^* = u''(x_1) + \frac{h^2}{12} u^{(4)}(x_1) + r_2, \qquad |r_2| \leq Ch^3 \|u^{(5)}\|_{L^\infty}. \tag{51}$$

Inserting the estimates (50), (51) in (49), we obtain

$$\delta_x^4 u_1^* = u^{(4)}(x_1) + r_3, \qquad |r_3| \leq Ch \|u^{(5)}\|_{L^\infty}. \tag{52}$$

Next, for x_2 we have

$$\delta_x^4 u_2^* = \frac{12}{h^2} \left((\delta_x u_x^*)_2 - \delta_x^2 u_2^* \right). \tag{53}$$

Expanding on the term $(\delta_x u_x^*)_2$ and using again (33), we have

$$\left(\delta_x u_x^* \right)_2 = \frac{u_{x,3}^* - u_{x,1}^*}{2h} = u''(x_2) + \frac{h^2}{6} u^{(4)}(x_2) + s_1, \qquad |s_1| \leq Ch^3 \|u^{(5)}\|_{L^\infty}. \tag{54}$$

For the second term $\delta_x^2 u_2^*$ we have, as in (51),

$$\delta_x^2 u_2^* = u''(x_2) + \frac{h^2}{12} u^{(4)}(x_2) + s_2, \qquad |s_2| \leq Ch^3 \|u^{(5)}\|_{L^\infty}. \tag{55}$$

Inserting the estimates (54), (55) in (53), we obtain, as in (52),

$$\delta_x^4 u_2^* = u^{(4)}(x_2) + s_3, \qquad |s_3| \leq Ch \|u^{(5)}\|_{L^\infty}. \tag{56}$$

Combining the estimates for r_3 and s_3 and inserting them in (48), we obtain

$$\left| \sigma_x \delta_x^4 u_1^* - \sigma_x \left(u^{(4)} \right)_1^* \right| \leq Ch \|u^{(5)}\|_{L^\infty}, \tag{57}$$

which proves (42).

(iii) Let $\mathfrak{t}_i = \delta_x^4 u_i^* - (u^{(4)})_i^*$ be the truncation error for the fourth-order derivative approximation. We have

$$\sigma_x \mathfrak{t} = \mathfrak{v}, \tag{58}$$

where $\mathfrak{v} \in l_{h,0}^2$ satisfies the estimates established in the previous parts of the lemma

$$|\mathfrak{v}_1|, |\mathfrak{v}_{N-1}| \leq Ch \|u^{(5)}\|_{L^\infty}, \qquad |\mathfrak{v}_i| \leq Ch^4 \|u^{(8)}\|_{L^\infty}, \quad 2 \leq i \leq N - 2. \tag{59}$$

The representative matrix of σ_x restricted to $l_{h,0}^2$ is $P/6 = I - T/6$. The eigenvalues of $P/6$ are

$$1 - \frac{2}{3} \mu_k = 1 - \frac{2}{3} \sin^2 \left(\frac{k\pi}{2N} \right). \tag{60}$$

The matrix norm of its inverse is

$$\left| (P/6)^{-1} \right|_2 = \max_{k=1,\ldots,N-1} \left| \frac{1}{1 - \frac{2}{3} \sin^2 (\frac{k\pi}{2N})} \right| \leq 3. \tag{61}$$

From (58), (59) and (61), we obtain,

$$|\mathfrak{t}|_h \leq C |\mathfrak{v}|_h. \tag{62}$$

Finally, since

$$|\mathfrak{v}|_h^2 \le Ch\left(2h^2 + \sum_{i=2}^{N-2} h^8\right)\left(\left\|u^{(5)}\right\|_{L^\infty}^2 + \left\|u^{(8)}\right\|_{L^\infty}^2\right) \le Ch^3\left(\left\|u^{(5)}\right\|_{L^\infty}^2 + \left\|u^{(8)}\right\|_{L^\infty}^2\right), \quad (63)$$

we get (43).   □

5 Optimal Rate of Convergence of the One-Dimensional Stephenson Scheme

In order to prove the fourth-order convergence of the scheme, we invoke the matrix representation for the discrete biharmonic operator.

5.1 Matrix Representation of the DBO

We have shown in [4] that the matrix form of the DBO (see Definition 2) is obtained from the matrix form of operators $\mathfrak{u} \mapsto \mathfrak{u}_x$ (see (21)), $\mathfrak{u} \mapsto \delta_x \mathfrak{u}$ (see (18)) and $\mathfrak{u} \mapsto \delta_x^2 \mathfrak{u}$ (see (22)). Let $U \in \mathbb{R}^{N-1}$ be the vector corresponding to the grid function $\mathfrak{u} \in l_{h,0}^2$.

Therefore, the matrix representation of $\mathfrak{u} \mapsto \delta_x^4 \mathfrak{u}$ is

$$SU = \frac{12}{h^2}\left[\frac{3}{2h^2}KP^{-1}K + \frac{1}{h^2}T\right]U = \frac{6}{h^4}\left[3KP^{-1}K + 2T\right]U. \quad (64)$$

The fact that we deal with a boundary value problem, rather than a periodic one, means that $PK - KP \ne 0$. However, the commutator is non-zero only at near-boundary points. Using the precise form of this commutator, we get the following proposition.

Proposition 4

(i) *The operator $\sigma_x \delta_x^4$ has the matrix form*

$$PS = \frac{6}{h^4}T^2 + \frac{6}{h^4}\left[e_1\left(e_1 + KP^{-1}e_1\right)^T + e_{N-1}\left(e_{N-1} - KP^{-1}e_{N-1}\right)^T\right], \quad (65)$$

where

$$e_1 = (1, 0, \ldots, 0)^T, \qquad e_{N-1} = (0, \ldots, 0, 1)^T. \quad (66)$$

(ii) *The symmetric positive definite operator δ_x^4 (see (64)) has the matrix form*

$$S = \frac{6}{h^4}P^{-1}T^2 + \frac{36}{h^4}\left(V_1 V_1^T + V_2 V_2^T\right), \quad (67)$$

where the vectors V_1, V_2 are

$$\begin{cases} V_1 = (\alpha - \beta)^{1/2} P^{-1}\left(\frac{\sqrt{2}}{2}e_1 - \frac{\sqrt{2}}{2}e_{N-1}\right) \\ V_2 = (\alpha + \beta)^{1/2} P^{-1}\left(\frac{\sqrt{2}}{2}e_1 + \frac{\sqrt{2}}{2}e_{N-1}\right). \end{cases} \quad (68)$$

The constants α, β are

$$\begin{cases} \alpha = 2(2 - e_1^T P^{-1}e_1) \\ \beta = 2e_{N-1}^T P^{-1}e_1. \end{cases} \quad (69)$$

Remark 5 In view of the positivity of P^{-1}, we have $0 \le e_1^T P^{-1} e_1 \le 1/2$ and $|e_{N-1}^T P^{-1} e_1| \le 1/2$, so that $3 \le \alpha \le 4$ and $|\beta| \le 1$. Thus, $(\alpha \pm \beta)^{1/2}$ are well defined.

5.2 Error Estimate for the One-Dimensional Stephenson Scheme

In [3] we carried out an error analysis based on the coercivity of δ_x^4. The analysis presented there was based on an energy (l^2) method and led to a "sub-optimal" convergence rate of $h^{\frac{3}{2}}$. In [10] we have improved this result by showing that the convergence rate is almost three (the error is bounded by $Ch^3 \log(|h|)$. Here we prove the optimal (fourth-order) convergence of the scheme.

In order to obtain an optimal convergence rate, we use the matrix structure of δ_x^4 given in (67). Let u be the exact solution of (1) and let $\mathfrak{u}$ be its approximation by the Stephenson scheme (35). Let u^* be the grid function corresponding to u. We consider the error between the approximated solution $\mathfrak{u}$ and the collocated exact solution u^*,

$$\mathfrak{e} = \mathfrak{u} - u^*.$$

The grid function u^* satisfies

$$\delta_x^4 u_i^* = f^*(x_i) + \mathfrak{r}_i, \quad 1 \le i \le N - 1, \tag{70}$$

where $\mathfrak{r}$ is by definition the truncation error. We later refer to Proposition 3 for estimates on $\mathfrak{r}$.

The error $\mathfrak{e} = \mathfrak{u} - u^*$ satisfies

$$\delta_x^4 \mathfrak{e}_i = -\mathfrak{r}_i, \quad 1 \le i \le N - 1,$$
$$\mathfrak{e}_0 = 0, \qquad \mathfrak{e}_N = 0, \qquad \mathfrak{e}_{x,0} = 0, \qquad \mathfrak{e}_{x,N} = 0. \tag{71}$$

We prove the following error estimate.

Theorem 6 *Let u be the exact solution of* (1) *and assume that u has continuous derivatives up to order eight on* $[a, b]$. *Let* $\mathfrak{u}$ *be the approximation to u, given by the Stephenson scheme* (35). *Let* u^* *be the grid function corresponding to u. Then, the error* $\mathfrak{e} = \mathfrak{u} - u^*$ *satisfies*

$$|\mathfrak{e}|_h \le Ch^4, \tag{72}$$

where C depends only on f.

Proof Let $U, U^* \in \mathbb{R}^{N-1}$ be the vectors corresponding to $\mathfrak{u}, u^*$, respectively, and let F be the vector corresponding to f^*. We denote by $E = U - U^*$ and R the vectors corresponding to $\mathfrak{e} = \mathfrak{u} - u^*$ and $\mathfrak{r}$, respectively.

Using the matrix representation (67), we can write (1) and (70) in the form

$$SU = F, \tag{73}$$

and

$$SU^* = F + R. \tag{74}$$

We therefore have

$$SE = -R. \tag{75}$$

In view of (67) we have that

$$PSP = \frac{6}{h^4}T^2 P + \frac{36}{h^4}JJ^T, \tag{76}$$

where

$$J = \frac{\sqrt{2}}{2}\big[(\alpha - \beta)^{1/2}(e_1 - e_{N-1}), (\alpha + \beta)^{1/2}(e_1 + e_{N-1})\big]. \tag{77}$$

Inverting PSP and multiplying by PR, we have

$$-P^{-1}E = P^{-1}S^{-1}R = (PSP)^{-1}PR. \tag{78}$$

Our goal is to bound the elements of $P^{-1}E$ by Ch^4. Note that by Proposition 3 we have

$$\big|(PR)_1\big|, \big|(PR)_{N-1}\big| \le Ch, \\ \big|(PR)_j\big| \le Ch^4, \quad 2 \le j \le N-2. \tag{79}$$

Thus, we need to estimate $(PSP)^{-1}PR$. We decompose PSP as follows

$$PSP = GH^{-1}, \tag{80}$$

where

$$G = I + 6JJ^T P^{-1}T^{-2}, \qquad H = \frac{h^4}{6}P^{-1}T^{-2}, \tag{81}$$

so that

$$(PSP)^{-1} = HG^{-1}. \tag{82}$$

Note that with $L = (6/h^4)H$, $Q = 6JJ^T$, we have

$$G = I + QL. \tag{83}$$

We first estimate the elements of the matrix H.

Estimate of the Elements of H. In what follows we use C as expressing various constants that do not depend on h. As in (27), we can diagonalize H by

$$H = Z\Lambda'Z^T,$$

where the j-th column of the matrix Z is Z^j, as defined in (25). Recall that $P = 6I - T$ (see (23)), and that the eigenvalues λ_j of T are given by (24). Therefore, the eigenvalues κ_j, $1 \le j \le N-1$ of P are given by

$$\kappa_j = 6 - \lambda_j = 6 - 4\sin^2\left(\frac{j\pi}{2N}\right), \quad 1 \le j \le N-1. \tag{84}$$

The diagonal matrix Λ' contains the eigenvalues of H, which can be written as

$$\theta_j = \frac{h^4}{6}\lambda_j^{-2}\kappa_j^{-1} = \frac{h^4}{96}\frac{1}{\sin^4(\frac{j\pi}{2N})(6 - 4\sin^2(\frac{j\pi}{2N}))}, \quad j = 1, \ldots, N-1.$$

The element $H_{i,k}$ of the matrix H is

$$H_{i,k} = \sum_{j=1}^{N-1} Z_{i,j} \theta_j Z_{j,k}.$$

$$H_{i,k} = \sum_{j=1}^{N-1} \frac{h^4}{96} \frac{2}{N} \sin\left(\frac{ij\pi}{N}\right) \frac{\sin(\frac{ik\pi}{N})}{\sin^4(\frac{j\pi}{2N})(6 - 4\sin^2(\frac{j\pi}{2N}))}. \tag{85}$$

We can now estimate the order of magnitude of the elements of H as functions of h. In fact, we shall inspect separately the first and last columns of H and the rest ($k = 2, \ldots, N - 2$). The reason is that writing

$$\left(H G^{-1} P R\right)_i = \sum_{k=1}^{N-1} H_{i,k} \left(G^{-1} P R\right)_k, \tag{86}$$

we shall see that $(G^{-1} P R)_1$, $(G^{-1} P R)_{N-1}$ can only be estimated by Ch^2 (see (112) below), so that the additional accuracy should come from $H_{i,1}$, $H_{i,N-1}$. Consider first the elements (i, k) of H for $k = 1, N - 1$. It suffices to consider $k = 1$.

$$H_{i,1} = \sum_{j=1}^{N-1} \frac{h^4}{96} \frac{2}{N} \sin\left(\frac{ij\pi}{N}\right) \frac{1}{\sin^4(\frac{j\pi}{2N})(6 - 4\sin^2(\frac{j\pi}{2N}))} \sin\left(\frac{j\pi}{N}\right). \tag{87}$$

Recall the elementary inequalities

$$\sin x \geq \frac{2}{\pi} x, \quad 0 \leq x \leq \frac{\pi}{2}, \tag{88}$$

$$|\sin x| \leq |x|, \quad 2 \leq 6 - 4\sin^2\left(\frac{j\pi}{2N}\right) \leq 6. \tag{89}$$

Noting that $h = 1/N$ and using the estimate $|\sin(\frac{ij\pi}{N})| \leq 1$, we obtain

$$|H_{i,1}| = |H_{1,i}| \leq C \sum_{j=1}^{N-1} h^5 \frac{1}{(jh)^4} (jh) \leq Ch^2, \quad i = 2, \ldots, N - 2. \tag{90}$$

Similarly, we have

$$C_1 h^3 \leq H_{1,1} \leq C \sum_{j=1}^{N-1} h^5 \frac{1}{(jh)^4} (jh)^2 \leq C_2 h^3. \tag{91}$$

This estimate holds equally for $H_{N-1,N-1}$. For the other corner elements of H we have

$$|H_{1,N-1}| = |H_{N-1,1}| \leq C_2 h^3. \tag{92}$$

For $i, k = 2, \ldots, N - 2$ we have

$$|H_{i,k}| \leq C \sum_{j=1}^{N-1} h^5 \frac{1}{(jh)^4} \leq Ch. \tag{93}$$

Therefore, the orders of magnitude of the elements of H are bounded by

$$
\begin{bmatrix}
Ch^3 & Ch^2 & \dots & Ch^2 & Ch^3 \\
Ch^2 & Ch & \dots & Ch & Ch^2 \\
\vdots & \vdots & \dots & \dots & \vdots \\
Ch^2 & Ch & \dots & Ch & Ch^2 \\
Ch^3 & Ch^2 & \dots & Ch^2 & Ch^3
\end{bmatrix}.
\tag{94}
$$

Estimate of the Elements of G^{-1}. We show that G is invertible and we estimate its elements. First note that the elements of L are the elements of H multiplied by $6/h^4$.

The matrix Q is $(N-1) \times (N-1)$, but it has only four non-zero components at the corner positions,

$$
Q_{1,1} = Q_{N-1,N-1} = 6\alpha, \qquad Q_{1,N-1} = Q_{N-1,1} = 6\beta.
\tag{95}
$$

Therefore, QL has only two non-zero rows—the first and the last. The first row is given by

$$
(QL)_{1,j} = 6(\alpha L_{1,j} + \beta L_{N-1,j}), \quad j = 1, \dots, N-1
$$

and the last row is given by

$$
(QL)_{N-1,j} = 6(\beta L_{1,j} + \alpha L_{N-1,j}), \quad j = 1, \dots, N-1.
$$

Thus,

$$
\begin{cases}
G_{1,1} = 1 + 6(\alpha L_{1,1} + \beta L_{N-1,1}) =: a_1, \\
G_{1,N-1} = 6(\alpha L_{1,N-1} + \beta L_{N-1,N-1}) =: a_{N-1}, \\
G_{1,j} = 6(\alpha L_{1,j} + \beta L_{N-1,j}) =: b_j, \quad j = 2, \dots, N-2
\end{cases}
\tag{96}
$$

and

$$
\begin{cases}
G_{N-1,1} = 6(\beta L_{1,1} + \alpha L_{N-1,1}) = G_{1,N-1} = a_{N-1}, \\
G_{N-1,N-1} = 1 + 6(\beta L_{1,N-1} + \alpha L_{N-1,N-1}) = G_{1,1} = a_1, \\
G_{N-1,j} = 6(\beta L_{1,j} + \alpha L_{N-1,j}) = b_{N-j}, \quad j = 2, \dots, N-2,
\end{cases}
\tag{97}
$$

where the symmetries of L have been used. In rows $2, 3, \dots, N-2$ the matrix G has 1 on the diagonal and otherwise it is zero.

The orders of magnitude of a_1, a_{N-1} and b_j ($2 \le j \le N-2$) follow from those of the elements of L. Namely, $|a_1|$, $|a_{N-1}| \le C/h$ and $|b_j| \le C/h^2$ for $j = 2, \dots, N-2$. In what follows we shall need lower bounds for a_1 and $a_1^2 - a_{N-1}^2$. From their definitions above it is seen that we need an inspection of the terms $L_{1,1} = (6/h^4)H_{1,1}$, $L_{1,N-1} = (6/h^4)H_{1,N-1}$. Using the definitions of $L_{1,1}$, and $L_{1,N-1}$, we obtain

$$
L_{1,1} > |L_{1,N-1}|
$$

$$
L_{1,1} \mp L_{1,N-1} = \frac{h}{4} \sum_{\substack{j=1 \\ j \text{ even or odd}}}^{N-1} \frac{\sin^2 \frac{j\pi}{N}}{\sin^4 \frac{j\pi}{2N}(6 - 4\sin^2 \frac{j\pi}{2N})}
\tag{98}
$$

$$
\ge Ch \sum_{\substack{j=1 \\ j \text{ even or odd}}}^{N-1} \frac{(jh)^2}{(jh)^4} = \frac{C}{h}.
$$

In the above, take "$j = even$" for "$-$" and "$j = odd$" for "$+$".

Using (96) and the bounds $3 \leq \alpha \leq 4$, $|\beta| \leq 1$ (Remark 5), we get in view of (98) and (91)

$$|a_1| \geq 6(3L_{1,1} - |L_{N-1,1}|) - 1 \geq 12L_{1,1} - 1 \geq \frac{C}{h}. \tag{99}$$

Next, we treat the difference $a_1^2 - a_{N-1}^2$. Since we have the upper bound $|a_1^2 - a_{N-1}^2| \leq C_1/h^2$, we again need only a lower bound. We write the difference $a_1^2 - a_{N-1}^2$ as

$$a_1^2 - a_{N-1}^2 = \left[1 + 6(\alpha + \beta)(L_{1,1} + L_{1,N-1})\right] \cdot \left[1 + 6(\alpha - \beta)(L_{1,1} - L_{1,N-1})\right] \tag{100}$$

(using the symmetries of L). In view of (98) and $\alpha \geq 3$, $|\beta| \leq 1$, we obtain

$$\left|a_1^2 - a_{N-1}^2\right| \geq C_2/h^2. \tag{101}$$

To compute the inverse of G, we apply Gaussian elimination using the following method. We perform operations on rows of G and apply the same operations to the identity matrix I. When G is transformed to the identity matrix, I is transformed to G^{-1}.

We first divide the first and the last row of G by a_1 and annihilate the terms $j = 2, \ldots, N - 2$ of both rows by subtracting suitable multiplies of rows $2, \ldots, N - 2$. Add the result to the first row, for $j = 2, 3, \ldots, N - 2$. The result is G_1, where

$$G_1 = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \frac{a_{N-1}}{a_1} \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \ldots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ \frac{a_{N-1}}{a_1} & 0 & 0 & \ldots & 0 & 0 & 1 \end{bmatrix}. \tag{102}$$

The same operations on the identity matrix yield the matrix

$$I_1 = \begin{bmatrix} \frac{1}{a_1} & \frac{-b_2}{a_1} & \frac{-b_3}{a_1} & \ldots & \frac{-b_{N-3}}{a_1} & \frac{-b_{N-2}}{a_1} & 0 \\ 0 & 1 & -0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \ldots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & \frac{-b_{N-2}}{a_1} & \frac{-b_{N-3}}{a_1} & \ldots & \frac{-b_3}{a_1} & \frac{-b_2}{a_1} & \frac{1}{a_1} \end{bmatrix}. \tag{103}$$

In order to eliminate the non-zero element of G_1 in position $(N-1, 1)$, we subtract a suitable multiple of the first row and add the result to the last row, thus getting the transformed matrix G_2

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \frac{a_{N-1}}{a_1} \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \ldots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & \frac{a_1^2 - a_{N-1}^2}{a_1^2} \end{bmatrix}. \tag{104}$$

The corresponding matrix I_2 (obtained similarly from I_1) is

$$I_2 = \begin{bmatrix} \frac{1}{a_1} & \frac{-b_2}{a_1} & \frac{-b_3}{a_1} & \cdots & \frac{-b_{N-3}}{a_1} & \frac{-b_{N-2}}{a_1} & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ \frac{-a_{N-1}}{a_1^2} & \frac{b_2 a_{N-1} - a_1 b_{N-2}}{a_1^2} & \frac{b_3 a_{N-1} - a_1 b_{N-3}}{a_1^2} & \cdots & \frac{b_{N-3} a_{N-1} - a_1 b_3}{a_1^2} & \frac{b_{N-2} a_{N-1} - a_1 b_2}{a_1^2} & \frac{1}{a_1} \end{bmatrix}. \tag{105}$$

Now we divide the last row of G_2 and I_2 by $\frac{a_1^2 - a_{N-1}^2}{a_1^2}$. We get

$$G_3 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \frac{a_{N-1}}{a_1} \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{106}$$

and

$$I_3 = \begin{bmatrix} \frac{1}{a_1} & \frac{-b_2}{a_1} & \frac{-b_3}{a_1} & \cdots & \frac{-b_{N-3}}{a_1} & \frac{-b_{N-2}}{a_1} & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ \frac{-a_{N-1}}{a_1^2 - a_{N-1}^2} & \frac{b_2 a_{N-1} - a_1 b_{N-2}}{a_1^2 - a_{N-1}^2} & \frac{b_3 a_{N-1} - a_1 b_{N-3}}{a_1^2 - a_{N-1}^2} & \cdots & \frac{b_{N-3} a_{N-1} - a_1 b_3}{a_1^2 - a_{N-1}^2} & \frac{b_{N-2} a_{N-1} - a_1 b_2}{a_1^2 - a_{N-1}^2} & \frac{a_1}{a_1^2 - a_{N-1}^2} \end{bmatrix}. \tag{107}$$

Finally, we eliminate the $(1, N-1)$ element in G_3 by subtracting a multiple of the last row. The corresponding operation on I_3 yields the inverse G^{-1} as

$$G^{-1} = \begin{bmatrix} \frac{a_1}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_{N-2} - a_1 b_2}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_{N-3} - a_1 b_3}{a_1^2 - a_{N-1}^2} & \cdots & \frac{a_{N-1} b_3 - a_1 b_{N-3}}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_2 - a_1 b_{N-2}}{a_1^2 - a_{N-1}^2} & \frac{-a_{N-1}}{a_1^2 - a_{N-1}^2} \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ \frac{-a_{N-1}}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_2 - a_1 b_{N-2}}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_3 - a_1 b_{N-3}}{a_1^2 - a_{N-1}^2} & \cdots & \frac{a_{N-1} b_{N-3} - a_1 b_3}{a_1^2 - a_{N-1}^2} & \frac{a_{N-1} b_{N-2} - a_1 b_2}{a_1^2 - a_{N-1}^2} & \frac{a_1}{a_1^2 - a_{N-1}^2} \end{bmatrix}. \tag{108}$$

We give accurate estimates for the non-trivial elements of G^{-1}, those in the first and last rows. Using (99), (101) and the corresponding upper bounds, one readily observes that

$$\left|(G^{-1})_{1,1}\right|, \left|(G^{-1})_{N-1,N-1}\right|, \left|(G^{-1})_{N-1,1}\right|, \left|(G^{-1})_{1,N-1}\right| \le Ch. \tag{109}$$

Similarly, and using also $|b_j| \le C/h^2$, we get

$$\left|(G^{-1})_{1,j}\right|, \left|(G^{-1})_{N-1,j}\right| \le \frac{C}{h}, \quad j = 2, \ldots, N-2. \tag{110}$$

Therefore, the elements of G^{-1} are bounded by

$$G^{-1} = \begin{bmatrix} Ch & C/h & C/h & \ldots & C/h & C/h & Ch \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \ldots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ Ch & C/h & C/h & \ldots & C/h & C/h & Ch \end{bmatrix}. \tag{111}$$

We can now bound the elements of $G^{-1}PR$ using (109)–(110) and (79).

$$\begin{aligned} \left|(G^{-1}PR)_1\right| &\le \sum_{k=1}^{N-1} \left|(G^{-1})_{1,k}\right| \cdot \left|(PR)_k\right| \\ &= \left|(G^{-1})_{1,1}\right| \cdot \left|(PR)_1\right| + \sum_{k=2}^{N-2} \left|(G^{-1})_{1,k}\right| \cdot \left|(PR)_k\right| \\ &\quad + \left|(G^{-1})_{1,N-1}\right| \cdot \left|(PR)_{N-1}\right| \\ &\le C_1 h \cdot h + C_2(N-3)(1/h) \cdot h^4 \le Ch^2. \end{aligned} \tag{112}$$

Similarly, we have that $\left|(G^{-1}PR)_{N-1}\right| \le Ch^2$.

For $i = 2, \ldots, N-2$

$$\left|(G^{-1}PR)_i\right| = \left|(PR)_i\right| \le Ch^4. \tag{113}$$

Finally we consider the product $HG^{-1}PR$ (see (78), (82))

$$-P^{-1}E = HG^{-1}PR. \tag{114}$$

Combining the estimates (90)–(93) with (112)–(113), we obtain

$$\begin{aligned} \left|(HG^{-1}PR)_i\right| &\le \sum_{k=1}^{N-1} |H_{i,k}| \cdot \left|(G^{-1}PR)_k\right| \\ &= |H_{i,1}| \cdot \left|(G^{-1}PR)_1\right| + \sum_{k=2}^{N-2} |H_{i,k}| \cdot \left|(G^{-1}PR)_k\right| \\ &\quad + |H_{i,N-1}| \cdot \left|(G^{-1}PR)_{N-1}\right| \\ &\le C_1 h^2 h^2 + C_2(N-3)hh^4 \le Ch^4. \end{aligned} \tag{115}$$

Therefore,

$$\left|\left(P^{-1}E\right)_i\right| = \left|\left(HG^{-1}PR\right)_i\right| \le Ch^4, \quad 1 \le i \le N-1. \tag{116}$$

Conclusion of the Proof of Theorem 6. Using (116) we obtain that the Euclidean norm of the vector $E = U - U^*$ satisfies the estimate

$$|E| = \left|PP^{-1}E\right| \le C\left|P^{-1}E\right| \overset{(116)}{\le} C\sqrt{\sum_{i=1}^{N-1}(h^4)^2} = Ch^{-1/2}h^4. \tag{117}$$

Thus, in view of the definition of the l_2 norm

$$|\mathfrak{e}|_h \le Ch^4. \tag{118}$$

This proves the fourth order error estimate result. □

6 Numerical Results

In order to assess the spatial fourth-order accuracy of the scheme, we performed several numerical tests. In the tables below we show $e_{\max}$—the error in the maximum norm, and e_2—the error in the l_2 norm.

$$e_{\max} = \max|u_{\text{comp}} - u_{\text{exact}}|,$$

$$e_2 = \|u_{\text{comp}} - u_{\text{exact}}\|_{l^2} = |u_{\text{comp}} - u_{\text{exact}}|_h.$$

Here, u_{comp} and u_{exact} are the computed and the exact solutions, respectively.

We illustrate the numerical properties of the scheme (35) as follows.

- The scheme (35) is observed to be fourth-order accurate in the maximum and the discrete l^2 norms, whenever homogeneous or nonhomogeneous boundary conditions are applied. This is shown in Case 1.
- In case of highly oscillatory solutions, the scheme (35) behaves remarkably well. Case 2 describes the convergence of the scheme for such a family of solutions. In this case too fourth-order accuracy is observed. In addition, the magnitude of the errors is very small even for coarse grids.
- Finally, in Case 3 we show that the scheme can be also used for nonlinear biharmonic equations, retaining the fourth-order accuracy.

6.1 Case 1: Polynomial Solutions

We consider two polynomial solutions. The first corresponds to homogeneous boundary conditions and the second to inhomogeneous conditions.

6.1.1 Homogeneous Boundary Conditions

Consider the polynomial solution of $u^{(4)} = f$

$$u(x) = x^4(x-1)^2 \quad \text{on } [a,b] = [0,1].$$

Table 1 Compact scheme for $u^{(4)} = f$ with exact solution: $u = x^4(x-1)^2$ on [0, 1]. We present e_{max} the error in the maximum norm, and e_2 the error in the l_2 norm

Mesh	$N = 16$	Rate	$N = 32$	Rate	$N = 64$	Rate	$N = 128$
e_{max}	7.8231(−6)	4.00	4.8894(−7)	4.00	3.0589(−8)	4.00	1.9106(−9)
e_2	5.6157(−6)	4.00	3.5099(−7)	4.00	2.1937(−8)	4.00	1.3739(−9)

It satisfies

$$u(0) = u'(0) = u(1) = u'(1) = 0. \tag{119}$$

Thus, choosing

$$f(x) = u^{(4)} = 360x^2 - 240x + 24, \tag{120}$$

then $u(x)$ is the unique solution of the biharmonic problem

$$\begin{cases} u^{(4)} = f, & 0 < x < 1, \\ u(0) = u(1) = 0, \\ u'(0) = u'(1) = 0. \end{cases} \tag{121}$$

Our objective is to recover approximations u_i of $u(x_i)$ from the knowledge of the discrete data $f(x_i)$ on the grid $0 = x_0 < x_1 < \cdots < x_{N-1} < x_N = 1$. The problem (121) is approximated by

$$\begin{cases} \delta_x^4 u_j = f(x_j), & 1 \leq j \leq N-1, \\ u_0 = u_N = 0, \\ u_{x,0} = u_{x,N} = 0. \end{cases} \tag{122}$$

In Table 1 we display numerical results for the fourth-order scheme (122). Observe that fourth-order accuracy is achieved in both the maximum and the l^2 norms.

6.1.2 Nonhomogeneous Boundary Conditions

Here we consider a polynomial solution, but with nonhomogeneous values at the two end points,

$$u(x) = x^5, \quad \text{on } [a, b] = [0, 1].$$

The function $u(x)$ is the solution of the biharmonic problem

$$\begin{cases} u^{(4)} = f, & 0 < x < 1, \\ u(0) = 0, & u(1) = 1, \\ u'(0) = 0, & u'(1) = 5, \end{cases} \tag{123}$$

where the function $f(x)$ is

$$f(x) = u^{(4)} = 120x. \tag{124}$$

Table 2 Compact scheme for $u^{(4)} = f$ with exact solution: $u = x^5$ on [0, 1]. We present e_{max} the error in the maximum norm, and e_2 the error in the l_2 norm

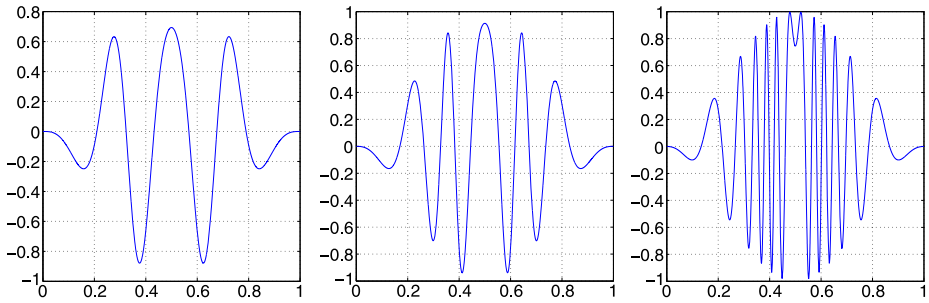| Mesh	$N = 16$	Rate	$N = 32$	Rate	$N = 64$	Rate	$N = 128$
e_{max}	9.6857(−7)	3.99	6.1118(−8)	4.00	3.8200(−9)	3.98	2.4129(−10)
e_2	7.0187(−7)	4.00	4.3873(−8)	4.00	2.7420(−9)	4.00	1.7062(−10)

Fig. 1 The oscillating function $x \mapsto 16x^2(1 - x)^2 \sin(1/((x - 0.5)^2 + \varepsilon))$ for $\varepsilon = 7.5 \times 10^{-2}$ (*left*), $\varepsilon = 5.0 \times 10^{-2}$ (*center*), $\varepsilon = 2.5 \times 10^{-2}$ (*right*)

Thus, we resolve numerically

$$\begin{cases} \delta_x^4 u_j = f(x_j), & 1 \le j \le N - 1, \\ u_0 = 0, & u_N = 1, \\ u_{x,0} = 0, & u_{x,N} = 5. \end{cases} \tag{125}$$

Our purpose is to demonstrate the fourth-order accuracy of the scheme for the case of nonhomogeneous boundary conditions. Indeed, the numerical results reported in Table 2 assesses the fourth-order accuracy of the scheme in this case too.

6.2 Case 2: Oscillating Solutions

We consider a family of functions defined by

$$u_\varepsilon(x) = p(x) \sin\big(1/q_\varepsilon(x)\big), \tag{126}$$

where the polynomial functions $p(x)$ and $q_\varepsilon(x)$ are given by

$$p(x) = 16x^2(1 - x)^2, \qquad q_\varepsilon(x) = 1/\big((x - 1/2)^2 + \varepsilon\big), \quad \varepsilon > 0. \tag{127}$$

For small ε the function u_ϵ oscillates in the middle of the interval. The parameter ε serves as a tuning parameter for the frequency of the oscillations.

In Fig. 1 we display the functions $u_\varepsilon(x)$ corresponding to

$$\varepsilon = 7.5 10^{-2}, \qquad \varepsilon = 5.0 10^{-2}, \qquad \varepsilon = 2.5 10^{-2}. \tag{128}$$
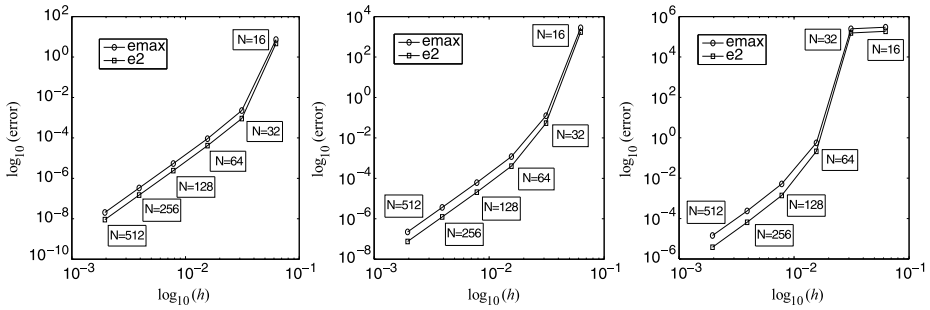
Fig. 2 Convergence rates for the problem (129) with $\varepsilon = 7.5 \times 10^{-2}$ (*left*), $\varepsilon = 5.0 \times 10^{-2}$ (*center*), $\varepsilon = 2.5 \times 10^{-2}$ (*right*). *Circles* correspond to the maximum norm and *squares* correspond to the l^2 norm

As in Case 1, we consider the approximation

$$\begin{cases} \delta_x^4 u_j = f_\varepsilon(x_j), & 1 \leq j \leq N-1, \\ u_0 = 0, \quad u_N = 0, \\ u_{x,0} = 0, \quad u_{x,N} = 0, \end{cases} \tag{129}$$

where the function $f_\varepsilon(x)$ is defined by

$$f_\varepsilon(x) = u_\varepsilon^{(4)}(x). \tag{130}$$

The results are reported in Fig. 2 on a LogLog scale. In addition, in Table 3 we display the errors for different values of ε and N. The results clearly demonstrate the asymptotic fourth-order convergence in both norms. The magnitude of the errors on relatively coarse grids is remarkably small. Observe that a maximum error of order 10^{-3} is obtained for $\varepsilon = 7.5 \times 10^{-2}$, $\varepsilon = 5.0 \times 10^{-2}$ and $\varepsilon = 2.5 \times 10^{-2}$ with $N = 32$, $N = 64$ and $N = 128$, respectively.

6.3 Case 3: A Nonlinear Biharmonic Equation

As a final example we consider the nonlinear problem

$$\begin{cases} u^{(4)} - H(u) = f, & 0 < x < 1, \\ u(0) = 0, \quad u(1) = 0, \\ u'(0) = 0, \quad u'(1) = 0, \end{cases} \tag{131}$$

where H is assumed to be a k-Lipschitz function.

The approximation of (131) is obtained via the (nonlinear) scheme

$$\begin{cases} \delta_x^4 u_j - H(u_j) = f(x_j), & 1 \leq j \leq N-1, \\ u_0 = 0, \quad u_N = 0, \\ u_{x,0} = 0, \quad u_{x,N} = 0. \end{cases} \tag{132}$$

Equation (131) has a unique solution under the sufficient condition

$$k < \lambda_{\min}, \tag{133}$$

Table 3 Compact scheme for $u^{(4)} = f_\varepsilon$ for $\varepsilon = 7.5 \times 10^{-2}$, $\varepsilon = 5.0 \times 10^{-2}$, $\varepsilon = 2.5 \times 10^{-2}$. We present e_{max} the error in the maximum norm, and e_2 the error in the l_2 norm

Mesh	N = 16	Rate	N = 32	Rate	N = 64	Rate	N = 128	Rate	N = 256	Rate	N = 512
e_{max}, $\varepsilon = 7.5 \times 10^{-2}$	7.3136(+0)	11.67	2.2523(−3)	4.62	9.1729(−5)	4.09	5.3870(−6)	4.02	3.3124(−7)	4.01	2.0501(−8)
$\varepsilon = 5.0 \times 10^{-2}$	2.7999(+3)	14.45	1.2469(−1)	6.72	1.1839(−3)	4.27	6.1165(−5)	4.06	3.6542(−6)	4.04	2.2279(−7)
$\varepsilon = 2.5 \times 10^{-2}$	2.9713(+5)	0.26	2.4744(+5)	18.75	5.6056(−1)	6.75	5.1981(−3)	4.44	2.3902(−4)	4.04	1.4556(−6)
e_2, $\varepsilon = 7.5 \times 10^{-2}$	4.7042(0)	12.34	9.0697(−4)	4.45	4.1464(−5)	4.11	2.4066(−6)	4.03	1.4768(−7)	4.03	9.0506(−9)
$\varepsilon = 5.0 \times 10^{-2}$	1.7151(+3)	14.97	5.3507(−2)	7.06	3.9973(−4)	4.28	2.0575(−5)	4.07	1.2252(−6)	4.02	7.5503(−8)
$\varepsilon = 2.5 \times 10^{-2}$	1.8588(+5)	0.27	1.5403(+5)	19.46	2.1295(−1)	7.25	1.3947(−3)	4.41	6.5468(−5)	4.11	3.8041(−6)

Table 4 Compact scheme for $u^{(4)} - 100 \sin^2 u = f$ with exact solution: $u = u_\varepsilon(x)$ on $[0, 1]$, with $\varepsilon = 5.0 \times 10^{-2}$. We present $e_{\max}$ the error in the maximum norm, and e_2 the error in the l_2 norm

Mesh	$N = 64$	Rate	$N = 128$	Rate	$N = 256$	Rate	$N = 512$
$e_{\max}$	$1.2391(-3)$	4.27	$6.4175(-5)$	4.06	$3.8378(-6)$	4.04	$2.3401(-7)$
e_2	$4.0399(-4)$	4.28	$2.0839(-5)$	4.07	$1.2417(-6)$	4.03	$7.6153(-8)$

where $\lambda_{\min}$ is the smallest eigenvalue of the problem

$$\begin{cases} u^{(4)} = \lambda u, \\ u(0) = 0, \qquad u(1) = 0, \\ u'(0) = 0, \qquad u'(1) = 0. \end{cases} \tag{134}$$

A sufficient condition for (133) to hold is that

$$k < \left(\frac{3\pi}{2}\right)^4. \tag{135}$$

In Table 4 we display numerical results for the function $H(u) = 100 \sin^2 u$. Here the right-hand side f is selected as $u_\varepsilon^{(4)} - H(u_\varepsilon)$, with $\varepsilon = 5.0 \times 10^{-2}$ (see Case 2). Observe the fourth-order accuracy of the scheme.

References

1. Abarbanel, S., Ditkowski, A., Gustafsson, B.: On error bound of finite difference approximations for partial differential equations. J. Sci. Comput. **15**(1), 79–116 (2000)
2. Altas, I., Dym, J., Gupta, M.M., Manohar, R.P.: Multigrid solution of automatically generated high-order discretizations for the biharmonic equation. SIAM J. Sci. Comput. **19**, 1575–1585 (1998)
3. Ben-Artzi, M., Croisille, J.-P., Fishelov, D.: Convergence of a compact scheme for the pure streamfunction formulation of the unsteady Navier-Stokes system. SIAM J. Numer. Anal. **44**(5), 1997–2024 (2006)
4. Ben-Artzi, M., Croisille, J.-P., Fishelov, D.: A fast direct solver for the biharmonic problem in a rectangular grid. SIAM J. Sci. Comput. **31**(1), 303–333 (2008)
5. Brüger, A., Gustafsson, B., Lötstedt, P., Nilsson, J.: High order accurate solution of the incompressible Navier-Stokes equations. J. Comput. Phys. **203**, 49–71 (2005)
6. Carey, G.F., Spotz, W.F.: High-order compact scheme for the stream-function vorticity equations. Int. J. Numer. Methods Eng. **38**, 3497–3512 (1995)
7. Carey, G.F., Spotz, W.F.: Extension of high-order compact schemes to time dependent problems. Numer. Methods Partial Differ. Equ. **17**(6), 657–672 (2001)
8. Carpenter, M.H., Gottlieb, D., Abarbanel, S.: The stability of numerical boundary treatments for compact high-order schemes finite difference schemes. J. Comput. Phys. **108**, 272–295 (1993)
9. Weinan, E., Liu, J.-G.: Essentially compact schemes for unsteady viscous incompressible flows. J. Comput. Phys. **126**, 122–138 (1996)
10. Fishelov, D., Ben-Artzi, M., Croisille, J.-P.: Recent developments in the pure streamfunction formulation of the Navier-Stokes system. J. Sci. Comput. **45**(1–3), 238–258 (2010)
11. Gupta, M.M., Manohar, R.P., Stephenson, J.W.: Single cell high order scheme for the convection-diffusion equation with variable coefficients. Int. J. Numer. Methods Fluids **4**, 641–651 (1984)
12. Gustafsson, B.: The convergence rate for difference approximations to mixed initial boundary value problems. Math. Comput. **29**, 396–406 (1975)
13. Gustafsson, B.: The convergence rate for difference approximations to general mixed initial boundary value problems. SIAM J. Numer. Anal. **18**, 179–190 (1981)

14. Lele, S.K.: Compact finite-difference schemes with spectral-like resolution. J. Comput. Phys. **103**, 16–42 (1992)
15. Li, Ming, Tang, Tao: A compact fourth-order finite difference scheme for unsteady viscous incompressible flows. J. Sci. Comput. **16**(1), 29–45 (2001)
16. Stephenson, J.W.: Single cell discretizations of order two and four for biharmonic problems. J. Comput. Phys. **55**, 65–80 (1984)