Journal of Scientific Computing (2019) 78:772–815
https://doi.org/10.1007/s10915-018-0789-4

An Ultra-weak Discontinuous Galerkin Method for Schrödinger Equation in One Dimension

Anqi Chen[1] · Fengyan Li[2] · Yingda Cheng[1,3]

Received: 4 April 2018 / Revised: 15 June 2018 / Accepted: 18 July 2018 / Published online: 25 July 2018

Abstract
In this paper, we develop an ultra-weak discontinuous Galerkin method to solve the one-dimensional nonlinear Schrödinger equation. Stability conditions and error estimates are derived for the scheme with a general class of numerical fluxes. The error estimates are based on detailed analysis of the projection operator associated with each individual flux choice. Depending on the parameters, we find out that in some cases, the projection can be defined element-wise, facilitating analysis. In most cases, the projection is global, and its analysis depends on the resulting 2×2 block-circulant matrix structures. For a large class of parameter choices, optimal a priori L^2 error estimates can be obtained. Numerical examples are provided verifying theoretical results.

Keywords Ultra-weak discontinuous Galerkin method · Stability · Error estimates · Projection · One-dimensional Schrödinger equation

1 Introduction

In this paper, we develop and analyze a discontinuous Galerkin (DG) method for one-dimensional nonlinear Schrödinger (NLS) equation:

$$iu_t + u_{xx} + f(|u|^2)u = 0, \tag{1}$$

Fengyan Li: Research is supported by NSF Grant DMS-1719942. Yingda Cheng: Research is supported by NSF Grants DMS-1453661, DMS-1720023 and the Simons Foundation.

✉ Yingda Cheng
ycheng@msu.edu

Anqi Chen
chenanq3@msu.edu

Fengyan Li
lif@rpi.edu

[1] Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

[2] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

[3] Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

where $f(u)$ is a nonlinear real function and u is a complex function. The Schrödinger equation is the fundamental equation in quantum mechanics, reaching out to many applications in fluid dynamics, nonlinear optics and plasma physics. It is also called Schrödinger wave equation as it can describe how the wave functions of a physical system evolve over time. Many numerical methods have been applied to solve NLS equations [5,14,17,18,23,28,30]. In [5, 30], several important finite difference schemes are implemented, analyzed and compared. In [23], the author introduced a pseudo-spectral method for general NLS equations. Many finite element methods have been tested, such as quadratic B-spline for NLS in [14,28] and space–time DG method for nonlinear (cubic) Schrödinger equation in [17,18]. In this paper, we focus on the DG methods, which is a class of finite element methods using completely discontinuous piecewise function space for test functions and numerical solution, to solve the Schrödinger equation. The first DG method was introduced by Reed and Hill [24]. A major development of DG methods is the Runge–Kutta DG (RKDG) framework introduced for solving hyperbolic conservation laws containing only first order spatial derivatives in a series of papers [9–13]. Because of the completely discontinuous basis, DG methods have several attractive properties. It can be used on many types of meshes, even those with hanging nodes. The methods have h-p adaptivity and very high parallel efficiency.

Various types of DG schemes for discretizing the second order spatial derivatives have been used to compute (1). One group of such methods is the so-called local DG (LDG) method invented in [12] for convection–diffusion equations. The algorithm is based on introducing auxiliary variables and reformulating the equation into its first order form. In [32], an LDG method using alternating fluxes is developed with L^2 stability and proved $(k + \frac{1}{2})$-th order of accuracy. Later in [33], Xu and Shu proved optimal accuracy for both the solution and the auxiliary variables in the LDG method for high order wave equations based on refined energy estimates. In [19], the authors presented an LDG method with exponential time differencing Runge–Kutta scheme and investigated the energy conservation performance of the scheme. Another group of method involves treating the second order spatial derivative directly in the weak formulations, such as IPDG method [15,31] and NIPG method [26,27]. Those schemes enforce a penalty jump term in the weak formulation, and they have been extensively applied to acoustic and elastic wave propagations [1,16,25]. As for Schrödinger equations, the direct DG (DDG) method was applied to Schrödinger equation in [21] and achieved energy conservation and optimal accuracy. Among all those various formulations, the work in this paper focus on the ultra-weak DG methods, which can be traced backed to [4], and refer to those DG methods [29] that rely on repeatedly applying integration by parts so all the spatial derivatives are shifted from the solution to the test function in the weak formulations. In [7], Cheng and Shu developed ultra-weak DG methods for general time dependent problems with higher order spatial derivatives. In [3], Bona et al. proposed an ultra-weak DG scheme for generalized KdV equation and performed error estimates.

The focus of this paper is the investigation of a most general form of the numerical flux functions that ensures stability along with our ultra-weak formulation. The fluxes under consideration include the alternating fluxes, and also the fluxes considered in [21], and therefore allows for flexibility for the design of the schemes. It is widely known that the choice of flux can have significant impact on the convergence order of the scheme as evidenced in DG methods for linear first-order transport equations, two-way wave equations [6], and the KdV equations [3,7] and many others. The main contribution of the work is a systematic study of error estimates based on the flux parameters. To this end, we define and analyze projection operator associated with each specific parameter choice. We assume the dependence of parameters on the mesh size can be freely enforced, therefore many cases shall follow. We find out that under certain conditions, the projections are "local", meaning that

they can be defined element-wise. In the most general setting, the projections are global, and detailed analysis based on block-circulant matrices are necessary. This type of analysis has been done in [3,22] for circulant matrices and in [20] for block-circulant matrices, but our case is more involved due to the 2×2 block-circulant structure, for which several cases need to be distinguished based on the eigenvalues of the block matrices, and some requires tools from Fourier analysis. Our analysis reveals that under a large class of parameter choices, our method is optimally convergent in L^2 norm, which is verified by extensive numerical tests for both the projection operators and the numerical schemes for (1).

The remainder of this paper is organized as follows. In Sect. 2, we introduce an ultra-weak DG method with general flux definitions for one-dimensional nonlinear Schrödinger equations and study its stability properties. The main body of the paper, the error estimates, is contained in Sect. 3. We introduce a new projection operator and analyze its properties in Sect. 3.1, which is later used in Sect. 3.2 to obtain the convergence results of the schemes. Numerical validations are provided in Sect. 4. Conclusions are made in Sect. 5. Some technical details, including proof of most lemmas are collected in the "Appendix".

2 A DG Method for One-Dimensional Schrödinger Equations

In this section, we formulate and discuss stability results of a DG scheme for one-dimensional NLS equation (1) on interval $I = [a, b]$ with initial condition $u(x, 0) = u_0(x)$ and periodic boundary conditions. Here $f(u)$ is a given real function. Our method can be defined for general boundary conditions, but the error analysis will require slightly different tools, and therefore we only consider periodic boundary conditions in this paper.

To facilitate the discussion, first we introduce some notations and definitions. For a 1-D interval $I = [a, b]$, the usual DG meshes are defined as:

$$\begin{aligned} &a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = b, \\ &I_j = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right), \quad x_j = \frac{1}{2}\left(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}}\right), \end{aligned}$$

and

$$h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}, \quad h = \max_j h_j,$$

with mesh regularity requirement $\frac{h}{\min h_j} < \sigma$, σ is fixed during mesh refinement.

The approximation space is defined as:

$$V_h^k = \{v_h : v_h|_{I_j} \in P^k(I_j),\ j = 1, \ldots, N\},$$

meaning v_h is a polynomial of degree up to k on each cell I_j. For a function $v_h \in V_h^k$, we use $(v_h)^-_{j-\frac{1}{2}}$ and $(v_h)^+_{j-\frac{1}{2}}$ to refer to the value of v_h at $x_{j-\frac{1}{2}}$ from the left cell I_{j-1} and the right cell I_j respectively. The jump and average are defined as $[v_h] = v_h^+ - v_h^-$ and $\{v_h\} = \frac{1}{2}(v_h^+ + v_h^-)$ at cell interfaces.

In this paper, we consider a DG scheme motivated by Cheng and Shu [7] and based on integration by parts twice, or the so-called ultra-weak formulation. In particular, we look for

the unique function $u_h = u_h(t) \in V_h^k$, $t \in (0, T]$, such that

$$i\int_{I_j}(u_h)_t v_h dx + \int_{I_j} u_h (v_h)_{xx} dx - \hat{u}_h (v_h)_x^-|_{j+\frac{1}{2}} + \hat{u}_h (v_h)_x^+|_{j-\frac{1}{2}} + \widetilde{(u_h)_x} v_h^-|_{j+\frac{1}{2}} - \widetilde{(u_h)_x} v_h^+|_{j-\frac{1}{2}} + \int_{I_j} f(|u_h|^2) u_h v_h dx = 0 \tag{2}$$

holds for all $v_h \in V_h^k$ and all $j = 1, \ldots, N$. Here, we require $k \geq 1$, because $k = 0$ yields an inconsistent scheme. Notice that (2) can be written equivalently in a weak formulation by performing another integration by parts back as:

$$i\int_{I_j}(u_h)_t v_h dx - \int_{I_j}(u_h)_x (v_h)_x dx + (u_h^- - \hat{u}_h)(v_h)_x^-|_{j+\frac{1}{2}} + (\hat{u}_h - u_h^+)(v_h)_x^+|_{j-\frac{1}{2}} + \widetilde{(u_h)_x} v_h^-|_{j+\frac{1}{2}} - \widetilde{(u_h)_x} v_h^+|_{j-\frac{1}{2}} + \int_{I_j} f(|u_h|^2) u_h v_h dx = 0 \tag{3}$$

The "hat" and"tilde" terms are the numerical fluxes we pick for u and u_x at cell boundaries, which are single valued functions defined as:

$$\widetilde{(u_h)_x} = \{(u_h)_x\} + \alpha_1 [(u_h)_x] + \beta_1 [u_h], \quad \hat{u}_h = \{u_h\} + \alpha_2 [u_h] + \beta_2 [(u_h)_x], \quad \alpha_1, \alpha_2 \in \mathbb{C}, \quad \beta_1, \beta_2 \in \mathbb{C}, \tag{4}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are prescribed parameters. They may depend on the mesh parameter h. Commonly used fluxes such as the central flux (by setting $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$) and alternating fluxes (by setting $\alpha_1 = -\alpha_2 = \pm\frac{1}{2}, \beta_1 = \beta_2 = 0$) belong to this flux family. The direct DG scheme considered in [21] is a special case of our method when $\alpha_1 = -\alpha_2, \beta_1 = \frac{c}{h}, \beta_2 = 0, c > 0, \alpha_1 \in \mathbb{R}$. The IPDG method can also be casted in this framework as $\alpha_1 = \alpha_2 = \beta_2 = 0, \beta_1 = \frac{c}{h}, c > 0$.

Using periodic boundary condition, we can sum up on j for the numerical scheme (2) and reduce it into the following short-hand notation

$$a_{\alpha_1,\alpha_2,\beta_1,\beta_2}(u_h, v_h) - i\int_I f(|u_h|^2) u_h v_h dx = 0, \tag{5}$$

where

$$a_{\alpha_1,\alpha_2,\beta_1,\beta_2}(u_h, v_h) = \int_I (u_h)_t v_h dx - i\int_I u_h (v_h)_{xx} dx - i\sum_j (\hat{u}_h [(v_h)_x] - \widetilde{(u_h)_x} [v_h])|_{j+\frac{1}{2}}.$$

The following theorem contains the results on semi-discrete L^2 stability.

Theorem 2.1 (Stability) *The solution of semi-discrete DG scheme* (2) *using numerical fluxes* (4) *satisfies* L^2 *stability condition*

$$\frac{d}{dt}\int_I |u_h|^2 dx \leq 0,$$

if

$$Im\beta_2 \geq 0,\ Im\beta_1 \leq 0,\ |\alpha_1 + \overline{\alpha_2}|^2 \leq -4 Im\beta_1 Im\beta_2. \tag{6}$$

In particular, when all parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ *are restricted to be real, this condition amounts to*

$$\alpha_1 + \alpha_2 = 0 \tag{7}$$

without any requirement on β_1, β_2.

Proof From integration by parts, we have, for $\forall v_h \in V_h^k$

$$a_{\alpha_1,\alpha_2,\beta_1,\beta_2}(u_h, v_h) = \int_I (u_h)_t v_h dx + i \int_I (u_h)_x (v_h)_x dx + i \sum_j ([u_h (v_h)_x] - \hat{u}_h [(v_h)_x] + \widetilde{(u_h)_x}[v_h])|_{j+\frac{1}{2}}.$$

Taking $v_h = \bar{u}_h$ in (5) and compute its conjugate as well, we get

$$\begin{aligned} 0 &= i \int_I f(|u_h|^2)|u_h|^2 dx + \overline{i \int_I f(|u_h|^2)|u_h|^2 dx} \\ &= a_{\alpha_1,\alpha_2,\beta_1,\beta_2}(u_h, \bar{u}_h) + \overline{a_{\alpha_1,\alpha_2,\beta_1,\beta_2}(u_h, \bar{u}_h)} \\ &= \frac{d}{dt} \int_I |u_h|^2 dx - 2\mathrm{Im} \sum_j ([u_h (\bar{u}_h)_x] - \hat{u}_h [(\bar{u}_h)_x] + \widetilde{(u_h)_x}[\bar{u}_h])|_{j+\frac{1}{2}}. \end{aligned} \tag{8}$$

Define

$$\begin{aligned} A(u_h, \bar{u}_h) &= \sum_j ([u_h (\bar{u}_h)_x] - \hat{u}_h [(\bar{u}_h)_x] + \widetilde{(u_h)_x}[\bar{u}_h])|_{j+\frac{1}{2}} \\ &= \sum_j \Big(\{u_h\}[(\bar{u}_h)_x] + [u_h]\{(\bar{u}_h)_x\} - \big(\{u_h\} + \alpha_2 [u_h] + \beta_2 [(u_h)_x]\big)[(\bar{u}_h)_x] \\ &\quad + (\{(u_h)_x\} + \alpha_1 [(u_h)_x] + \beta_1 [u_h])[\bar{u}_h] \Big)|_{j+\frac{1}{2}} \\ &= \sum_j \big(2\mathrm{Re}([u_h]\{(\bar{u}_h)_x\}) - \beta_2 |[(u_h)_x]|^2 + \beta_1 |[u_h]|^2 \\ &\quad + \alpha_1 [(u_h)_x][\bar{u}_h] - \alpha_2 [u_h][(\bar{u}_h)_x] \big)|_{j+\frac{1}{2}}. \end{aligned}$$

Therefore, $\mathrm{Im} A(u_h, \bar{u}_h) = \sum_j (-\mathrm{Im}\beta_2 |[(u_h)_x]|^2 + \mathrm{Im}\beta_1 |[u_h]|^2 + \mathrm{Im}\{(\alpha_1 + \overline{\alpha_2})[\bar{u}_h][(u_h)_x]\})|_{j+\frac{1}{2}}$. Plug it back into (8):

$$\frac{d}{dt} \int_I |u_h|^2 dx + \sum_j 2\mathrm{Im}\beta_2 |[(u_h)_x]|^2 - 2\mathrm{Im}\beta_1 |[u_h]|^2 - 2\mathrm{Im}\{(\alpha_1 + \overline{\alpha_2})[\bar{u}_h][(u_h)_x]\}|_{j+\frac{1}{2}} = 0. \tag{9}$$

If the stability condition (6) is satisfied, we have

$$\frac{d}{dt} \int_I |u_h|^2 dx \le 0.$$

If all parameters are real and (7) is satisfied, then (9) further yields:

$$\frac{d}{dt} \int_I |u_h|^2 dx = 0,$$

which implies energy conservation. □

Remark 2.1 For simplicity of the discussion, in the next section, we will only consider real parameters, i.e. when $\alpha_1, \alpha_2, \beta_1, \beta_2$ are real and $\alpha_1 + \alpha_2 = 0$.

3 Error Estimates

In this section, we will derive error estimates of the DG scheme (2) for the model NLS equation (1). As mentioned before, we consider L^2 stable real parameter choices, which means the numerical fluxes are defined by three parameters as,

$$\widetilde{(u_h)_x} = \{(u_h)_x\} + \alpha_1[(u_h)_x] + \beta_1[u_h], \quad \hat{u}_h = \{u_h\} - \alpha_1[u_h] + \beta_2[(u_h)_x], \quad \alpha_1, \beta_1, \beta_2 \in \mathbb{R}. \tag{10}$$

We will focus on the impact of the choice of the parameters $\alpha_1, \beta_1, \beta_2$ on the accuracy of the scheme. We proceed as follows: first, we define and discuss the properties of projection operator $P_h^\star$ in Sect. 3.1. Then, we use the projection error estimates to obtain convergence result for DG scheme in Sect. 3.2.

3.1 Projection Operator

In this subsection, we perform detailed studies of a projection operator defined as follows.

Definition 3.1 For our DG scheme with flux choice (10), we define the associated projection operator $P_h^\star$ for any periodic function $u \in W^{1,\infty}(I)$ to be the unique polynomial $P_h^\star u \in V_h^k$ (when $k \geq 1$) satisfying

$$\int_{I_j} P_h^\star u \, v_h dx = \int_{I_j} u \, v_h dx \quad \forall v_h \in P^{k-2}(I_j), \tag{11a}$$

$$\widehat{P_h^\star u} = \{P_h^\star u\} - \alpha_1[P_h^\star u] + \beta_2[(P_h^\star u)_x] = u \quad \text{at} \quad x_{j+\frac{1}{2}}, \tag{11b}$$

$$\widetilde{(P_h^\star u)_x} = \{(P_h^\star u)_x\} + \alpha_1[(P_h^\star u)_x] + \beta_1[P_h^\star u] = u_x \quad \text{at} \quad x_{j+\frac{1}{2}}, \tag{11c}$$

for all j. When $k = 1$, only conditions (11b)–(11c) are needed.

This definition is to ensure $u - \widehat{P_h^\star u} = 0$ and $u_x - \widetilde{(P_h^\star u)_x} = 0$, which will be used in error estimates for the scheme. In the following, we analyze the projection when the parameter choice reduces it to a local projection in Sect. 3.1.1, and then we consider the more general global projection in Sect. 3.1.2.

3.1.1 Local Projection Results

In general, the projection $P_h^\star$ is globally defined, and its existence, uniqueness and approximation properties are quite complicated mathematically. However, with some special parameter choices, $P_h^\star$ can be reduced to a local projection, meaning that it can be solved element-wise, and hence the analysis can be greatly simplified.

For example, with the alternating fluxes $\alpha_1 = \pm\frac{1}{2}$, $\beta_1 = \beta_2 = 0$, $P_h^\star$ can be reduced to P_h^1 and P_h^2 defined below. $P_h^\star = P_h^1$ for parameter choice $\alpha_1 = \frac{1}{2}$, $\beta_1 = \beta_2 = 0$ is formulated as: for each cell I_j, we find the unique polynomial of degree k, $P_h^1 u$, satisfying

$$\int_{I_j} P_h^1 u \, v_h dx = \int_{I_j} u \, v_h dx \qquad \forall v_h \in P^{k-2}(I_j), \tag{12a}$$

$$(P_h^1 u)^- = u \qquad \text{at } x_{j+\frac{1}{2}}, \tag{12b}$$

$$(P_h^1 u)_x^+ = u_x \qquad \text{at } x_{j-\frac{1}{2}}. \tag{12c}$$

When $k = 1$, only conditions (12b)–(12c) are needed.

Similarly, we can define $P_h^\star = P_h^2$ for parameter choice $\alpha_1 = -\frac{1}{2}, \beta_1 = \beta_2 = 0$ as: for each cell I_j, we find the unique polynomial of degree k, $P_h^2 u$, satisfying

$$\int_{I_j} P_h^2 u\, v_h dx = \int_{I_j} u\, v_h dx \qquad \forall v_h \in P^{k-2}(I_j), \tag{13a}$$

$$(P_h^2 u)^+ = u \qquad \text{at } x_{j-\frac{1}{2}}, \tag{13b}$$

$$(P_h^2 u)_x^- = u_x \qquad \text{at } x_{j+\frac{1}{2}}. \tag{13c}$$

When $k = 1$, only conditions (13b)–(13c) are needed.

Similar local projections have been introduced and considered in [7]. It is obvious that $P_h^1 u$, $P_h^2 u$ can be solved element-wise, and their existence, uniqueness are straightforward. From a standard scaling argument by Bramble–Hilbert lemma in [8], P_h^1 and P_h^2 have the following error estimates: let $u \in W^{k+1,p}(I_j)(p = 2, \infty)$, then

$$\begin{aligned} &\|u - P_h^v u\|_{L^p(I_j)} \le C h_j^{k+1} |u|_{W^{k+1,p}(I_j)}, \quad p = 2, \infty, \ v = 1, 2, \\ &\|u_x - P_h^v u_x\|_{L^p(I_j)} \le C h_j^{k} |u|_{W^{k+1,p}(I_j)}, \quad p = 2, \infty, \ v = 1, 2, \end{aligned} \tag{14}$$

where here and below, C is a generic constant that is independent of the mesh size h_j, the parameters $\alpha_1, \beta_1, \beta_2$ and the function u, but may take different value in each occurrence.

Naturally, the next question is that if there are other parameter choices such that $P_h^\star$ can be reduced to a local projection. The following lemma addresses this issue.

Lemma 3.1 (The condition for reduction to a local projection) *If $\alpha_1^2 + \beta_1\beta_2 = \frac{1}{4}$, $P_h^\star$ is a local projection.*

Proof The definition (11a) provides $k - 1$ linearly independent equations for solving $P_h^\star u$ on each cell. If (11b)–(11c) can be locally decoupled, $P_h^\star$ is a local projection. We can write (11b)–(11c) in vector form as

$$\begin{bmatrix} \widehat{P_h^\star u} \\ (P_h^\star u)_x \end{bmatrix}\Bigg|_{x_{j+\frac{1}{2}}} = A_0 \begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Bigg|^-_{x_{j+\frac{1}{2}}} + B_0 \begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Bigg|^+_{x_{j+\frac{1}{2}}} = \begin{bmatrix} u\left(x_{j+\frac{1}{2}}\right) \\ u_x\left(x_{j+\frac{1}{2}}\right) \end{bmatrix}, \tag{15}$$

where

$$A_0 = \begin{bmatrix} \frac{1}{2} + \alpha_1 & -\beta_2 \\ -\beta_1 & \frac{1}{2} - \alpha_1 \end{bmatrix}, \quad B_0 = \begin{bmatrix} \frac{1}{2} - \alpha_1 & \beta_2 \\ \beta_1 & \frac{1}{2} + \alpha_1 \end{bmatrix}. \tag{16}$$

Note that $A_0 + B_0 = I_2$, $\det A_0 = \det B_0 = -(\alpha_1^2 + \beta_1\beta_2 - \frac{1}{4})$ and $A_0 B_0 = -(\det A_0) I_2$.

By assumption $\alpha_1^2 + \beta_1\beta_2 = \frac{1}{4}$, if $\beta_1 = \beta_2 = 0$, then $\alpha_1 = \pm\frac{1}{2}$ and $P_h^\star u = P_h^1$ or P_h^2, which are local projections. The rest of the cases are

- if $\beta_1 \neq 0$, left multiply (15) by a matrix, we have

$$\begin{bmatrix} \beta_1 & \frac{1}{2} + \alpha_1 \\ \beta_1 & -\left(\frac{1}{2} - \alpha_1\right) \end{bmatrix} \begin{bmatrix} u \\ u_x \end{bmatrix}\Bigg|_{x_{j+\frac{1}{2}}} = \begin{bmatrix} 0 & 0 \\ \beta_1 & -\left(\frac{1}{2} - \alpha_1\right) \end{bmatrix} \begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Bigg|^-_{x_{j+\frac{1}{2}}} + \begin{bmatrix} \beta_1 & \frac{1}{2} + \alpha_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Bigg|^+_{x_{j+\frac{1}{2}}},$$

which implies the following decoupled relations

$$\begin{aligned}(P_h^\star u)^+ + \frac{\frac{1}{2}+\alpha_1}{\beta_1}(P_h^\star u)_x^+ &= u + \frac{\frac{1}{2}+\alpha_1}{\beta_1}u_x \quad \text{at } x_{j-\frac{1}{2}},\\ (P_h^\star u)^- - \frac{\frac{1}{2}-\alpha_1}{\beta_1}(P_h^\star u)_x^- &= u - \frac{\frac{1}{2}-\alpha_1}{\beta_1}u_x \quad \text{at } x_{j+\frac{1}{2}}.\end{aligned} \tag{17}$$

- if $\beta_2 \neq 0$, by similar linear transformation, we have

$$\begin{aligned}\begin{bmatrix} \frac{1}{2}-\alpha_1 & \beta_2 \\ -\left(\frac{1}{2}+\alpha_1\right) & \beta_2 \end{bmatrix}\begin{bmatrix} u \\ u_x \end{bmatrix}\Big|_{x_{j+\frac{1}{2}}} &= \begin{bmatrix} 0 & 0 \\ -\left(\frac{1}{2}+\alpha_1\right) & \beta_2 \end{bmatrix}\begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Big|^-_{x_{j+\frac{1}{2}}}\\ &\quad + \begin{bmatrix} \frac{1}{2}-\alpha_1 & \beta_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} P_h^\star u \\ (P_h^\star u)_x \end{bmatrix}\Big|^+_{x_{j+\frac{1}{2}}},\end{aligned}$$

which implies

$$\begin{aligned}(P_h^\star u)_x^+ + \frac{\frac{1}{2}-\alpha_1}{\beta_2}(P_h^\star u)^+ &= u_x + \frac{\frac{1}{2}-\alpha_1}{\beta_2}u \quad \text{at } x_{j-\frac{1}{2}},\\ (P_h^\star u)_x^- - \frac{\frac{1}{2}+\alpha_1}{\beta_2}(P_h^\star u)^- &= u_x - \frac{\frac{1}{2}+\alpha_1}{\beta_2}u \quad \text{at } x_{j+\frac{1}{2}}.\end{aligned} \tag{18}$$

(17), (18) are the desired decoupled conditions on each cell I_j. Therefore the proof is complete. □

This lemma implies that for any parameter satisfying $\alpha_1^2+\beta_1\beta_2 = \frac{1}{4}$, $P_h^\star$ is locally defined. We remark that this condition turns out to be the same as the optimally convergent numerical flux families in [6] for two-way wave equations, although they arise in different contexts. Unfortunately, for the general definition of $P_h^\star$, unlike P_h^1 and P_h^2, we cannot directly use the Bramble–Hilbert lemma and the standard scaling argument to obtain optimal approximation property, since the second and third relations in (17) and (18) may break the scaling.

The next lemma performs a detailed analysis of this local projection when $\beta_1 \neq 0$ or $\beta_2 \neq 0$. Indeed for some parameter choices, only suboptimal convergence rate is obtained.

Lemma 3.2 (Local projection: existence, uniqueness and error estimates) *If $\alpha_1^2+\beta_1\beta_2 = \frac{1}{4}$ with $\beta_1 \neq 0$ or $\beta_2 \neq 0$, the local projection $P_h^\star$ exists and is uniquely defined when*

$$\Gamma_j = \beta_1 - \frac{k^2}{h_j} + \beta_2\frac{k^2(k^2-1)}{h_j^2} \neq 0, \quad \forall j. \tag{19}$$

In addition, the following error estimates hold for $p = 2, \infty$:

$$\|P_h^\star u - u\|_{L^p(I)} \leq Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\left(1 + \frac{\max\left(|\beta_1|, \min\left(\frac{\left|\frac{1}{2}-\alpha_1\right|}{h}, \frac{\left|\frac{1}{2}+\alpha_1\right|}{h}\right), \frac{|\beta_2|}{h^2}\right)}{\min_j|\Gamma_j|}\right). \tag{20}$$

Proof The proof of this lemma can be found in the "Appendix A.1". □

If we assume $\beta_1 = c/h, \beta_2 = ch$, then $\alpha_1 = constant$, and as long as the solvability condition (19) is satisfied, we have the optimal approximation property for $P_h^\star$. Such conclusions are not surprising, because (17) and (18) will maintain the correct scaling relation. However, for other parameter choices, the convergence rate in (20) may be suboptimal. The estimate (20) is verified by numerical experiments in Tables 1, 2, 3 and 4.

3.1.2 Global Projection Results

In this subsection, we consider $\alpha_1^2+\beta_1\beta_2 \neq \frac{1}{4}$, where $P_h^\star$ is a global projection. For simplicity, only uniform mesh is investigated, which makes the coefficient matrix of the linear system block-circulant. First, we analyze the existence and uniqueness of $P_h^\star$.

Lemma 3.3 (Global projection: existence and uniqueness) *If $\alpha_1^2 + \beta_1\beta_2 \neq \frac{1}{4}$, assuming a uniform mesh of size h, let $\Gamma := \beta_1 + \frac{k^2(k^2-1)}{h^2}\beta_2 - \frac{2k^2}{h}(\alpha_1^2 + \beta_1\beta_2 + \frac{1}{4})$ and $\Lambda := \frac{-2k}{h}(\alpha_1^2 + \beta_1\beta_2 - \frac{1}{4})$, then we have*

Case 1. if $|\Gamma| > |\Lambda|$, then $P_h^\star$ exists and is uniquely defined.

Case 2. if $|\Gamma| = |\Lambda|$, then $P_h^\star$ exists and is uniquely defined if N is odd, and furthermore, if k is odd, we require $\Gamma = -\Lambda$; if k is even, we require $\Gamma = \Lambda$.

Case 3. if $|\Gamma| < |\Lambda|$, then $P_h^\star$ exists and is uniquely defined if

$$(-1)^{(k+1)N}\left(\frac{\Gamma}{\Lambda} + \sqrt{\left(\frac{\Gamma}{\Lambda}\right)^2 - 1}\right)^N \neq 1.$$

Proof The proof of this lemma can be found in the "Appendix A.2". □

Next, we will focus on error estimates of the projection $P_h^\star$ based on the three cases as categorized in Lemma 3.3.

Lemma 3.4 (Global projection: error estimates for Case 1) *When the parameter choice belongs to Case 1 in Lemma 3.3, we have for $p = 2, \infty$,*

$$\begin{aligned}
&\|P_h^\star u - u\|_{L^p(I)} \\
&\leq Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\left(1 + \left(\frac{|\lambda_2|+1}{|\lambda_2|-1}\left(\|Q_1V_1\|_\infty + h^{-1}\|Q_1V_2\|_\infty\right)\right.\right. \\
&\quad \left.\left. + \frac{1}{|\lambda_2|-1}\left(\|V_1\|_\infty + h^{-1}\|V_2\|_\infty\right)\right)\right), \quad \text{if } \Gamma < 0, \\
&\|P_h^\star u - u\|_{L^p(I)} \\
&\leq Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\left(1 + \left(\frac{|\lambda_1|+1}{|\lambda_1|-1}\left(\|(I_2-Q_1)V_1\|_\infty + h^{-1}\|(I_2-Q_1)V_2\|_\infty\right)\right.\right. \\
&\quad \left.\left. + \frac{1}{|\lambda_1|-1}\left(\|V_1\|_\infty + h^{-1}\|V_2\|_\infty\right)\right)\right), \quad \text{if } \Gamma > 0,
\end{aligned} \tag{21}$$

where Q_1 is given by (60) or (61) depending on the parameter choices as shown in the proof; I_2 is the 2×2 identity matrix; V_1, V_2 are given by (65); and λ_1, λ_2 are the eigenvalues of Q as defined in (52).

Proof The proof of this lemma can be found in the "Appendix A.3". □

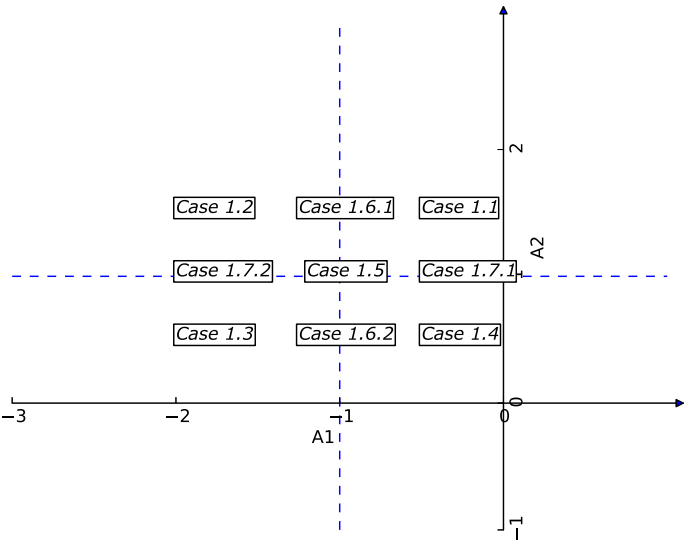

Fig. 1 A sketch to illustrate the different cases parameterized by the values of A_1, A_2

Equation (21) provides error bound that can be computed once the parameters $\alpha_1, \beta_1, \beta_2$ are given, yet its dependence on the mesh size h is not fully revealed, particularly when the parameters $\alpha_1, \beta_1, \beta_2$ also have h-dependence. To clarify such relations, next we will consider the following common choice of parameters, where α_1 has no dependence on h, $\beta_1 = \tilde{\beta}_1 h^{A_1}$, $\beta_2 = \tilde{\beta}_2 h^{A_2}$, $\tilde{\beta}_1, \tilde{\beta}_2$ are nonzero constants that do not depend on h. If indeed β_1 or β_2 is zero, it is equivalent to let $A_1, A_2 \to +\infty$ in the discussions below. We will discuss if the parameter choice will yield optimal $(k+1)$-th order accuracy. To distinguish different cases, we illustrate the choice of parameters A_1, A_2 in Fig. 1. For example, Case 1.1 means $A_1 > -1, A_2 > 1$, Case 1.5 means $A_1 = -1, A_2 = 1$ and Case 1.7.1 means $A_1 > -1, A_2 = 1$. The main results are summarized in Algorithm 1.

Algorithm 1: Interpretation of error estimate (21).

1 **if** $k = 1$ *and* $A_2 < 1$, **then**
2 | $P_h^\star$ is suboptimal and is $(k + A_2)$-th order accurate,
3 **else**
4 | **if** $\lim_{h\to 0} |\lambda_1, \lambda_2| = 1$ *with* $|\lambda_1, \lambda_2| = 1 + O(h^{\delta/2})$, **then**
5 | | $P_h^\star$ is suboptimal and is $(k + 1 - \delta)$-th order accurate,
6 | **else**
7 | | $P_h^\star$ has optimal $(k+1)$-th order error estimates.
8 | **end**
9 **end**

The main reason of order reduction for $k = 1, A_2 < 1$ in Statement 2 (i.e. line 2 of the algorithm above) is that the term such as $\frac{1}{|\lambda_2|-1}\|Q_1 V_1\|_\infty$ is of $O(h^{A_2-1})$ instead of $O(1)$, and this will cause $(1 - A_2)$-th order reduction. This happens for Cases 1.3, 1.4 and 1.6.2 when $k = 1$.

The main reason of order reduction in Statement 5 is because of the terms such as $\frac{1}{|\lambda_2|-1}$, $\frac{|\lambda_2|+1}{|\lambda_2|-1}$ in (21). The fractions $\frac{1}{|\lambda_2|-1}$, $\frac{|\lambda_2|+1}{|\lambda_2|-1}$ cannot be controlled by a constant if

$\lim_{h\to 0}|\lambda_2| = 1$. By definition of λ_1, λ_2 in (52), we know that $\left|\frac{\Gamma}{\Lambda}\right| \to 1 \Leftrightarrow |\lambda_1, \lambda_2| \to 1$. More precisely, if $\left|\frac{\Gamma}{\Lambda}\right| = 1 + O(h^\delta)$, $\delta > 0$, then $|\lambda_1, \lambda_2| = 1 + O(h^{\delta/2})$, then $\frac{|\lambda_2|+1}{|\lambda_2|-1}$ or $\frac{1}{|\lambda_2|-1} = O(h^{-\delta/2})$. The relation $\Gamma^2 - \Lambda^2 = (b_1 - b_2)(b_1 + b_2) + c_2^2$ also indicates that there is some cancellation of leading terms in $b_1 - b_2$ or $b_1 + b_2$, making $\|Q_1\|_\infty \sim O(h^{-\delta/2})$, multiplying these factors together will result in δ-th order reduction in the error estimation of $P_h^\star$. Note that b_1, b_2, c_2 and Q_1 are defined in (50), (51), (48) and (60).

Then we look at what parameter choices make $\left|\frac{\Gamma}{\Lambda}\right| \to 1$. Since

$$\frac{\Gamma}{\Lambda} = \begin{cases} k + \frac{\beta_1 + \frac{k^2(k^2-1)}{h^2}\beta_2 - \frac{k^2}{h}}{\Lambda} & k > 1, \\ 1 + \frac{\beta_1 - \frac{1}{h}}{\Lambda} & k = 1, \end{cases}$$

we have

1. Case 1.1 ($A_1 > -1, A_2 > 1$) with $k = 1, \alpha_1 = 0, \left|\frac{\Gamma}{\Lambda}\right| \to \left|\frac{\frac{1}{2}+2\alpha_1^2}{\frac{1}{2}-2\alpha_1^2}\right| = 1.$
2. Case 1.6.1 ($A_1 = -1, A_2 > 1$) $\tilde{\beta}_1 = \frac{k(k\pm 1)}{2} + 2\alpha_1^2 k(k \mp 1), \left|\frac{\Gamma}{\Lambda}\right| \to \left|k + \frac{\beta_1 - \frac{k^2}{h}}{\Lambda}\right| \to 1.$
3. Case 1.6.2 ($A_1 = -1, A_2 < 1$) with $k > 1$, $\tilde{\beta}_1 = \frac{k(k\pm 1)}{2}, \left|\frac{\Gamma}{\Lambda}\right| \to \left|k + \frac{\frac{k^2(k^2-1)\beta_2}{h^2}}{\Lambda}\right| \to 1.$
4. Case 1.7.1 ($A_1 > -1, A_2 = 1$) $\tilde{\beta}_2 = \frac{1}{2k(k\mp 1)} + \frac{2\alpha_1^2}{k(k\pm 1)}, \left|\frac{\Gamma}{\Lambda}\right| \to \left|k + \frac{\frac{k^2(k^2-1)\beta_2}{h^2} - \frac{k^2}{h}}{\Lambda}\right| \to 1.$
5. Case 1.7.2 ($A_1 < -1, A_2 = 1$) $\tilde{\beta}_2 = \frac{1}{2k(k\pm 1)}, \left|\frac{\Gamma}{\Lambda}\right| \to \left|k + \frac{\beta_1}{\Lambda}\right| \to 1.$

Remark 3.1 We only considered T given by (54) in the discussion above. By "Appendix A.6", we can conclude that under the parameter conditions in Case 1, $(b_1 + b_2)(b_1 - b_2) = 0$ only can happen if $A_1 = -1, A_2 = 1$ with (85) or (86). This is Case 1.5, for which we always have optimal error estimate.

Remark 3.2 Through numerical tests, we found that (21) is mostly sharp with two exceptions. When $\lim_{h\to 0}|\lambda_1, \lambda_2| = 1$, the estimates show that there will be order reduction for error of $P_h^\star$, while in numerical experiments (see e.g. Tables 7, 8), such order reduction is observed only when $\lim_{h\to 0}\lambda_1, \lambda_2 = 1$ but not -1. We believe when $\lim_{h\to 0}\lambda_1, \lambda_2 = -1$, a refined estimate can be obtained similar to Lemma 3.8 for Case 2. We have not carried out this estimate in this work.

Another example we find for which (21) is not sharp is $k = 2, A_1 = -2, -3, A_2 = 1, (\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, -1, \frac{1}{12})$, where parameters belong to Case 1.7.2, $\tilde{\beta}_2 = \frac{1}{2k(k+1)}$ and $\lambda_1, \lambda_2 \to 1 + O(h^{-(1+A_1)/2})$. The theoretical results predict accuracy order of $(k + 2 + A_1)$ but numerical experiments in Table 9 show the order to be $(k + 3 + A_1)$. Our estimations can't resolve this one order difference. This special parameter may trigger a cancellation we didn't capture in analysis. We will improve this estimate in our future work.

We can then generalize the approach to Cases 2 and 3.

Lemma 3.5 (Global projection: error estimates for Case 2) *When the parameter choice belongs to Case 2 in Lemma 3.3 and $P_h^\star$ is well defined, we have*

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} |u|_{W^{k+1,\infty}(I)} \left(1 + h^{-1}\left(1 + \frac{h^{-1}\|Q_2\|_\infty}{|\Gamma|}\right)\left(\|V_1\|_\infty + h^{-1}\|V_2\|_\infty\right)\right), \tag{22}$$

where $p = 2, \infty$, Q_2 is given by (72) and V_1, V_2 are given by (65).

Proof The proof of this lemma can be found in the "Appendix A.3". □

Remark 3.3 Detailed discussions on the parameter choices for Case 2 are contained in "Appendix A.7". Under these conditions, we actually have $\Gamma = C\left(\beta_1 - \frac{k^2}{h} + \frac{k^2(k^2-1)}{h^2}\beta_2\right)$, and by (66)

$$\|V_1\|_\infty + h^{-1}\|V_2\|_\infty \sim C\left(1 + \frac{\max(|\beta_1|, \left|\frac{1}{2} - \alpha_1\right|/h)}{|\Gamma|}\right), \tag{23}$$

in addition

$$\frac{\|Q_2\|_\infty}{|\Gamma|} \sim C\frac{\max\left(|\beta_1|, \frac{1}{h}, \frac{|\beta_2|}{h^2}\right)}{|\Gamma|}. \tag{24}$$

In the best-case scenario, the right hand side of the two equations above are bounded by a constant. Therefore, (22) yields the accuracy order to be $(k-1)$ at best.

Lemma 3.6 (Global projection: error estimates for Case 3) *When the parameter choice belongs to Case 3 in Lemma 3.3 and $P_h^\star$ is well defined, assuming $\left|1 - \lambda_1^N\right| \sim O(h^{\delta'})$, we have*

$$\|P_h^\star u - u\|_{L^p(I)} \le Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\left(1 + h^{-(\delta'+1)}\|Q_1\|_\infty\left(\|V_1\|_\infty + h^{-1}\|V_2\|_\infty\right)\right) \tag{25}$$

where $p = 2, \infty$ and Q_1, V_1, V_2 are given by (60) and (65).

Proof The proof of this lemma can be found in the "Appendix A.3". □

Remark 3.4 In the best-case scenario, the term $\|Q_1\|_\infty$ and $\|V_1\|_\infty + h^{-1}\|V_2\|_\infty$ are bounded by constants. While the term $h^{-(\delta'+1)}$ is of order at least h^{-1}, leading to loss of at least one order of accuracy.

Lemmas 3.5 and 3.6 only give suboptimal results. In what follows, we aim at improving the convergence order with stronger assumption on the regularity of the solution by using additional techniques involving cancellation of errors from neighboring terms and global approximation by Fourier expansions. We will need the following lemma that resembles Proposition 3.2 in [3], and also the fast decay property of Fourier coefficients of the exact solution. The proof of Lemma 3.7 follows the same line as in [3] and is skipped for brevity.

Lemma 3.7 (Detailed error estimates for P_h^1) *When P_h^1 is applied to a periodic and sufficiently smooth function u on uniform mesh, denote $\eta_j = (u - P_h^1 u)^+|_{j+\frac{1}{2}}$ and $\theta_j = (u_x - (P_h^1 u)_x)^-|_{j+\frac{1}{2}}$, $j = 0, \ldots, N-1$, we have:*

$$\eta_{j-1} = \mu h^{k+1}u^{(k+1)}\left(x_{j-\frac{1}{2}}\right) + \mu_2 h^{k+2}u^{(k+2)}\left(x_{j-\frac{1}{2}}\right) + C_2 h^{k+3}, \tag{26}$$

$$\theta_j = \rho h^k u^{(k+1)}\left(x_{j-\frac{1}{2}}\right) + \rho_2 h^{k+1}u^{(k+2)}\left(x_{j-\frac{1}{2}}\right) + C_3 h^{k+2}, \tag{27}$$

where μ, μ_2, ρ and ρ_2 are constants that depend only on k. C_2 and C_3 depend on k and $|u|_{W^{k+3,\infty}(I_j)}$. Thus, by using Mean-Value Theorem, an additional h can be extracted,

$$|\eta_j - \eta_{j+1}| \le Ch^{k+2}|u|_{W^{k+2,\infty}(I)}, \tag{28}$$

$$|\theta_j - \theta_{j+1}| \le Ch^{k+1}|u|_{W^{k+2,\infty}(I)}. \tag{29}$$

With Lemma 3.7 and Fourier analysis, we can prove the following two lemmas with refined error estimates.

Lemma 3.8 (Global projection: refined error estimates for Case 2) *When the parameter choice belongs to Case 2 in Lemma 3.3 and $P_h^\star$ is well defined, we have*

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} \|u\|_{W^{k+4,\infty}(I)} \left(1 + \left(1 + \frac{\|Q_2\|_\infty}{|\Gamma|}\right)(\|V_1\|_\infty + h^{-1}\|V_2\|_\infty)\right), \tag{30}$$

where $p = 2, \infty$, Q_2 is given by (72), V_1, V_2 are given by (65).

Proof The proof of this lemma can be found in the "Appendix A.4". □

Remark 3.5 The difference between (30) and (22) are the two h^{-1} factors and the norm of u, which corresponds to the different regularity requirement for the estimation. It is obvious that (30) is always a better estimate if the solution is smooth enough.

In most cases, (30) yields optimal accuracy order, except when $k = 1, \alpha_1 = 0, \beta_1 = 0, \beta_2 = O(h^{A_2})$, $A_2 < 1$, where the $P_h^\star$ is only $(k + A_2)$-th order accurate because $\frac{\|Q_2\|_\infty}{|\Lambda|} = \frac{|b_1+b_2|}{|\Lambda|} = \frac{|-\frac{-4}{h^2}\beta_2 + \frac{1}{2h}|}{\frac{1}{2h}} \sim O(h^{A_2-1})$ in (30). This is verified numerically in Table 11.

Lemma 3.9 (Global projection: refined error estimates for Case 3) *When the parameter choice belongs to Case 3 in Lemma 3.3 and $P_h^\star$ is well defined, assuming $|1 - \lambda_1^N| = O(h^{\delta'})$ and $|\lambda_1 - 1| = O(h^{\delta/2})$ with $0 \le \delta/2 \le 1$, we have*

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} \|u\|_{W^{k+3,\infty}(I)} \left(1 + h^{-(\delta'+\delta/2)} \|Q_1\|_\infty (\|V_1\|_\infty + h^{-1}\|V_2\|_\infty)\right), \tag{31}$$

where $p = 2, \infty$, λ_1 is the eigenvalue of Q defined in (52), Q_1 is given by (60), V_1, V_2 are given by (65).

Proof The proof of this lemma can be found in the "Appendix A.5". □

Remark 3.6 If $0 \le \delta/2 \le 1$, Lemma 3.9 is always a better estimate than Lemma 3.6 when the solution is smooth enough. If $\delta/2 > 1$, we can show $\delta/2 = \delta' + 1$. This is because $|1 - \lambda_1| = |1 - e^{i\theta}| = 2|\sin(\theta/2)|$, and $|1 - \lambda_1^N| = |1 - e^{iN\theta}| = 2|\sin(N\theta/2)|$. When $\delta/2 > 1$, one can assert that $|1 - \lambda_1| \sim \theta$, $|1 - \lambda_1^N| \sim N\theta$, i.e. $\delta/2 = \delta' + 1$. With this condition, we notice that Lemma 3.6 yields an reduction of δ-th order in convergence rate by checking the order of each term as is done for Case 1. This order reduction is consistent with numerical experiments in Example 4.4. Therefore, there is no need to further improve the estimates as is done for $0 \le \delta/2 \le 1$ in Lemma 3.9.

Now we can summarize the estimation of $P_h^\star$ for some frequently used flux parameters. For IPDG scheme with $\alpha_1 = \beta_2 = 0, \beta_1 = c/h$, and DDG scheme discussed in [21] with $\alpha_1 = constant, \beta_1 = c/h, \beta_2 = 0$, and the more general scale invariant parameter choice $\alpha_1 = constant, \beta_1 = c/h, \beta_2 = ch$, $P_h^\star$ always have optimal error estimates. For those parameters, we can show that the eigenvalues λ_1, λ_2 are always constants independent of h, therefore, either by estimates for local projection in Lemma 3.2 or global projection in Lemmas 3.4, 3.8 and 3.9, we will have optimal convergence rate. Corresponding numerical results are shown in Tables 2 and 6.

For a natural parameter choice where $\alpha_1, \beta_1, \beta_2$ are all real constants, if $\beta_2 \neq 0$, then $P_h^\star$ has first order convergence rate when $k = 1$ and optimal convergence rate when $k > 1$ by

Lemmas 3.2, 3.4, 3.8, 3.9. Corresponding numerical results are shown in Tables 1 and 11. Lastly, for central flux $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$, this parameter choice belongs to Case 2 when $k = 1$ and Case 1 when $k > 1$, thus we can verify that $P_h^\star$ has optimal convergence rate by Lemmas 3.4 and 3.8. Corresponding numerical results are shown in Table 10.

3.2 Error Estimates of the DG Scheme

We are now ready to state the main theorem, which is the semi-discrete L^2 error estimates of the DG scheme (2) with numerical flux (10).

Theorem 3.10 *Assume that the exact solution u and the nonlinear term $f(|u|^2)$ of (1) are sufficiently smooth with bounded derivatives for any time $t \in (0, T_e]$ and that the numerical flux parameters in (10) satisfy the existence conditions of $P_h^\star$ in Lemmas 3.2 or 3.3. Furthermore, assume $\epsilon_h = u - P_h^\star u$ has at least first order convergence rate in L^2 and L^∞ norm from the results in Sect. 3.1. With periodic boundary conditions solution space V_h^k ($k \geq 1$), the following error estimation holds for u_h, which is the numerical solution of (2) with flux (10):*

$$\|u - u_h\|_{L^2(I)} \leq C_\star \left(\|(u - u_h)|_{t=0}\|_{L^2(I)} + \|(\epsilon_h)_t\|_{L^2(I)} + \|\epsilon_h\|_{L^2(I)} \right), \tag{32}$$

where $C_\star$ depends on k, $\|f\|_{W^{2,\infty}}$, u as well as final time T_e, but not on h.

Moreover, the estimates for $\|\epsilon_h\|_{L^2(I)}$ has been established in Lemma 3.2 on non-uniform mesh and in Lemma 3.4–3.9 on uniform mesh for corresponding flux parameters. In other words, the error of the DG scheme (2) has the same accuracy as $P_h^\star u$, as long as $P_h^\star u$ is well-defined and the numerical initial condition is chosen sufficiently accurate.

Proof When $P_h^\star$ exists, we can decompose the error into two parts.

$$e = u - u_h = u - P_h^\star u + P_h^\star u - u_h := \epsilon_h + \zeta_h.$$

By Galerkin orthogonality

$$\begin{aligned} 0 &= a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(e, v_h) - i \int_I f(|u|^2) u v_h dx + i \int_I f(|u_h|^2) u_h v_h dx \qquad \forall v_h \in V_h^k \\ &= a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\epsilon_h, v_h) + a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\zeta_h, v_h) - i \int_I f(|u|^2) u v_h dx \\ &\quad + i \int_I f(|u_h|^2) u_h v_h dx. \end{aligned}$$

By letting $v_h = \overline{\zeta_h}$ and taking conjugate of above equation, we have

$$\begin{aligned} &a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\zeta_h, \overline{\zeta_h}) + \overline{a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\zeta_h, \overline{\zeta_h})} \\ &\quad = -a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\epsilon_h, \overline{\zeta_h}) - \overline{a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\epsilon_h, \overline{\zeta_h})} - 2 \int_I f(|u|^2) \mathrm{Im}(u \overline{\zeta_h}) dx \\ &\qquad + 2 \int_I f(|u_h|^2) \mathrm{Im}(u_h \overline{\zeta_h}) dx. \end{aligned} \tag{33}$$

By Taylor expansion

$$f(|u_h|^2) = f(|u|^2) + f'(|u|^2) E + \frac{1}{2} \hat{f}'' E^2,$$

where $\hat{f}'' = f''(c)$, c is a value between $|u_h|^2$ and $|u|^2$. $E = |u_h|^2 - |u|^2 = -2\mathrm{Re}(e\overline{u}) + |e|^2$. Therefore, the nonlinear part becomes

$$\begin{aligned}
&\int_I f(|u|^2)\mathrm{Im}(u\overline{\zeta_h})dx - \int_I f(|u_h|^2)\mathrm{Im}(u_h\overline{\zeta_h})dx \\
&= \int_I f(|u_h|^2)\mathrm{Im}\big(e\overline{\zeta_h}\big) + \big(f(|u|^2) - f(|u_h|^2)\big)\mathrm{Im}(u\overline{\zeta_h})dx \\
&= \mathcal{N}_1 + \mathcal{N}_2 + \mathcal{N}_3,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{N}_1 &= \int_I f(|u|^2)\mathrm{Im}\big(e\overline{\zeta_h}\big) - f'(|u|^2)E\mathrm{Im}(u\overline{\zeta_h})dx, \\
\mathcal{N}_2 &= \int_I f'(|u|^2)E\mathrm{Im}\big(e\overline{\zeta_h}\big) - \frac{1}{2}\hat{f}''E^2\mathrm{Im}(u\overline{\zeta_h})dx, \\
\mathcal{N}_3 &= \int_I \frac{1}{2}\hat{f}''E^2\mathrm{Im}\big(e\overline{\zeta_h}\big)dx,
\end{aligned}$$

will be estimated separately as follows.

- $\mathcal{N}_1$ and $\mathcal{N}_2$ terms.
  Since $e\overline{\zeta_h} = \epsilon_h\overline{\zeta_h} + |\zeta_h|^2$, $\big|E\mathrm{Im}(u\overline{\zeta_h})\big| = \big|(-2\mathrm{Re}(e\overline{u}) + |e|^2)\mathrm{Im}(u\overline{\zeta_h})\big| \le C(\|u\|^2_{L^\infty(I)} + \|u\|_{L^\infty(I)}\|e\|_{L^\infty(I)})(|\epsilon_h|^2 + |\zeta_h|^2)$, we have

$$\begin{aligned}
|\mathcal{N}_1| &\le C\|f\|_{W^{1,\infty}}\left(1 + \|u\|^2_{L^\infty(I)} + \|u\|_{L^\infty(I)}\|e\|_{L^\infty(I)}\right)(\|\epsilon_h\|^2_{L^2(I)} + \|\zeta_h\|^2_{L^2(I)}), \\
|\mathcal{N}_2| &\le C\|f\|_{W^{2,\infty}}\|E\|_{L^\infty(I)}\left(1 + \|u\|^2_{L^\infty(I)} + \|u\|_{L^\infty(I)}\|e\|_{L^\infty(I)}\right)(\|\epsilon_h\|^2_{L^2(I)} + \|\zeta_h\|^2_{L^2(I)}).
\end{aligned}$$

- $\mathcal{N}_3$ term.

$$|\mathcal{N}_3| \le C\|f''\|_{L^\infty}\|E\|^2_{L^\infty(I)}\left(\|\epsilon_h\|^2_{L^2(I)} + \|\zeta_h\|^2_{L^2(I)}\right).$$

To conduct a proper estimate for the nonlinear part, we would like to make an a priori assumption that, for h small enough,

$$\|e\|_{L^2(I)} = \|u - u_h\|_{L^2(I)} \le h^{0.5}. \tag{34}$$

By our assumption on $P_h^\star$, $\|\epsilon_h\|_{L^p(I)} \le C_1 h$, $p = 2, \infty$, thus $\|\zeta_h\|_{L^2(I)} \le C_1 h^{0.5}$ and $\|\zeta_h\|_{L^\infty(I)} \le C_1$ by inverse inequality, then $\|e\|_{L^\infty(I)} \le C_1$, $\|E\|_{L^\infty(I)} \le C_1$. Here and below, C_1 is a generic constant that has no dependence on h, but may depend on u according to the lemma used to estimate ϵ_h.

Therefore, we get the estimate:

$$|\mathcal{N}_1| + |\mathcal{N}_2| + |\mathcal{N}_3| \le C_1\left(\|\epsilon_h\|^2_{L^2(I)} + \|\zeta_h\|^2_{L^2(I)}\right), \tag{35}$$

where C_1 depends on $\|f\|_{W^{2,\infty}}$ and u.

For linear part of the right hand side in (33), we have

$$
\begin{aligned}
&a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\epsilon_h,\overline{\zeta_h})+\overline{a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\epsilon_h,\overline{\zeta_h})}=\int_I(\epsilon_h)_t\overline{\zeta_h}+\overline{(\epsilon_h)_t}\zeta_h dx-i\int_I(\epsilon_h)(\overline{\zeta_h})_{xx}dx\\
&\quad+i\int_I\overline{(\epsilon_h)}(\zeta_h)_{xx}dx-i\sum_j(\widehat{\epsilon_h}[(\overline{\zeta_h})_x]-\widetilde{(\epsilon_h)_x}[\overline{\zeta_h}])|_{j+\frac{1}{2}}\\
&\quad+i\sum_j\overline{(\widehat{\epsilon_h}[(\overline{\zeta_h})_x]-\widetilde{(\epsilon_h)_x}[\overline{\zeta_h}])|_{j+\frac{1}{2}}},\\
&=2\int_I\mathrm{Re}\big((\epsilon_h)_t\overline{\zeta_h}\big)dx.
\end{aligned}
$$

The last equality holds because of the definition of $P_h^\star u$. For the left hand side of (33), by similar computation in stability analysis we have

$$
a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\zeta_h,\overline{\zeta_h})+\overline{a_{\alpha_1,-\alpha_1,\beta_1,\beta_2}(\zeta_h,\overline{\zeta_h})}=\frac{d}{dt}\int_I|\zeta_h|^2dx. \tag{36}
$$

Combine these two equations with (35):

$$
\frac{d}{dt}\|\zeta_h\|^2_{L^2(I)}\le\|(\epsilon_h)_t\|^2_{L^2(I)}+\|\zeta_h\|^2_{L^2(I)}+C_1(\|\epsilon_h\|^2_{L^2(I)}+\|\zeta_h\|^2_{L^2(I)}).
$$

Assuming u_t, u have sufficient smoothness, then by Gronwall's inequality, we can get:

$$
\|\zeta_h\|^2_{L^2(I)}\le C_1\left(\|\zeta_h|_{t=0}\|^2_{L^2(I)}+\|(\epsilon_h)_t\|^2_{L^2(I)}+\|(\epsilon_h)\|^2_{L^2(I)}\right),
$$

and we obtain (32).

To complete the proof, we shall justify the a priori assumption. To be more precise, we consider h_0, s.t., $\forall h<h_0, C_\star h\le\frac{1}{2}h^{0.5}$, where $C_\star$ is defined in (32), dependent on T_e, but not on h. Suppose $\exists t^*=\sup\{t:\|u(t^*)-u_h(t^*)\|_{L^2(I)}\le h^{0.5}\}$, we would have $\|u(t^*)-u_h(t^*)\|_{L^2(I)}=h^{0.5}$ by continuity if t^* is finite. By (32), we obtain $\|e\|_{L^2(I)}\le C_\star h\le\frac{1}{2}h^{0.5}$ if $t^*\le T_e$, which contradicts the definition of t^*. Therefore, $t^*>T_e$ and the a priori assumption is justified. □

Remark 3.7 If f is a constant function, we can prove the same error estimates without using the a priori assumption. Therefore, the assumption that $\epsilon_h=u-P_h^\star u$ has at least first order convergence rate in L^2 and L^∞ norm is no longer needed.

4 Numerical Experiments

In this section, we present numerical experiments to validate our theoretical results. Particularly, in Sect. 4.1, we provide numerical validations of convergence rate for the projection $P_h^\star$ as discussed in Sect. 3.1 with focus on the dependence of the errors on parameters α_1, β_1, β_2. Section 4.2 illustrates the energy conservation property and validates theoretical convergence rate of DG scheme for NLS equation (1).

4.1 Numerical Results of the Projection Operator $P_h^\star$

Example 4.1 In this example, we focus on local projection where $\alpha_1^2+\beta_1\beta_2=\frac{1}{4}$, and verify the conclusions in Lemma 3.2 by considering a smooth test function $u=\cos(x)$, $x\in[0,2\pi]$

Table 1 Example 4.1

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	160	1.98E−02	–	1.56E−02	–	1.81E−02	–
	320	9.98E−03	0.99	7.87E−03	0.99	9.20E−03	0.97
	640	5.01E−03	0.99	3.95E−03	0.99	4.55E−03	1.02
	1280	2.51E−03	1.00	1.98E−03	1.00	2.27E−03	1.00
P^2	160	2.18E−06	–	1.91E−06	–	3.73E−06	–
	320	2.71E−07	3.01	2.39E−07	3.00	5.14E−07	2.86
	640	3.37E−08	3.01	2.97E−08	3.01	6.71E−08	2.94
	1280	4.19E−09	3.01	3.69E−09	3.01	7.99E−09	3.07
P^3	160	2.82E−09	–	2.45E−09	–	5.67E−09	–
	320	1.76E−10	4.00	1.53E−10	4.00	3.76E−10	3.92
	640	1.10E−11	4.00	9.50E−12	4.01	2.25E−11	4.06
	1280	6.86E−13	4.00	5.93E−13	4.00	1.46E−12	3.95

Error of local projection $P_h^\star u - u$ on a nonuniform mesh. Flux parameters: $\alpha_1 = 0.3, \beta_1 = 0.4, \beta_2 = 0.4$

on a nonuniform mesh and $k = 1, 2, 3$ for various sets of parameters $(\alpha_1, \beta_1, \beta_2)$. The nonuniform mesh is generated by perturbing the nodes of a uniform mesh of N cells by at most 10%.

We first consider two sets of parameters $(\alpha_1, \beta_1, \beta_2) = (0.3, 0.4, 0.4)$ and $(\alpha_1, \beta_1, \beta_2) = (0.3, 0.4/h_j, 0.4h_j)$. The results with $(\alpha_1, \beta_1, \beta_2) = (0.3, 0.4, 0.4)$ are listed in Table 1. By plugging in the parameters into (20), we have that when $k = 1$, the projection has suboptimal first order convergence rate, while for $k > 1$, optimal $(k + 1)$-th order convergence rate should be achieved. For $k = 1$, $\Gamma_j = \beta_1 - \frac{1}{h_j}$, which does not depend on β_2 any more. This technical difference cause the discrepancy of the convergence order between $k = 1$ and $k > 1$ in Table 1. Results in Table 1 agree well with the theoretical prediction. On the other hand, when we choose parameters $(\alpha_1, \beta_1, \beta_2) = (0.3, 0.4/h_j, 0.4h_j)$, by Lemma 3.2, we should observe optimal convergence rate for all $k \geq 1$, and this is verified by the numerical results in Table 2.

Then, we choose the parameters as $(\alpha_1, \beta_1, \beta_2) = (0.5, 1, 0)$ to verify the super-closeness claim (42), i.e., the difference between $P_h^\star$ and P_h^1 can have convergence rates higher than $k + 1$. The results are listed in Table 3. The difference of the two projections is indeed of $(k+2)$-th order for any $k \geq 1$ in all norms. Finally, we take $(\alpha_1, \beta_1, \beta_2) = (0.5, \frac{k^2}{h_j(1+h_j)}, 0)$. In this case, $\Gamma_j = O(1)$. The numerical results in Table 4 verify the order reduction to k-th order accuracy for all $k \geq 1$ as predicted by (20).

Example 4.2 In this example, we consider global projection when the parameter choices belong to Case 1. We consider a smooth test function $u = e^{\cos(x)}$ on $[0, 2\pi]$ with a uniform mesh of size $h = 2\pi/N$ and $k = 1, 2, 3$ for various sets of parameters $(\alpha_1, \beta_1, \beta_2)$.

We first test the situation when $\lim_{h\to 0} |\lambda_1, \lambda_2| \neq 1$ by setting the parameters $(\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, 1, 1), A_1 = -0.5, A_2 = 2$. Another example is $(\alpha_1, \beta_1, \beta_2) = (0, \frac{1}{2h}, h)$, for which the eigenvalues λ_1, λ_2 are constant dependent on k but not h. These two parameter choices belong to Case 1.1 and Case 1.5, respectively. The numerical results shown in Tables 5 and 6 verify the optimal $(k + 1)$-th order convergence rate predicted by Lemma 3.4.

Table 2 Example 4.1

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	160	3.42E−04	–	3.50E−04	–	8.62E−04	–
	320	8.55E−05	2.00	8.75E−05	2.00	2.21E−04	1.96
	640	2.14E−05	2.00	2.19E−05	2.00	5.45E−05	2.02
	1280	5.34E−06	2.00	5.47E−06	2.00	1.36E−05	2.00
P^2	160	6.36E−06	–	6.06E−06	–	2.06E−05	–
	320	8.17E−07	2.96	7.99E−07	2.92	3.09E−06	2.73
	640	1.02E−07	3.00	1.00E−07	2.99	4.51E−07	2.78
	1280	1.27E−08	3.01	1.24E−08	3.02	5.12E−08	3.14
P^3	160	3.32E−09	–	2.93E−09	–	7.58E−09	–
	320	2.08E−10	4.00	1.83E−10	4.00	5.08E−10	3.90
	640	1.30E−11	4.00	1.14E−11	4.01	3.04E−11	4.06
	1280	8.09E−13	4.00	7.12E−13	4.00	1.99E−12	3.93

Error of local projection $P_h^\star u - u$ on a nonuniform mesh. Flux parameters: $\alpha_1 = 0.3$, $\beta_1 = 0.4/h_j$, $\beta_2 = 0.4h_j$

Table 3 Example 4.1

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	160	2.09E−05	–	1.96E−05	–	4.66E−05	–
	320	2.56E−06	3.03	2.40E−06	3.03	5.99E−06	2.96
	640	3.17E−07	3.01	2.96E−07	3.02	7.17E−07	3.06
	1280	3.94E−08	3.01	3.67E−08	3.01	9.11E−08	2.98
P^2	160	5.00E−09	–	5.05E−09	–	1.82E−08	–
	320	3.14E−10	3.99	3.21E−10	3.98	1.28E−09	3.83
	640	1.96E−11	4.00	2.00E−11	4.00	8.56E−11	3.90
	1280	1.22E−12	4.01	1.24E−12	4.01	5.02E−12	4.09
P^3	160	2.91E−12	–	3.38E−12	–	1.40E−11	–
	320	9.11E−14	5.00	1.06E−13	5.00	4.72E−13	4.89
	640	2.84E−15	5.00	3.27E−15	5.01	1.40E−14	5.08
	1280	8.84E−17	5.00	1.02E−16	5.00	4.63E−16	4.92

Difference of local projection $P_h^\star$ with P_h^1: $P_h^\star u - P_h^1 u$ on a nonuniform mesh. Flux parameters: $\alpha_1 = 0.5$, $\beta_1 = 1$, $\beta_2 = 0$

Then we test the situation when $\lim_{h\to 0} |\lambda_1, \lambda_2| = 1$ by using two sets of parameters $(\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, \frac{k(k-1)}{2} + \frac{k(k+1)}{8}, 1)$, $A_1 = -1$, $A_2 = 2, 3$, and $(\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, \frac{2}{k(k-1)}, 1)$, $A_1 = -2, -3$, $A_2 = 1$. The first set of parameters belongs to Case 1.6.1 and we can verify that $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^k$. Lemma 3.4 and Algorithm 1 imply $(k+2-A_2)$th convergence order. The numerical results listed in Table 7 show that the expected order reduction only happens when $\lim_{h\to 0} \lambda_1, \lambda_2 = 1$, but not for $\lim_{h\to 0} \lambda_1, \lambda_2 = -1$. The second set of parameters belongs to Case 1.7.2 and we can verify that $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^{k+1}$. Lemma 3.4 and Algorithm 1 imply $(k+2+A_1)$th convergence order. The numerical results listed in Table 8 also show that order reduction is only observed when $\lim_{h\to 0} \lambda_1, \lambda_2 = 1$.

Table 4 Example 4.1

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	160	1.33E−02	–	1.20E−02	–	2.28E−02	–
	320	6.46E−03	1.04	5.87E−03	1.03	1.14E−02	1.00
	640	3.19E−03	1.02	2.90E−03	1.02	5.55E−03	1.03
	1280	1.58E−03	1.01	1.44E−03	1.01	2.76E−03	1.01
P^2	160	1.31E−05	–	1.28E−05	–	3.68E−05	–
	320	3.19E−06	2.04	3.13E−06	2.03	9.85E−06	1.90
	640	7.85E−07	2.02	7.71E−07	2.02	2.51E−06	1.97
	1280	1.95E−07	2.01	1.91E−07	2.01	6.06E−07	2.05
P^3	160	1.72E−08	–	1.89E−08	–	6.49E−08	–
	320	2.08E−09	3.05	2.28E−09	3.05	8.28E−09	2.97
	640	2.54E−10	3.03	2.80E−10	3.03	9.88E−10	3.07
	1280	3.14E−11	3.02	3.46E−11	3.01	1.25E−10	2.98

Error of local projection $P_h^\star u - u$ on a nonuniform mesh. Flux parameters: $\alpha_1 = 0.5$, $\beta_1 = \frac{k^2}{h_j(1+h_j)}$, $\beta_2 = 0$

Table 5 Example 4.2

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	160	0.10E−03	–	0.69E−03	–	0.89E−03	–
	320	0.26E−04	1.93	0.18E−03	1.93	0.23E−03	1.94
	640	0.67E−05	1.98	0.46E−04	1.97	0.58E−04	1.98
	1280	0.17E−05	1.99	0.12E−04	1.99	0.15E−04	2.00
P^2	160	0.63E−06	–	0.52E−05	–	0.87E−05	–
	320	0.88E−07	2.85	0.71E−06	2.88	0.11E−05	2.95
	640	0.11E−07	2.95	0.91E−07	2.97	0.14E−06	3.00
	1280	0.14E−08	2.99	0.11E−07	2.99	0.17E−07	3.01
P^3	320	0.64E−10	–	0.49E−09	–	0.72E−09	–
	640	0.45E−11	3.82	0.35E−10	3.80	0.52E−10	3.79
	1280	0.29E−12	3.93	0.23E−11	3.91	0.34E−11	3.92
	2560	0.19E−13	3.97	0.15E−12	3.96	0.22E−12	3.96

Error of global projection $P_h^\star u - u$. Flux parameters (Case 1.1): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = 1$, $\tilde{\beta}_2 = 1$, $A_1 = -0.5$, $A_2 = 2$

Lastly, we test $(\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, -1, \frac{1}{12})$ with $k = 2$, $A_1 = -2, -3$, $A_2 = 1$, where our theoretical results predict accuracy order of $(k + 2 + A_1)$, but numerical experiments show the order to be $(k + 3 + A_1)$ in Table 9. This is one of the exceptions that Lemma 3.4 is not sharp and has been commented in Remark 3.2.

Example 4.3 In this example, we consider global projection when the parameter choices are central-like fluxes belonging to Cases 1 and 2, for smooth function $u = e^{\cos(x)}$ on $[0, 2\pi]$ with a uniform mesh of size $h = 2\pi/N$ and $k = 1, 2, 3$.

For central flux $(\alpha_1, \beta_1, \beta_2) = (0, 0, 0)$, $\Gamma = -\frac{k^2}{2h}$, $\Lambda = \frac{k}{2h}$. If $k > 1$, $\frac{|\Gamma|}{|\Lambda|} = k > 1$, it belongs to Case 1, and if $k = 1$, $\Gamma = -\Lambda$ and it belongs to Case 2. We conclude that $P_h^\star$

Table 6 Example 4.2

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	320	0.11E−03	–	0.63E−03	–	0.38E−03	–
	640	0.28E−04	2.00	0.16E−03	2.00	0.95E−04	2.00
	1280	0.70E−05	2.00	0.39E−04	2.00	0.24E−04	2.00
	2560	0.18E−05	2.00	0.98E−05	2.00	0.60E−05	2.00
P^2	320	0.11E−06	–	0.71E−06	–	0.62E−06	–
	640	0.14E−07	3.00	0.89E−07	3.00	0.77E−07	3.00
	1280	0.18E−08	3.00	0.11E−07	3.00	0.96E−08	3.00
	2560	0.22E−09	3.00	0.14E−08	3.00	0.12E−08	3.00
P^3	320	0.38E−10	–	0.25E−09	–	0.22E−09	–
	640	0.24E−11	4.00	0.16E−10	4.00	0.14E−10	4.00
	1280	0.15E−12	4.00	0.99E−12	4.00	0.86E−12	4.00
	2560	0.92E−14	4.00	0.62E−13	4.00	0.54E−13	3.99

Error of global projection $P_h^\star u - u$. Flux parameters (Case 1.5): $\alpha_1 = 0, \beta_1 = \frac{1}{2h}, \beta_2 = h$

Table 7 Example 4.2

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	640	0.75E−05	–	0.52E−04	–	0.66E−04	–
$A_2 = 2$	1280	0.19E−05	1.97	0.13E−04	1.97	0.17E−04	1.97
$\tilde{\beta_1} = \frac{1}{4}$	2560	0.48E−06	1.99	0.34E−05	1.98	0.42E−05	1.99
	5120	0.12E−06	1.99	0.84E−06	1.99	0.11E−05	1.99
P^2	640	0.15E−06	–	0.12E−05		0.23E−05	–
$A_2 = 2$	1280	0.39E−07	1.94	0.32E−06	1.93	0.61E−06	1.94
$\tilde{\beta_1} = \frac{7}{4}$	2560	0.98E−08	1.97	0.82E−07	1.97	0.16E−06	1.97
	5120	0.25E−08	1.98	0.21E−07	1.98	0.39E−07	1.99
P^2	640	0.14E−04	–	0.12E−03	–	0.21E−03	–
$A_2 = 3$	1280	0.71E−05	1.00	0.58E−04	1.00	0.11E−03	1.00
$\tilde{\beta_1} = \frac{7}{4}$	2560	0.35E−05	1.00	0.29E−04	1.00	0.54E−04	1.00
	5120	0.18E−05	1.00	0.15E−04	1.00	0.27E−04	1.00
P^3	320	0.12E−09	–	0.95E−09	–	0.20E−08	–
$A_2 = 2$	640	0.78E−11	3.99	0.60E−10	3.99	0.13E−09	3.99
$\tilde{\beta_1} = \frac{9}{2}$	1280	0.49E−12	3.99	0.38E−11	3.99	0.80E−11	3.99
	2560	0.31E−13	4.00	0.24E−12	3.99	0.51E−12	3.97

Error of global projection $P_h^\star u - u$. Flux parameters (Case 1.6.1): $\alpha_1 = 0.25, \tilde{\beta}_1 = \frac{k(k-1)}{2} + \frac{k(k+1)}{8}, \tilde{\beta}_2 = 1.0, A_1 = -1, A_2 = 2, 3$. Note here $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^k$

Table 8 Example 4.2

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^2	320	0.28E−07	–	0.21E−06	–	0.24E−06	–
$A_1 = -3$	640	0.35E−08	3.00	0.27E−07	3.00	0.31E−07	3.00
$\tilde{\beta_2} = \frac{1}{4}$	1280	0.44E−09	3.00	0.33E−08	3.00	0.38E−08	3.00
	2560	0.55E−10	3.00	0.41E−09	3.00	0.48E−09	3.00
P^3	320	0.70E−08	–	0.57E−07	–	0.12E−06	–
$A_1 = -2$	640	0.94E−09	2.90	0.77E−08	2.90	0.16E−07	2.91
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.12E−09	2.95	0.99E−09	2.95	0.20E−08	2.95
	2560	0.15E−10	2.98	0.13E−09	2.98	0.26E−09	2.98
P^3	320	0.16E−06	–	0.13E−05	–	0.24E−05	–
$A_1 = -3$	640	0.40E−07	2.00	0.32E−06	2.00	0.61E−06	2.00
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.10E−07	2.00	0.79E−07	2.00	0.15E−06	2.00
	2560	0.25E−08	2.00	0.20E−07	2.00	0.38E−07	2.00

Error of global projection $P_h^\star u - u$. Flux parameters (Case 1.7.2): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = 1$, $\tilde{\beta}_2 = \frac{1}{2k(k-1)}$, $A_1 = -2, -3$, $A_2 = 1$. Note here $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^{k+1}$

Table 9 Example 4.2

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^2	320	0.72E−07	2.99	0.56E−06	2.98	0.94E−06	2.97
$A_1 = -2$	640	0.90E−08	2.99	0.71E−07	2.99	0.12E−06	2.99
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.11E−08	3.00	0.89E−08	3.00	0.15E−07	2.99
	2560	0.14E−09	3.00	0.11E−08	3.00	0.19E−08	3.00
P^2	320	0.80E−06	2.01	0.63E−05	2.01	0.12E−04	2.01
$A_1 = -3$	640	0.20E−06	2.00	0.16E−05	2.00	0.30E−05	2.00
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.50E−07	2.00	0.39E−06	2.00	0.75E−06	2.00
	2560	0.13E−07	2.00	0.98E−07	2.00	0.19E−06	2.00

Error of global projection $P_h^\star u - u$. Flux parameters (Case 1.7.2): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = -1$, $\tilde{\beta}_2 = \frac{1}{2k(k+1)}$, $A_1 = -2, -3$, $A_2 = 1$. Note that $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^k = 1$

exists and is unique for $k = 1$ when N is odd and $k > 1$ for arbitrary N. $P_h^\star$ has optimal error estimates as proved in Lemmas 3.4 and 3.8. Our numerical test in Table 10 demonstrates optimal convergence rate for all k.

A similar flux is $(\alpha_1, \beta_1, \beta_2) = (0, 0, 1)$. Lemma 3.8 yields first order convergence rate when $k = 1$ as discussed in Remark 3.5. When $k = 2, 3$, similar to central flux, this parameter choice belongs to Case 1, showing optimal convergence rate. The numerical test in Table 11 verifies the theoretical results.

Example 4.4 In this example, we consider global projection when the parameter choices belong to Case 3 for the smooth function $u = e^{\cos(x)}$ on $[0, 2\pi]$ with uniform mesh size $h = 2\pi/N$ and $k = 1, 2, 3$.

An example of Case 3 is shown in Table 12, where the parameters are $(\alpha_1, \tilde{\beta}_1, \tilde{\beta}_2) = (0.25, -1, \frac{1}{2k(k-1)})$, $A_1 = -2, -3$, $A_2 = 1$, similar to the parameters in Table 8. The

Table 10 Example 4.3

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	93	0.12E−03	–	0.74E−03	–	0.55E−03	–
	279	0.13E−04	2.00	0.82E−04	2.00	0.61E−04	2.00
	837	0.15E−05	2.00	0.91E−05	2.00	0.68E−05	2.00
	2511	0.17E−06	2.00	0.10E−05	2.00	0.76E−06	2.00
P^2	160	0.11E−05	–	0.85E−05	–	0.10E−04	–
	320	0.14E−06	3.00	0.11E−05	3.00	0.13E−05	2.99
	640	0.17E−07	3.00	0.13E−06	3.00	0.16E−06	3.00
	1280	0.22E−08	3.00	0.17E−07	3.00	0.20E−07	3.00
P^3	160	0.11E−08	–	0.83E−08	–	0.11E−07	–
	320	0.68E−10	4.00	0.52E−09	4.00	0.68E−09	4.00
	640	0.42E−11	4.00	0.32E−10	4.00	0.42E−10	4.00
	1280	0.27E−12	4.00	0.20E−11	4.00	0.26E−11	4.00

Error of global projection $P_h^\star u - u$. (Central flux) Flux parameters: $\alpha_1 = 0, \beta_1 = 0, \beta_2 = 0$

Table 11 Example 4.3

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	93	0.21E−01	–	0.12E+00	–	0.68E−01	–
	279	0.72E−02	1.00	0.40E−01	1.00	0.23E−01	1.00
	837	0.24E−02	1.00	0.13E−01	1.00	0.75E−02	1.00
	2511	0.80E−03	1.00	0.44E−02	1.00	0.25E−02	1.00
P^2	160	0.11E−05	–	0.86E−05	–	0.10E−04	–
	320	0.14E−06	3.00	0.11E−05	3.00	0.13E−05	3.00
	640	0.17E−07	3.00	0.13E−06	3.00	0.16E−06	3.00
	1280	0.22E−08	3.00	0.17E−07	3.00	0.20E−07	3.00
	2560	0.27E−09	3.00	0.21E−08	3.00	0.25E−08	3.00
P^3	160	0.27E−08	–	0.23E−07	–	0.36E−07	–
	320	0.17E−09	4.00	0.14E−08	4.00	0.22E−08	4.00
	640	0.11E−10	4.00	0.89E−10	4.00	0.14E−09	4.00
	1280	0.66E−12	4.00	0.55E−11	4.00	0.87E−11	4.00

Error of global projection $P_h^\star u - u$. Flux parameters: $\alpha_1 = 0, \beta_1 = 0, \beta_2 = 1$

asymptotic behavior of λ_1, λ_2 when h approaches 0 is indeed similar to Table 8, that is, $|\lambda_1, \lambda_2| = 1 + O(h^{-(A_1+1)/2})$ and $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^{k+1}$. Same as previous examples, order reductions are only observed when $\lim_{h\to 0} \lambda_1, \lambda_2 = 1$, that is for $k = 3$.

We use this example to compare the error bounds obtained in Lemmas 3.6 and 3.9. When $A_1 = -2$, $\delta = -(A_1 + 1) = 1$, we can verify $\left|1 - \lambda_1^N\right| \sim O(1)$, i.e., $\delta' = 0$, thus by Lemma 3.9, the convergence rate of $P_h^\star$ is k, which agrees with the simulation and is better than the one in Lemma 3.6 by half order. When $A_1 = -3$, $\delta = -(A_1 + 1) = 2, \delta' = 0$, Lemmas 3.6 and 3.9 both show a convergence rate of $k - 1$. These estimations are confirmed by the numerical results in Table 12 when $k = 3$.

Table 12 Example 4.4

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^2	320	0.28E−07	–	0.21E−06	–	0.24E−06	–
$A_1 = -3$	640	0.35E−08	3.00	0.27E−07	3.00	0.31E−07	3.00
$\tilde{\beta_2} = \frac{1}{4}$	1280	0.44E−09	3.00	0.33E−08	3.00	0.38E−08	3.00
	2560	0.55E−10	3.00	0.41E−09	3.00	0.48E−09	3.00
P^3	320	0.70E−08	–	0.57E−07	–	0.12E−06	–
$A_1 = -2$	640	0.94E−09	2.90	0.77E−08	2.90	0.16E−07	2.91
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.12E−09	2.95	0.99E−09	2.95	0.20E−08	2.95
	2560	0.15E−10	2.98	0.13E−09	2.98	0.26E−09	2.98
P^3	320	0.16E−06	–	0.13E−05	–	0.24E−05	–
$A_1 = -3$	640	0.40E−07	2.00	0.32E−06	2.00	0.61E−06	2.00
$\tilde{\beta_2} = \frac{1}{12}$	1280	0.10E−07	2.00	0.79E−07	2.00	0.15E−06	2.00
	2560	0.25E−08	2.00	0.20E−07	2.00	0.38E−07	2.00

Error of global projection $P_h^\star u - u$. Flux parameters (Case 3, and similar to Case 1.7.2 in Table 8): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = -1$, $\tilde{\beta}_2 = \frac{1}{2k(k-1)}$, $A_1 = -2, -3$, $A_2 = 1$. Note here $\lim_{h\to 0} \lambda_1, \lambda_2 = (-1)^{k+1}$

We performed more numerical results of Case 3, and all are similar to those of Case 1 as long as the eigenvalues λ_1, λ_2 are approaching 1 at the same rate. Hence, we will not show more examples about Case 3.

4.2 Numerical Results of the DG Scheme

In this subsection, we show the numerical results of the DG scheme applied to the NLS equation. For the time discretization, we use third order IMEX Runge–Kutta method [2] and fix $\Delta t = 1/10000$, which is small enough to guarantee that the spatial errors dominate. To be more precise, we treat the DG discretization of linear term u_{xx} implicitly and nonlinear term $f(|u|^2)u$ explicitly.

Example 4.5 In this example, we verify the energy conservation property of our scheme by considering the following linear equation

$$iu_t + u_{xx} = 0,$$

with the progressive plane wave solution: $u(x,t) = Aexp(i(x-t))$, with $A = 1$.

We use L^2 projection as the numerical initial condition. In the discussion of stability condition, we derive that when $\mathrm{Im}\beta_2 \geq 0, \mathrm{Im}\beta_1 \leq 0, |\alpha_1 + \overline{\alpha_2}|^2 \leq -4\mathrm{Im}\beta_1\mathrm{Im}\beta_2$, our scheme for Schrödinger equation is stable. Furthermore, when $\alpha_1 + \alpha_2 = 0, \beta_1, \beta_2$ are real numbers, the scheme is energy conservative. In this example, we compare two different parameter choices to verify the energy conservation property. The parameter choices are $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (0.25, -0.25, 1-i, 1+i)$, and $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (0.25, -0.25, 1, 1)$ when $k = 2$, $N = 40$, ending time $T = 100$. Both are numerically stable flux parameters. For the first set of parameters, we expect energy decay due to the contributions from the imaginary part of β_1, β_2 as in (9). For the second set of parameter, energy should be conserved.

In Fig. 2, we verify that as t increases from 0 to 100, the flux with only real parameters preserve $\|u_h\|_{L^2(I)}$, while the flux with complex numbers have much larger errors. More

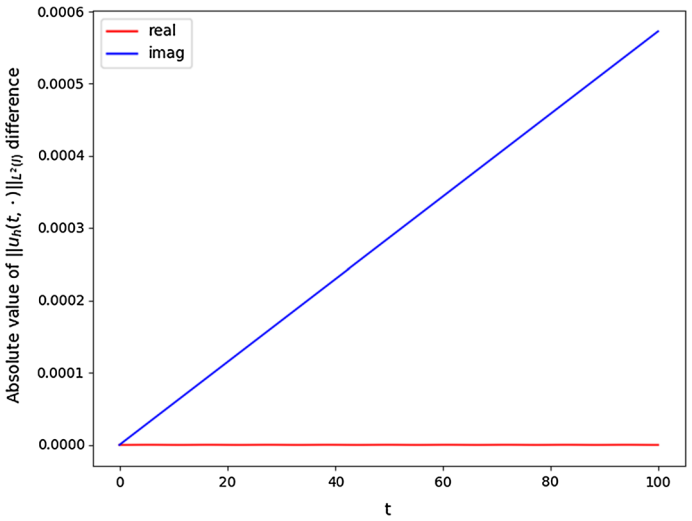

Fig. 2 Example 4.5. Absolute difference of $\|u_h(t,\cdot)\|_{L^2(I)}$ with $\|u_h(0,\cdot)\|_{L^2(I)}$ with two sets of parameters $(\alpha_1,\alpha_2,\beta_1,\beta_2) = (0.25,-0.25,1-i,1+i)$ (denoted by "imag") and $(\alpha_1,\alpha_2,\beta_1,\beta_2) = (0.25,-0.25,1,1)$ (denoted by "real") when $k=2$, $N=40$, ending time $T_e = 100$

precisely, for real parameters, $\|u_h(0,\cdot)\|_{L^2(I)} - \|u_h(100,\cdot)\|_{L^2(I)} = 7.9E\text{-}09$, for complex parameters, $\|u_h(0,\cdot)\|_{L^2(I)} - \|u_h(100,\cdot)\|_{L^2(I)} = 5.7E\text{-}04$.

Example 4.6 Accuracy test for NLS equation

$$iu_t + u_{xx} + |u|^2u + |u|^4u = 0, \tag{37}$$

which admits a progressive plane wave solution: $u(x,t) = A\exp(i(cx-\omega t))$, where $\omega = c^2 - |A|^2 - |A|^4$ with $c=1$, $A=1$.

For numerical initial condition, $P_h^\star$ is used when applicable, otherwise standard L^2 projection is applied. On uniform mesh, we use four sets of parameters. The numerical errors and orders are shown in Tables 13, 14, 15, 16, 17 and 18, where corresponding projection results are listed in Tables 1, 2, 10, 11, 7 and 9 respectively. Our numerical experiments show that the order of convergence for the scheme is the same as the order of error estimates for the projection $P_h^\star$.

We would like to make some additional comments on Tables 15 and 16, whose parameter choices belong to Case 2 when $k=1$. The existence of $P_h^\star$ requires N to be odd for this case. However, this assumption is not needed for the optimal convergence rate of the numerical scheme for (37) as shown in Tables 15 and 16. Similar comments have been made in [3].

Example 4.7 A simulation for the NLS equation

$$iu_t + u_{xx} + 2|u|^2u = 0 \tag{38}$$

with double-soliton collision

$$u(x,t) = \mathrm{sech}(x+10-4t)\exp(i(2(x+10)-3t))+\mathrm{sech}(x-10+4t)\exp(i(-2(x-10)-3t)). \tag{39}$$

Table 13 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	40	2.86E−02	–	2.48E−02	–	3.92E−02	–
	80	1.26E−02	1.18	1.02E−02	1.28	1.56E−02	1.33
	160	6.34E−03	1.00	4.99E−03	1.03	6.77E−03	1.20
	320	3.18E−03	1.00	2.56E−03	0.96	3.47E−03	0.96
	640	1.58E−03	1.01	1.27E−03	1.01	1.85E−03	0.91
P^2	40	2.22E−04	–	2.13E−04	–	6.06E−04	–
	80	1.99E−05	3.48	2.13E−05	3.33	7.28E−05	3.06
	160	3.17E−06	2.65	3.03E−06	2.81	9.01E−06	3.02
	320	3.49E−07	3.18	3.34E−07	3.18	1.23E−06	2.87
P^3	40	1.54E−06	–	1.35E−06	–	3.29E−06	–
	80	4.96E−08	4.96	4.36E−08	4.95	1.29E−07	4.67
	160	2.81E−09	4.14	2.60E−09	4.07	8.37E−09	3.95
	320	1.61E−10	4.13	1.57E−10	4.05	7.68E−10	3.45

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) on a nonuniform mesh using flux parameters (corresponding to Table 1) $\alpha_1 = 0.3$, $\beta_1 = \beta_2 = 0.4$, ending time $T_e = 0.3$

Table 14 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	40	7.47E−03	–	6.50E−03	–	1.29E−02	–
	80	2.10E−03	1.83	1.76E−03	1.89	4.22E−03	1.62
	160	4.82E−04	2.12	4.18E−04	2.07	1.16E−03	1.86
	320	1.21E−04	1.99	1.05E−04	1.99	2.87E−04	2.01
	640	3.12E−05	1.96	2.71E−05	1.95	7.40E−05	1.96
P^2	40	5.14E−04	–	5.37E−04	–	1.74E−03	–
	80	6.81E−05	2.92	7.00E−05	2.94	2.99E−04	2.54
	160	8.04E−06	3.08	8.06E−06	3.12	3.58E−05	3.06
	320	9.53E−07	3.08	9.75E−07	3.05	3.92E−06	3.19
	640	1.68E−07	2.50	1.61E−07	2.60	4.90E−07	3.00
P^3	40	1.30E−06	–	1.25E−06	–	4.09E−06	–
	80	5.74E−08	4.51	6.00E−08	4.38	2.60E−07	3.98
	160	4.44E−09	3.69	4.12E−09	3.86	1.49E−08	4.13
	320	2.25E−10	4.30	2.13E−10	4.28	9.65E−10	3.94

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) on a nonuniform mesh using flux parameters (corresponding to Table 2) $\alpha_1 = 0.3$, $\beta_1 = 0.4h_j$, $\beta_2 = 0.4/h_j$, ending time $T_e = 1$

We use periodic boundary condition and L^2 projection initialization to run the simulation for double-soliton collision solution. The two waves propagate in opposite directions and collide at $t = 2.5$, after that, the two waves separate. Such behaviors are accurately captured by our numerical simulations, see Fig. 3 for details.

Table 15 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	40	0.28E−02	–	0.22E−02	–	0.27E−02	–
	80	0.71E−03	2.00	0.56E−03	2.00	0.67E−03	2.02
	160	0.18E−03	2.00	0.14E−03	2.00	0.17E−03	2.01
	320	0.45E−04	2.00	0.35E−04	2.00	0.41E−04	2.00
	640	0.11E−04	2.00	0.88E−05	2.00	0.10E−04	2.00
P^2	40	0.13E−03	–	0.11E−03	–	0.16E−03	–
	80	0.16E−04	2.99	0.14E−04	2.99	0.20E−04	3.00
	160	0.21E−05	3.00	0.18E−05	3.00	0.25E−05	3.01
	320	0.26E−06	3.00	0.22E−06	3.00	0.31E−06	3.00
	640	0.32E−07	3.00	0.27E−07	3.00	0.39E−07	3.00
P^3	40	0.22E−06	–	0.18E−06	–	0.24E−06	–
	80	0.16E−07	3.76	0.13E−07	3.80	0.13E−07	4.16
	160	0.10E−08	4.00	0.79E−09	4.00	0.84E−09	4.00
	320	0.62E−10	4.00	0.49E−10	4.00	0.52E−10	4.00
	640	0.39E−11	3.99	0.31E−11	3.99	0.33E−11	3.96

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) using central flux (corresponding to Case 2 in Table 10) $\alpha_1 = \beta_1 = \beta_2 = 0$, ending time $T_e = 1$

Table 16 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	40	0.17E+00	–	0.13E+00	–	0.14E+00	–
	80	0.92E−01	0.90	0.72E−01	0.89	0.75E−01	0.87
	160	0.48E−01	0.94	0.38E−01	0.94	0.38E−01	0.97
	320	0.24E−01	0.97	0.19E−01	0.97	0.19E−01	0.98
	640	0.12E−01	0.98	0.97E−02	0.98	0.98E−02	0.99
P^2	40	0.13E−03	–	0.11E−03	–	0.17E−03	–
	80	0.16E−04	3.00	0.14E−04	3.00	0.20E−04	3.02
	160	0.21E−05	3.00	0.18E−05	3.00	0.25E−05	3.01
	320	0.26E−06	3.00	0.22E−06	3.00	0.31E−06	3.01
	640	0.32E−07	3.00	0.27E−07	3.00	0.39E−07	3.00
P^3	40	0.68E−06	–	0.56E−06	–	0.83E−06	–
	80	0.42E−07	4.00	0.35E−07	4.01	0.51E−07	4.01
	160	0.26E−08	4.00	0.22E−08	4.00	0.32E−08	4.00
	320	0.16E−09	4.00	0.14E−09	4.00	0.20E−09	4.00
	640	0.10E−10	4.00	0.85E−11	4.00	0.13E−10	4.00

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) using flux parameters (corresponding to Case 2 in Table 11): $\alpha_1 = \beta_1 = 0$, $\beta_2 = 1$, ending time $T_e = 1$

Table 17 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^1	40	0.41E−02	–	0.37E−02	–	0.72E−02	–
$A_2 = 2$	80	0.12E−02	1.77	0.10E−02	1.82	0.21E−02	1.80
$\tilde{\beta}_1 = \frac{1}{4}$	160	0.31E−03	1.93	0.25E−03	2.05	0.39E−03	2.39
	320	0.87E−04	1.86	0.69E−04	1.87	0.10E−03	1.94
	640	0.23E−04	1.93	0.18E−04	1.94	0.26E−04	1.97
P^2	40	0.49E−04	–	0.49E−04	–	0.13E−03	–
$A_2 = 2$	80	0.83E−05	2.55	0.73E−05	2.74	0.14E−04	3.23
$\tilde{\beta}_1 = \frac{7}{4}$	160	0.31E−05	1.44	0.29E−05	1.32	0.65E−05	1.12
	320	0.95E−06	1.69	0.92E−06	1.69	0.20E−05	1.70
	640	0.26E−06	1.85	0.25E−06	1.86	0.55E−06	1.87
P^2	40	0.36E−03	–	0.34E−03	–	0.74E−03	–
$A_2 = 3$	80	0.21E−03	0.78	0.20E−03	0.76	0.43E−03	0.77
$\tilde{\beta}_1 = \frac{7}{4}$	160	0.11E−03	0.92	0.11E−03	0.92	0.23E−03	0.92
	320	0.56E−04	1.00	0.53E−04	1.00	0.11E−03	0.99
	640	0.28E−04	1.00	0.27E−04	1.00	0.58E−04	1.00
P^3	40	0.19E−05	–	0.19E−05	–	0.43E−05	–
$A_2 = 2$	80	0.43E−07	5.50	0.38E−07	5.65	0.84E−07	5.66
$\tilde{\beta}_1 = \frac{9}{2}$	160	0.15E−08	4.88	0.15E−08	4.68	0.26E−08	5.00
	320	0.91E−10	4.00	0.90E−10	4.02	0.17E−09	3.94
	640	0.58E−11	3.96	0.57E−11	3.99	0.11E−10	3.98

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) using flux parameters (corresponding to Case 1.6.1 in Table 7): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = \frac{k(k-1)}{2} + \frac{k(k+1)}{8}$, $\tilde{\beta}_2 = 1.0$, $A_1 = -1$, $A_2 = 2, 3$, ending time $T_e = 1$

Table 18 Example 4.6

	N	L^1 error	Order	L^2 error	Order	L^∞ error	Order
P^2	40	0.60E−04	–	0.54E−04	–	0.95E−04	–
$A_1 = -2$	80	0.76E−05	2.99	0.68E−05	2.98	0.12E−04	2.96
$\tilde{\beta}_1 = \frac{1}{12}$	160	0.96E−06	3.00	0.85E−06	3.00	0.15E−05	2.99
	320	0.12E−06	3.00	0.11E−06	3.00	0.19E−06	2.99
	640	0.15E−07	3.00	0.13E−07	3.00	0.24E−07	3.00
P^2	40	0.95E−04	–	0.85E−04	–	0.15E−03	–
$A_1 = -3$	80	0.21E−04	2.22	0.18E−04	2.20	0.33E−04	2.18
$\tilde{\beta}_1 = \frac{1}{12}$	160	0.49E−05	2.08	0.44E−05	2.07	0.79E−05	2.06
	320	0.12E−05	2.02	0.11E−05	2.02	0.20E−05	2.02
	640	0.29E−06	2.02	0.27E−06	2.02	0.48E−06	2.02

Error in L^1, L^2 and L^∞ norm for solving NLS equation (37) using flux parameters (corresponding to Case 1.7.2 in Table 9): $\alpha_1 = 0.25$, $\tilde{\beta}_1 = -1$, $\tilde{\beta}_2 = \frac{1}{2k(k+1)}$, $A_1 = -2, -3$, $A_2 = 1$, ending time $T_e = 1$

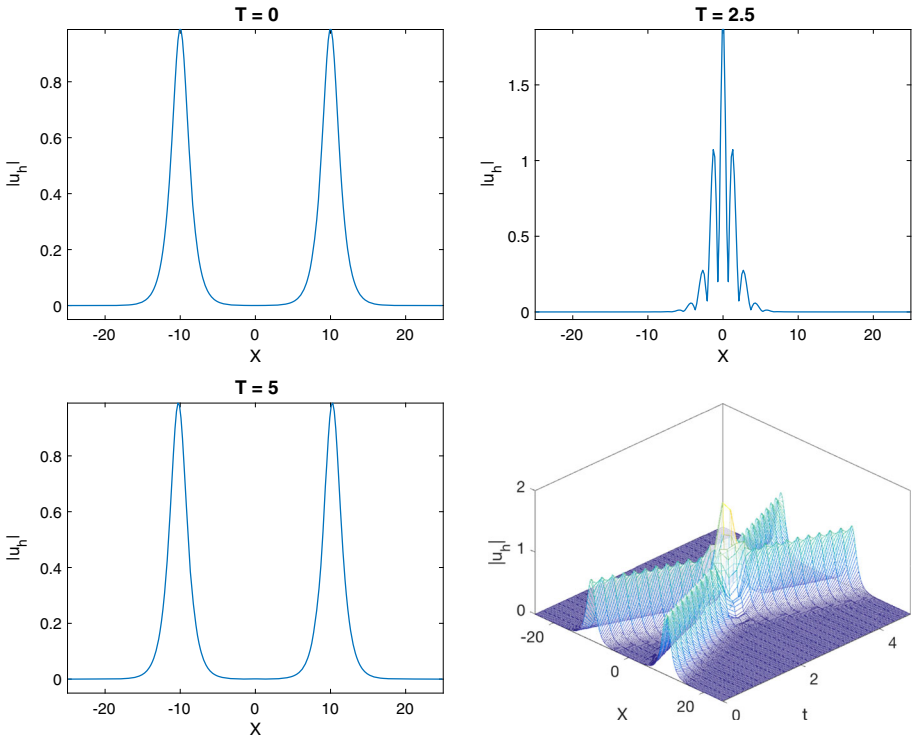

Fig. 3 Example 4.7. Double soliton collision graphs at $t = 0, 2.5, 5$ and a $x - t$ plot of the numerical solution. $N = 250$, P^2 elements with periodic boundary conditions on [-25,25]. Central flux ($\alpha_1 = \beta_1 = \beta_2 = 0$) is used

5 Conclusions and Future Work

In this paper, we studied the ultra-weak DG method with a general class of numerical fluxes for solving one-dimensional nonlinear Schrödinger equation with periodic boundary conditions. Semi-discrete L^2 stability and error estimates are obtained when the polynomial degree $k \geq 1$. Focusing on the real parameters, we performed detailed investigation of the associated projection operators. Our analysis assume the dependence of parameters on the mesh size h can be freely enforced, hence several cases follow. A variety of analytic tools are employed, including decoupling of global projection into local projection, analysis of block-circulant matrix and Fourier analysis. We acquire error bounds that are sharp in most cases from numerical verifications. Future work includes improvement of the error bounds for some suboptimal cases, superconvergence studies and generalization to higher-dimensions.

A Appendix

A.1 Proof of Lemma 3.2

First, we consider the case when $\beta_1 \neq 0$. Define the difference operator $Wu = P_h^\star u - P_h^1 u$, then (17) implies:

$$\begin{aligned} &\int_{I_j} Wu\, v_h dx = 0 && \forall v_h \in P^{k-2}(I_j), \\ &Wu^+ + \frac{\frac{1}{2}+\alpha_1}{\beta_1}(Wu)_x^+ = u - (P_h^1 u)^+ && \text{at } x_{j-\frac{1}{2}}, \\ &Wu^- - \frac{\frac{1}{2}-\alpha_1}{\beta_1}(Wu)_x^- = -\frac{\frac{1}{2}-\alpha_1}{\beta_1}\left(u_x - \left(P_h^1 u\right)_x^-\right) && \text{at } x_{j+\frac{1}{2}}. \end{aligned} \tag{40}$$

For $l \geq 0$, let $P_l(\xi)$ be the l-th order Legendre polynomials on $[-1,1]$, with $\xi = \frac{2(x-x_j)}{h_j}$ on I_j, and define $P_{j,l}(x) = P_l(\frac{2(x-x_j)}{h_j}) = P_l(\xi)$. Then Wu can be expressed as:

$$Wu(x) = \sum_{l=0}^{k} a_{j,l} P_{j,l}(x) = \sum_{l=0}^{k} a_{j,l} P_l(\xi).$$

By the first equation in (40) and orthogonality of Legendre polynomials, one can get:

$$a_{j,l} = 0, \quad l = 0, \ldots, k-2, \quad j = 1, \ldots, N.$$

We can then move on to solve for $a_{j,k-1}$ and $a_{j,k}$ on each cell directly by the second and third equations in (40). By properties of Legendre polynomials: $P_l(\pm 1) = (\pm 1)^l$, $P_l'(\pm 1) = \frac{1}{2}(\pm 1)^{l-1} l(l+1)$, the following 2×2 linear system holds on each cell I_j:

$$\mathcal{M}_j \begin{bmatrix} a_{j,k-1} \\ a_{j,k} \end{bmatrix} = \begin{bmatrix} \phi_j \\ \psi_j \end{bmatrix},$$

where

$$\mathcal{M}_j = \begin{bmatrix} (\mathcal{M}_j)_{11} & (\mathcal{M}_j)_{12} \\ (\mathcal{M}_j)_{21} & (\mathcal{M}_j)_{22} \end{bmatrix} = \begin{bmatrix} (-1)^{k-1} + (-1)^k \frac{\frac{1}{2}+\alpha_1}{\beta_1}\frac{k(k-1)}{h_j} & (-1)^k + (-1)^{k-1}\frac{\frac{1}{2}+\alpha_1}{\beta_1}\frac{k(k+1)}{h_j} \\ 1 - \frac{\frac{1}{2}-\alpha_1}{\beta_1}\frac{k(k-1)}{h_j} & 1 - \frac{\frac{1}{2}-\alpha_1}{\beta_1}\frac{k(k+1)}{h_j} \end{bmatrix}$$

and $\phi_j = (u - (P_h^1 u)^+)|_{x_{j-\frac{1}{2}}}$ and $\psi_j = -\frac{\frac{1}{2}-\alpha_1}{\beta_1}(u_x - (P_h^1 u)_x^-)|_{x_{j+\frac{1}{2}}}$.

We can calculate the determinant of the matrix $\mathcal{M}_j$ to be $2(-1)^{k-1} + 2(-1)^k \frac{k^2}{\beta_1 h_j} + 2(-1)^{k-1}\frac{\beta_2 k^2(k^2-1)}{\beta_1 h_j^2} = 2(-1)^{k-1}\Gamma_j/\beta_1$. Hence, when $\Gamma_j \neq 0$, $\forall j$, $P_h^\star$ exists and is unique. We now move on to estimate the $a_{j,k-1}, a_{j,k}$. Clearly,

$$\begin{aligned} a_{j,k-1} &= \frac{1}{\det \mathcal{M}_j}((\mathcal{M}_j)_{22}\phi_j - (\mathcal{M}_j)_{12}\psi_j) \\ a_{j,k} &= \frac{1}{\det \mathcal{M}_j}(-(\mathcal{M}_j)_{21}\phi_j + (\mathcal{M}_j)_{11}\psi_j), \end{aligned}$$

and from the projection property of P_h^1, $|\phi_j| \leq C h_j^{k+1}|u|_{W^{k+1,\infty}(I_j)}$, $|\psi_j| \leq C h_j^k |\frac{\frac{1}{2}-\alpha_1}{\beta_1}|$ $||u|_{W^{k+1,\infty}(I_j)}$. The error estimates can be obtained based on the following cases.

- If $k = 1$, then

$$a_{j,0} = \frac{1}{2\Gamma_j}\left(\left(\beta_1 - \left(\frac{1}{2} - \alpha_1\right)\frac{2}{h_j}\right)(u - P_h^1 u)^+|_{x_{j-\frac{1}{2}}} - \left(\frac{1}{2} - \alpha_1 - \frac{2\beta_2}{h_j}\right)(u_x - (P_h^1 u)_x)^-|_{x_{j+\frac{1}{2}}}\right),$$

$$a_{j,1} = \frac{1}{2\Gamma_j}\left(-\beta_1(u - P_h^1 u)^+|_{x_{j-\frac{1}{2}}} - \left(\frac{1}{2} - \alpha_1\right)(u_x - (P_h^1 u)_x)^-|_{x_{j+\frac{1}{2}}}\right).$$

Thus we have estimates

$$|a_{j,0}| \le \frac{Ch_j^2|u|_{W^{2,\infty}(I_j)}}{|\Gamma_j|}\max\left(\left|\beta_1 - \frac{1-2\alpha_1}{h_j}\right|, \left|\frac{\frac{1}{2}-\alpha_1}{h_j} - \frac{2\beta_2}{h_j^2}\right|\right),$$

$$|a_{j,1}| \le \frac{Ch_j^2|u|_{W^{2,\infty}(I_j)}}{|\Gamma_j|}\max\left(|\beta_1|, \left|\frac{\frac{1}{2}-\alpha_1}{h_j}\right|\right).$$

Then,

$$\begin{aligned}\|P_h^\star u - P_h^1 u\|_{L^\infty(I_j)} &= \|a_{j,0}P_0(\xi) + \alpha_{j,1}P_1(\xi)\|_{L^\infty(I_j)} \\ &\le \frac{Ch_j^2|u|_{W^{2,\infty}(I_j)}}{|\Gamma_j|}\max\left(|\beta_1|, \left|\frac{\frac{1}{2}-\alpha_1}{h_j}\right|, \left|\frac{\beta_2}{h_j^2}\right|\right). \end{aligned} \tag{41}$$

Combining with the error estimates for P_h^1 and the mesh regularity assumption, we get

$$\|P_h^\star u - u\|_{L^p(I)} \le Ch^2|u|_{W^{2,\infty}(I)}\left(1 + \frac{\max\left(|\beta_1|, \frac{\left|\frac{1}{2}-\alpha_1\right|}{h}, \frac{|\beta_2|}{h^2}\right)}{\min_j|\Gamma_j|}\right), \quad p = 2, \infty.$$

- If $k > 1$, then we need to discuss the case when $\beta_2 = 0$ or $\beta_2 \ne 0$.
If $\beta_2 = 0$, then $\alpha_1 = \pm\frac{1}{2}$. When $\alpha_1 = \frac{1}{2}$, we have $\psi_j = 0$, and

$$|a_{j,k-1}| \le Ch_j^{k+1}|u|_{W^{k+1,\infty}(I_j)}\frac{|\beta_1|}{|\Gamma_j|},$$

$$|a_{j,k}| \le Ch_j^{k+1}|u|_{W^{k+1,\infty}(I_j)}\frac{|\beta_1|}{|\Gamma_j|}.$$

Therefore,

$$\|P_h^\star u - P_h^1 u\|_{L^\infty(I_j)} \le Ch_j^{k+1}|u|_{W^{k+1,\infty}(I_j)}\frac{|\beta_1|}{|\Gamma_j|}, \tag{42}$$

implying a supercloseness between $P_h^\star$ and P_h^1 if $\beta_1/\Gamma_j = o(1)$. In summary, we have

$$\|P_h^\star u - u\|_{L^p(I)} \le Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\left(1 + \frac{|\beta_1|}{\min_j|\Gamma_j|}\right), \quad p = 2, \infty.$$

When $\alpha_1 = -\frac{1}{2}$, we then should compare the projection with P_h^2 instead of P_h^1. We skip the details of the calculations. The conclusion is similar, i.e.

$$\|P_h^\star u - P_h^2 u\|_{L^\infty(I_j)} \le Ch_j^{k+1}|u|_{W^{k+1,\infty}(I_j)}\frac{|\beta_1|}{|\Gamma_j|},$$

and

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} |u|_{W^{k+1,\infty}(I)} \left(1 + \frac{|\beta_1|}{\min_j |\Gamma_j|}\right), \quad p = 2, \infty.$$

If $\beta_2 \neq 0$, similar to previous case, we can show

$$|a_{j,0}| \le \frac{C h_j^{k+1} |u|_{W^{k+1,\infty}(I_j)}}{|\Gamma_j|} \max\left(|\beta_1|, \left|\frac{\frac{1}{2} - \alpha_1}{h_j}\right|, \left|\frac{\beta_2}{h_j^2}\right|\right),$$
$$|a_{j,1}| \le \frac{C h_j^{k+1} |u|_{W^{k+1,\infty}(I_j)}}{|\Gamma_j|} \max\left(|\beta_1|, \left|\frac{\frac{1}{2} - \alpha_1}{h_j}\right|, \left|\frac{\beta_2}{h_j^2}\right|\right).$$

Therefore,

$$\|P_h^\star u - u\|_{L^\infty(I_j)} \le C h_j^{k+1} |u|_{W^{k+1,\infty}(I_j)} \left(1 + \frac{\max\left(|\beta_1|, \left|\frac{\frac{1}{2} - \alpha_1}{h_j}\right|, \left|\frac{\beta_2}{h_j^2}\right|\right)}{|\Gamma_j|}\right)$$

and it leads to

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} |u|_{W^{k+1,\infty}(I)} \left(1 + \frac{\max\left(|\beta_1|, \frac{|\frac{1}{2} - \alpha_1|}{h}, \frac{|\beta_2|}{h^2}\right)}{\min_j |\Gamma_j|}\right), \quad p = 2, \infty.$$

Finally, when $\beta_1 = 0, \beta_2 \neq 0, \alpha_1 = \pm\frac{1}{2}$, we have the following estimates

$$\|P_h^\star u - u\|_{L^p(I)} \le C h^{k+1} |u|_{W^{k+1,\infty}(I)} \left(1 + \frac{|\beta_2|}{h^2 \min_j |\Gamma_j|}\right), \quad p = 2, \infty.$$

Summarizing all the estimates, we have shown (20) for all cases.

A.2 Proof of Lemma 3.3

We adopt similar notations as in the proof of Lemma 3.2. Define $\xi = \frac{2(x - x_j)}{h}$, and let

$$P_h^\star u(x)|_{I_j} = \sum_{l=0}^{k} \gamma_{j,l} P_{j,l}(x) = \sum_{l=0}^{k} \gamma_{j,l} P_l(\xi).$$

By (11a) and orthogonality of Legendre polynomials, one can get:

$$\gamma_{j,l} = \frac{2l+1}{2} \int_{-1}^{1} u\left(x_j + \frac{h}{2}\xi\right) P_l(\xi) d\xi, \quad l = 0, \ldots, k-2, \quad j = 1, \ldots, N.$$

We can then move on to solve for $\gamma_{j,k-1}$ and $\gamma_{j,k}$ from (11b)–(11c). At $x_{j+\frac{1}{2}}$,

$$\begin{bmatrix} \widehat{P_h^\star u} \\ (P_h^\star u)_x \end{bmatrix} = \sum_{l=0}^{k} \gamma_{j,l} A_0 \begin{bmatrix} P_l(1) \\ \frac{2}{h} P_l'(1) \end{bmatrix} + \gamma_{j+1,l} B_0 \begin{bmatrix} P_l(-1) \\ \frac{2}{h} P_l'(-1) \end{bmatrix} = \begin{bmatrix} u\left(x_{j+\frac{1}{2}}\right) \\ u_x\left(x_{j+\frac{1}{2}}\right) \end{bmatrix}, \tag{43}$$

where A_0, B_0 are defined in (16).

Combining (43) for all j and using the periodic boundary condition will result in the following $2N \times 2N$ linear system

$$M \begin{bmatrix} \gamma_{1,k-1} \\ \gamma_{1,k} \\ \cdots \\ \gamma_{N-1,k-1} \\ \gamma_{N-1,k} \\ \gamma_{N,k-1} \\ \gamma_{N,k} \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \psi_1 \\ \cdots \\ \phi_N - 1 \\ \psi_N - 1 \\ \phi_N \\ \psi_N \end{bmatrix} \tag{44}$$

where $M = circ(A, B, 0_2, \ldots, 0_2)$, denoting a $2N \times 2N$ block-circulant matrix with first two rows as $(A, B, 0_2, \ldots, 0_2)$, with 0_2 as a 2×2 zero matrix, and

$$A = A_0 \begin{bmatrix} P_{k-1}(1) & P_k(1) \\ \frac{2}{h} P'_{k-1}(1) & \frac{2}{h} P'_k(1) \end{bmatrix}, \tag{45}$$

$$B = B_0 \begin{bmatrix} P_{k-1}(-1) & P_k(-1) \\ \frac{2}{h} P'_{k-1}(-1) & \frac{2}{h} P'_k(-1) \end{bmatrix},$$

$$\begin{bmatrix} \phi_j \\ \psi_j \end{bmatrix} = \begin{bmatrix} u\left(x_{j+\frac{1}{2}}\right) \\ u_x\left(x_{j+\frac{1}{2}}\right) \end{bmatrix} - \sum_{l=0}^{k-2} \left(\gamma_{j,l} A_0 \begin{bmatrix} P_l(1) \\ \frac{2}{h} P'_l(1) \end{bmatrix} + \gamma_{j+1,l} B_0 \begin{bmatrix} P_l(-1) \\ \frac{2}{h} P'_l(-1) \end{bmatrix} \right), \tag{46}$$

where A_0, B_0 are defined in (16). We can calculate that

$$\det A = \det B = \frac{-2k}{h} \left(\alpha_1^2 + \beta_1 \beta_2 - \frac{1}{4} \right) := \Lambda \neq 0. \tag{47}$$

It is clear that the existence and uniqueness of $P_h^\star$ is equivalent to $\det M \neq 0$. By a direct computation, $\det M = \det A^N \det(I_2 - Q^N)$, where I_2 denotes the 2×2 identity matrix, and

$$Q = -A^{-1} B = \frac{(-1)^{k+1}}{\Lambda} \begin{bmatrix} c_1 + c_2 & b_1 + b_2 \\ b_1 - b_2 & c_1 - c_2 \end{bmatrix},$$

with

$$c_1 = \beta_1 + \frac{k^2(k^2-1)}{h^2} \beta_2 - \frac{2k^2}{h} \left(\alpha_1^2 + \beta_1 \beta_2 + \frac{1}{4} \right) := \Gamma, \tag{48}$$

$$c_2 = \frac{k}{h} (2\alpha_1), \tag{49}$$

$$b_1 = -\beta_1 - \frac{k^2(k^2+1)}{h^2} \beta_2 + \frac{2k^2}{h} \left(\alpha_1^2 + \beta_1 \beta_2 + \frac{1}{4} \right), \tag{50}$$

$$b_2 = -\frac{2k^3}{h^2} \beta_2 + \frac{2k}{h} \left(\alpha_1^2 + \beta_1 \beta_2 + \frac{1}{4} \right). \tag{51}$$

The eigenvalues of Q are

$$\lambda_1 = \frac{(-1)^{(k+1)}}{\Lambda} (\Gamma + \sqrt{\Gamma^2 - \Lambda^2}), \quad \lambda_2 = \frac{(-1)^{(k+1)}}{\Lambda} (\Gamma - \sqrt{\Gamma^2 - \Lambda^2}). \tag{52}$$

Since $\det Q = \det B / \det A = 1$, we have the relations $\lambda_1 \lambda_2 = 1$ and

$$b_1^2 - b_2^2 = \Gamma^2 - \Lambda^2 - c_2^2. \tag{53}$$

Below we will discuss the existence and uniqueness of $P_h^\star$ based on three cases depending on the relation of Γ and Λ.

Case 1. If $|\Gamma| > |\Lambda|$, then $\lambda_{1,2}$ are real and different. Therefore, we can perform eigenvalue decomposition of Q,

$$Q = TDT^{-1},$$

where

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

and

$$T = \begin{bmatrix} 1 & -\frac{b_1+b_2}{c_2+\sqrt{\Gamma^2-\Lambda^2}} \\ \frac{b_1-b_2}{c_2+\sqrt{\Gamma^2-\Lambda^2}} & 1 \end{bmatrix}, T^{-1} = \frac{1}{\det T}\begin{bmatrix} 1 & \frac{b_1+b_2}{c_2+\sqrt{\Gamma^2-\Lambda^2}} \\ -\frac{b_1-b_2}{c_2+\sqrt{\Gamma^2-\Lambda^2}} & 1 \end{bmatrix}, \tag{54}$$

where $\det T = \frac{2\sqrt{\Gamma^2-\Lambda^2}}{c_2+\sqrt{\Gamma^2-\Lambda^2}}$, except for the case when $(b_1 - b_2)(b_1 + b_2) = 0$ and $c_2 < 0$, where

$$T = \begin{bmatrix} 1 & -\frac{b_1+b_2}{2c_2} \\ \frac{b_1-b_2}{2c_2} & 1 \end{bmatrix}, T^{-1} = \begin{bmatrix} 1 & \frac{b_1+b_2}{2c_2} \\ -\frac{b_1-b_2}{2c_2} & 1 \end{bmatrix}. \tag{55}$$

In both situations, we have

$$\det M = \det A^N \det\left(I_2 - \begin{bmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{bmatrix}\right) = \det A^N \det\left(\begin{bmatrix} 1-\lambda_1^N & 0 \\ 0 & 1-\lambda_2^N \end{bmatrix}\right).$$

$\det M \neq 0$ if and only if $(\lambda_1)^N \neq 1$ and $(\lambda_2)^N \neq 1$. This is clearly true since $|\lambda_1, \lambda_2| \neq 1$.

Case 2. If $|\Gamma| = |\Lambda|$, then $\lambda_1 = \lambda_2 = (-1)^{k+1}\frac{\Gamma}{\Lambda}$ and we have two repeated eigenvalues. Perform Jordan decomposition:

$$\begin{bmatrix} c_1 + c_2 & b_1 + b_2 \\ b_1 - b_2 & c_1 - c_2 \end{bmatrix} = \mathcal{T}\begin{bmatrix} c_1 & 1 \\ 0 & c_1 \end{bmatrix}\mathcal{T}^{-1},$$

and

$$\mathcal{T} = \begin{bmatrix} c_2 & 1 \\ b_1 - b_2 & 0 \end{bmatrix}, \qquad \text{if } b_1 \neq b_2,$$
$$\mathcal{T} = \begin{bmatrix} 2b_1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \text{if } b_1 = b_2. \tag{56}$$

We define

$$J = \begin{bmatrix} c_1 & 1 \\ 0 & c_1 \end{bmatrix}, \qquad \mathcal{J} = \frac{(-1)^{k+1}}{\Lambda}\begin{bmatrix} c_1 & 1 \\ 0 & c_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \frac{(-1)^{k+1}}{\Lambda} \\ 0 & \lambda_1 \end{bmatrix},$$

then

$$Q^j = \mathcal{T}\mathcal{J}^j\mathcal{T}^{-1}, \quad \mathcal{J}^j = \begin{bmatrix} \lambda_1^j & \kappa_j \\ 0 & \lambda_1^j \end{bmatrix},$$
$$I_2 - Q^N = \mathcal{T}\begin{bmatrix} 1-(\lambda_1)^N & -\kappa_N \\ 0 & 1-(\lambda_1)^N \end{bmatrix}\mathcal{T}^{-1},$$

where $\kappa_j = \frac{(-1)^{(k+1)j}}{\Lambda^j} j\Gamma^{j-1}$.

In both situations, $\det M \neq 0$ if and only if $(\lambda_1)^N \neq 1$, meaning that we require N to be odd and further, if k is odd, we require $\Gamma = -\Lambda$; if k is even, we require $\Gamma = \Lambda$. In both cases, $\lambda_1 = \lambda_2 = -1$.

Case 3. If $|\Gamma| < |\Lambda|$, then $\lambda_{1,2}$ are complex, $|\lambda_{1,2}| = 1$, $\lambda_1 = \overline{\lambda_2}$, still Q is diagonalizable, and similar to Case 1, $\det M \neq 0$ turns to $(\lambda_1)^N \neq 1$ and $(\lambda_2)^N = \overline{(\lambda_1)^N} \neq 1$, i.e. we require

$$(-1)^{(k+1)N}\left(\frac{\Gamma}{\Lambda}+\sqrt{\left(\frac{\Gamma}{\Lambda}\right)^2-1}\right)^N \neq 1.$$

A.3 Proof of Lemmas 3.4–3.6

A.3.1 Proof of Lemma 3.4

In the proof, we still use the difference operator $Wu = P_h^\star u - P_h^1 u = \sum_{l=0}^k \alpha_{j,l} P_{j,l}(x) = \sum_{l=0}^k \alpha_{j,l} P_l(\xi)$, with $\alpha_{j,l} = 0, \quad l = 0, \ldots, k-2, \quad j = 1, \ldots, N$, and

$$M\begin{bmatrix}\alpha_{1,k-1}\\ \alpha_{1,k}\\ \cdots\\ \alpha_{N-1,k-1}\\ \alpha_{N-1,k}\\ \alpha_{N,k-1}\\ \alpha_{N,k}\end{bmatrix} = \begin{bmatrix}\tau_1\\ \iota_1\\ \cdots\\ \tau_{N-1}\\ \iota_{N-1}\\ \tau_N\\ \iota_N\end{bmatrix}, \tag{57}$$

where

$$\begin{bmatrix}\tau_j\\ \iota_j\end{bmatrix} = \begin{bmatrix}\frac{1}{2}-\alpha_1 & -\beta_2\\ \beta_1 & \frac{1}{2}-\alpha_1\end{bmatrix}\begin{bmatrix}\eta_j\\ \theta_j\end{bmatrix}, \quad \begin{bmatrix}\eta_j\\ \theta_j\end{bmatrix} = \begin{bmatrix}u-(P_h^1 u)^+\\ u_x-(P_h^1 u)_x^-\end{bmatrix}_{j+\frac{1}{2}}.$$

We will now analyze the inverse of the matrix M. It is known that the inverse of a nonsingular circulant matrix is also circulant, so is a block-circulant matrix. In particular,

$$M^{-1} = circ(r_0, r_1, \ldots, r_{N-1}) \otimes A^{-1}$$

where $\otimes$ means Kronecker product for block matrices and r_j is a 2×2 matrix defined as,

$$\begin{aligned} r_j &= Q^j(I_2 - Q^N)^{-1}, \quad j = 0, \ldots, N-1\\ &= T D^j (I_2 - D^N)^{-1} T^{-1},\\ D^j(I_2 - D^N)^{-1} &= \begin{bmatrix}\frac{\lambda_1^j}{1-\lambda_1^N} & 0\\ 0 & \frac{\lambda_2^j}{1-\lambda_2^N}\end{bmatrix} := \begin{bmatrix}d_1^j & 0\\ 0 & d_2^j\end{bmatrix}, \quad \text{and} \quad d_2^j = -d_1^{N-j}. \end{aligned} \tag{58}$$

For the convenience of further analysis, we separate r_j in terms of d_1^j and d_2^j,

$$\begin{aligned} r_j &= d_1^j T\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}T^{-1} + d_2^j T\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}T^{-1}\\ &:= d_1^j Q_1 + d_2^j (I_2 - Q_1), \end{aligned} \tag{59}$$

where

$$Q_1 = \frac{1}{2\sqrt{\Gamma^2 - \Lambda^2}} \begin{bmatrix} c_2 + \sqrt{\Gamma^2 - \Lambda^2} & b_1 + b_2 \\ b_1 - b_2 & -c_2 + \sqrt{\Gamma^2 - \Lambda^2} \end{bmatrix}, \tag{60}$$

when T is given by (54), and

$$Q_1 = \frac{1}{2c_2} \begin{bmatrix} 2c_2 & b_1 + b_2 \\ b_1 - b_2 & 0 \end{bmatrix}, \tag{61}$$

when T is given by (55).

For Case 1, eigenvalues $\lambda_{1,2}$ are real. $\sum_{j=0}^{N-1} |d_{1,2}^j| = \frac{1}{1-|\lambda_{1,2}|} \frac{1-|\lambda_{1,2}|^N}{|1-\lambda_{1,2}^N|}$. Without loss of generality, we assume $|\lambda_1| < 1 < |\lambda_2|$, which is equivalent to $\Gamma < 0$, then

$$\sum_{j=0}^{N-1} |d_1^j| \le \frac{1}{1-|\lambda_1|} = \frac{|\lambda_2|}{|\lambda_2| - 1}, \tag{62}$$

$$\sum_{j=0}^{N-1} |d_2^j| \le \frac{1}{|\lambda_2| - 1}. \tag{63}$$

We let

$$\begin{bmatrix} \Xi_j \\ \Theta_j \end{bmatrix} := A^{-1} \begin{bmatrix} \tau_j \\ \iota_j \end{bmatrix} = \eta_j V_1 + \theta_j V_2, \qquad j = 1, \ldots, N \tag{64}$$

where

$$V_1 = \frac{1}{\Lambda} \begin{bmatrix} -\beta_1 + \frac{k(k+1)}{h} \left(\left(\frac{1}{2} - \alpha_1 \right)^2 + \beta_1 \beta_2 \right) \\ \beta_1 - \frac{k(k-1)}{h} \left(\left(\frac{1}{2} - \alpha_1 \right)^2 + \beta_1 \beta_2 \right) \end{bmatrix},$$

$$V_2 = \frac{1}{\Lambda} \begin{bmatrix} \alpha_1^2 + \beta_1 \beta_2 - \frac{1}{4} \\ -\left(\alpha_1^2 + \beta_1 \beta_2 - \frac{1}{4} \right) \end{bmatrix} = \begin{bmatrix} -\frac{h}{2k} \\ \frac{h}{2k} \end{bmatrix}, \tag{65}$$

$$\|V_1\|_\infty \le C \left(1 + \frac{\max(|\beta_1|, |\frac{1}{2} - \alpha_1|/h)}{|\Lambda|} \right), \qquad \|V_2\|_\infty \le Ch \tag{66}$$

from basic algebraic calculations. Therefore,

$$\begin{bmatrix} \alpha_{m,k-1} \\ \alpha_{m,k} \end{bmatrix} = \sum_{j=0}^{N-1} r_j \begin{bmatrix} \Xi_{j+m} \\ \Theta_{j+m} \end{bmatrix}, \quad m = 1, \ldots, N, \tag{67}$$

where by periodicity, when $j + m > N$, $\Xi_{j+m} = \Xi_{j+m-N}$, $\Theta_{j+m} = \Theta_{j+m-N}$.

In summary, we obtain the estimation when $|\lambda_1| < 1 < |\lambda_2|$,

$$\begin{aligned}\left\|\begin{bmatrix}\alpha_{m,k-1}\\ \alpha_{m,k}\end{bmatrix}\right\|_\infty &\le \sum_{j=0}^{N-1}(|d_1^j|+|d_2^j|)\left(\max_j|\eta_j|\,\|Q_1V_1\|_\infty+\max_j|\theta_j|\|Q_1V_2\|_\infty\right)\\ &\quad+\sum_{j=0}^{N-1}|d_2^j|\left(\max_j|\eta_j|\|V_1\|_\infty+\max_j|\theta_j|\|V_2\|_\infty\right),\\ &\le Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\Big(\frac{|\lambda_2|+1}{|\lambda_2|-1}\left(\|Q_1V_1\|_\infty+h^{-1}\|Q_1V_2\|_\infty\right)\\ &\quad+\frac{1}{|\lambda_2|-1}\left(\|V_1\|_\infty+h^{-1}\|V_2\|_\infty\right)\Big),\qquad m=1,\dots,N.\end{aligned}\tag{68}$$

Thus, the estimates for the difference between $P_h^\star$ and P_h^1 are

$$\begin{aligned}\|P_h^\star u-P_h^1u\|_{L^p(I)} &\le Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\Big(\frac{|\lambda_2|+1}{|\lambda_2|-1}\left(\|Q_1V_1\|_\infty+h^{-1}\|Q_1V_2\|_\infty\right)\\ &\quad+\frac{1}{|\lambda_2|-1}\left(\|V_1\|_\infty+h^{-1}\|V_2\|_\infty\right)\Big).\end{aligned}\tag{69}$$

Similar estimates can be proved when $\Gamma > 0$ and $|\lambda_1| > 1 > |\lambda_2|$,

$$\begin{aligned}&\|P_h^\star u-P_h^1u\|_{L^p(I)}\\ &\le Ch^{k+1}|u|_{W^{k+1,\infty}(I)}\Big(\frac{|\lambda_1|+1}{|\lambda_1|-1}\left(\|(I_2-Q_1)V_1\|_\infty+h^{-1}\|(I_2-Q_1)V_2\|_\infty\right)\\ &\quad+\frac{1}{|\lambda_1|-1}\left(\|V_1\|_\infty+h^{-1}\|V_2\|_\infty\right)\Big),\end{aligned}\tag{70}$$

and (21) is obtained.

A.3.2 Proof of Lemma 3.5

Since $P_h^\star$ is well defined, we know that $\lambda_1^N=-1$. Therefore, we can obtain

$$\begin{aligned}I_2-Q^N=\mathcal{T}\begin{bmatrix}2&\frac{N}{\Gamma}\\0&2\end{bmatrix}\mathcal{T}^{-1},\ (I_2-Q^N)^{-1}=\mathcal{T}\begin{bmatrix}\frac12&-\frac{N}{4\Gamma}\\0&\frac12\end{bmatrix}\mathcal{T}^{-1},\\ r_j=Q^j(I_2-Q^N)^{-1}=\frac{(-1)^j}{2}I_2+(-1)^j\frac{-N+2j}{4\Gamma}Q_2,\end{aligned}\tag{71}$$

where

$$Q_2=\mathcal{T}\begin{bmatrix}0&1\\0&0\end{bmatrix}\mathcal{T}^{-1}=\begin{bmatrix}c_2&b_1+b_2\\b_1-b_2&-c_2\end{bmatrix}.\tag{72}$$

Therefore, we have for $m=0,\dots,N-1$:

$$\begin{aligned}\left\|\begin{bmatrix}\alpha_{m,k-1}\\ \alpha_{m,k}\end{bmatrix}\right\|_\infty &\le CN\left(1+\frac{N}{|\Gamma|}\|Q_2\|_\infty\right)\max_j\left\|\begin{bmatrix}\Xi_j\\ \Theta_j\end{bmatrix}\right\|_\infty,\\ &\le Ch^k|u|_{W^{k+1,\infty}(I)}\left(1+\frac{h^{-1}\|Q_2\|_\infty}{|\Gamma|}\right)\left(\|V_1\|_\infty+h^{-1}\|V_2\|_\infty\right).\end{aligned}$$

Similar to Lemma 3.4, we can estimate $\|P_h^\star u-P_h^1u\|_{L^p(I)}$ and (22) follows.

A.3.3 Proof of Lemma 3.6

In Case 3, $\lambda_{1,2}$ are conjugate to each other and $|\lambda_{1,2}| = 1$. Therefore, $\delta' \geq 0$, and

$$\sum_{j=0}^{N-1} |d_1^j| = \sum_{j=0}^{N-1} \left| \frac{\lambda_1^j}{1-\lambda_1^N} \right| = \frac{N}{|1-\lambda_1^N|} = \sum_{j=0}^{N-1} |d_2^j|.$$

Similar to (68), we obtain

$$\begin{aligned}
\left\| \begin{bmatrix} \alpha_{m,k-1} \\ \alpha_{m,k} \end{bmatrix} \right\|_\infty &\leq \sum_{j=0}^{N-1} (|d_1^j| + |d_2^j|) \left(\max_j |\eta_j| \|Q_1 V_1\|_\infty + \max_j |\theta_j| \|Q_1 V_2\|_\infty \right) \\
&\quad + \sum_{j=0}^{N-1} |d_2^j| \left(\max_j |\eta_j| \|V_1\|_\infty + \max_j |\theta_j| \|V_2\|_\infty \right), \\
&\leq C h^{k+1} |u|_{W^{k+1,\infty}(I)} h^{-(\delta'+1)} \left(\|Q_1 V_1\|_\infty + h^{-1} \|Q_1 V_2\|_\infty + \|V_1\|_\infty + h^{-1} \|V_2\|_\infty \right) \\
&\leq C h^{k+1} |u|_{W^{k+1,\infty}(I)} h^{-(\delta'+1)} \|Q_1\|_\infty \left(\|V_1\|_\infty + h^{-1} \|V_2\|_\infty \right)
\end{aligned}$$

and we reach the estimation (25).

A.4 Proof of Lemma 3.8

From (64) and (67), we have

$$\begin{aligned}
\begin{bmatrix} \alpha_{m,k-1} \\ \alpha_{m,k} \end{bmatrix} &= \sum_{j=0}^{N-1} r_j \begin{bmatrix} \Xi_{j+m} \\ \Theta_{j+m} \end{bmatrix}, \quad m = 1, \ldots, N, \\
&=: U_1 V_1 + U_2 V_2,
\end{aligned}$$

where $U_1 = \sum_{j=0}^{N-1} r_j \eta_{j+m}$, $U_2 = \sum_{j=0}^{N-1} r_j \theta_{j+m}$. We first estimate U_1, then U_2 can be estimated in a similar way. From (71),

$$U_1 = \frac{1}{2} I_2 \sum_{j=0}^{N-1} (-1)^j \eta_{j+m} + \frac{Q_2}{2\Gamma} \sum_{j=0}^{N-1} (-1)^j \frac{-N+2j}{2} \eta_{j+m}. \tag{73}$$

By Lemma 3.7, the first term in (73) can be estimated by

$$\begin{aligned}
\left| \sum_{j=0}^{N-1} (-1)^j \eta_{j+m} \right| &= \left| \sum_{j'=0}^{\frac{N-1}{2}} (\eta_{2j'+m} - \eta_{2j'+1+m}) + \eta_{N-1+m} \right| \\
&\leq \frac{N-1}{2} C h^{k+2} |u|_{W^{k+2,\infty}(I)} + C h^{k+1} |u|_{W^{k+1,\infty}(I)} \\
&\leq C h^{k+1} \|u\|_{W^{k+2,\infty}(I)},
\end{aligned} \tag{74}$$

because N must be odd from Lemma 3.3. The second term in (73) can be estimated by using (26),

$$\left|\sum_{j=0}^{N-1}(-1)^j\frac{-N+2j}{2}\eta_{j+m}-\mu h^{k+1}S_{k+1}-\mu_2 h^{k+2}S_{k+2}\right|$$
$$\leq C\sum_{j=0}^{N-1}\left|\frac{-N+2j}{2}\right|h^{k+3}|u|_{W^{k+3,\infty}(I)}\leq Ch^{k+1}|u|_{W^{k+3,\infty}(I)}, \tag{75}$$

where S_{k+1}, S_{k+2} are defined as:

$$S_{k+\nu}:=\sum_{j=0}^{N-1}(-1)^j\frac{-N+2j}{2}u^{(k+\nu)}\left(x_{j+m+\frac{1}{2}}\right),\quad \nu=1,2.$$

We assume $u\in W^{k+4,1}(I)$. Then $u^{(k+1)}\in W^{3,1}(I)$, is periodic, and has the following Fourier series expansion $u^{(k+1)}(x)=\sum_{n=-\infty}^{\infty}\hat{f}(n)e^{2\pi inx/L}$, $L=b-a$, where its fourier coefficient $\hat{f}(n)$ satisfies:

$$\left|\hat{f}(n)\right|\leq C\frac{|u|_{W^{k+4,1}(I)}}{1+|n|^3}. \tag{76}$$

Since $x_{j+\frac{1}{2}}=j\Delta x=j\frac{L}{N}$, $j=0,\ldots,N-1$, then $u^{(k+1)}(x_{j+\frac{1}{2}})=\sum_{n=-\infty}^{\infty}\hat{f}(n)\omega^{jn}$ with $\omega=e^{i\frac{2\pi}{N}}$. Then

$$S_{k+1}=\sum_{j=0}^{N-1}(-1)^j\frac{-N+2j}{2}\sum_{n=-\infty}^{\infty}\hat{f}(n)\omega^{(j+m)n}.$$

Due to (76), $\sum|\hat{f}(n)|$ is convergent and we can switch the order of summation, which results in

$$S_{k+1}=\sum_{n=-\infty}^{\infty}\hat{f}(n)W(n),\quad \text{where}\quad W(n)=\frac{-2\omega^{(m+1)n}}{(1+\omega^n)^2}. \tag{77}$$

Since N is odd, $\omega^n=e^{2\pi i\frac{n}{N}}\neq -1,\forall n$. Hence, $W(n)$ and S_{k+1} are well defined. Because $W(n)$ is N-periodic, it's helpful to split S_{k+1} into blocks of size N as

$$S_{k+1}=\sum_{l=-\infty}^{\infty}S_{k+1}^l,\quad \text{where}\quad S_{k+1}^l=\sum_{n=lN-\frac{N-1}{2}}^{lN+\frac{N-1}{2}}\hat{f}(n)W(n).$$

Let's estimate S^0_{k+1} first. For $|n| \le [\frac{3N}{8}]$, $|W(n)| = \frac{2}{|1+\omega^n|^2} \le \frac{2}{|1+e^{i3\pi/4}|^2} = \frac{2}{2-\sqrt{2}}$. For other n, $|W(n)| \le |W(\frac{N-1}{2})| = \frac{2}{|1+\omega^{(N-1)/2}|^2} \le CN^2$ from Taylor expansions.

$$
\begin{aligned}
|S^0_{k+1}| &\le \sum_{n=-\left[\frac{3N}{8}\right]}^{\left[\frac{3N}{8}\right]} \left|\hat{f}(n)W(n)\right| + \sum_{-\frac{N-1}{2}}^{n=-\left[\frac{3N}{8}\right]-1} \left|\hat{f}(n)W(n)\right| + \sum_{n=\left[\frac{3N}{8}\right]+1}^{\frac{N-1}{2}} \left|\hat{f}(n)W(n)\right| \\
&\le \frac{2}{2-\sqrt{2}} \sum_{n=-\left[\frac{3N}{8}\right]}^{\left[\frac{3N}{8}\right]} \left|\hat{f}(n)\right| + CN^2 \sum_{-\frac{N-1}{2}}^{n=-\left[\frac{3N}{8}\right]-1} \left|\hat{f}(n)\right| + CN^2 \sum_{n=\left[\frac{3N}{8}\right]+1}^{\frac{N-1}{2}} \left|\hat{f}(n)\right| \\
&\le \frac{2}{2-\sqrt{2}} \sum_{n=-\left[\frac{3N}{8}\right]}^{\left[\frac{3N}{8}\right]} \left|\hat{f}(n)\right| + CN^2 \frac{1}{1+\left(\frac{3N}{8}\right)^3} \left(\frac{N}{4}+2\right) |u|_{W^{k+4,1}(I)} \\
&\le C \left(\sum_{n=-\frac{N-1}{2}}^{\frac{N-1}{2}} \frac{1}{1+|n|^3} + \frac{1}{\left(\frac{3}{8}\right)^3} \right) |u|_{W^{k+4,1}(I)}.
\end{aligned}
$$

Then, in a similar way,

$$
|S^l_{k+1}| \le C \left(\sum_{n=lN-\frac{N-1}{2}}^{lN+\frac{N-1}{2}} \frac{1}{1+|n|^3} + \frac{1}{\left(|l|+\frac{3}{8}\right)^3} \right) |u|_{W^{k+4,1}(I)}.
$$

Therefore,

$$
|S_{k+1}| \le C \left(\sum_{n=-\infty}^{\infty} \frac{1}{1+|n|^3} + \sum_{l=-\infty}^{\infty} \frac{1}{\left(|l|+\frac{3}{8}\right)^3} \right) |u|_{W^{k+4,1}(I)} \le C|u|_{W^{k+4,1}(I)}. \tag{78}
$$

By similar Fourier expansion technique, we can show

$$
|S_{k+2}| \le CN|u|_{W^{k+4,1}(I)} = Ch^{-1}|u|_{W^{k+4,1}(I)}. \tag{79}
$$

Combine (78), (79) with (73), (74) and (75), we get

$$
\|U_1\|_\infty \le Ch^{k+1}\|u\|_{W^{k+4,\infty}(I)} \left(1 + \frac{\|Q_2\|_\infty}{|\Gamma|}\right). \tag{80}
$$

Similarly, by (29) and the Fourier expansion technique

$$
\|U_2\|_\infty \le Ch^{k}\|u\|_{W^{k+4,\infty}(I)} \left(1 + \frac{\|Q_2\|_\infty}{|\Gamma|}\right). \tag{81}
$$

Therefore,

$$
\begin{aligned}
\left\| \begin{bmatrix} \alpha_{m,k-1} \\ \alpha_{m,k} \end{bmatrix} \right\|_\infty &\le \|U_1\|_\infty \|V_1\|_\infty + \|U_2\|_\infty \|V_2\|_\infty, \\
&\le Ch^{k+1}\|u\|_{W^{k+4,\infty}(I)} \left(1 + \frac{\|Q_2\|_\infty}{|\Gamma|}\right) (\|V_1\|_\infty + h^{-1}\|V_2\|_\infty), \quad m = 1, \ldots, N,
\end{aligned}
$$

and (30) is obtained.

A.5 Proof of Lemma 3.9

From the discussion in Lemma 3.3, we can write $\lambda_{1,2} = e^{\pm i\theta}$ and assume $\theta \in (0, \pi)$. First, we want to make clear of the conditions on δ, δ'. Since $|\lambda_1| = |\lambda_1^N| = 1$, we have $\delta, \delta' \geq 0$. Because $1 - \lambda_1^N = (\omega^n)^N - (e^{i\theta})^N = (\omega^n - e^{i\theta})(\sum_{l=0}^{N-1}(\omega^n)^{N-1-l}(e^{i\theta})^l)$, thus $\left|1 - \lambda_1^N\right| \leq N\left|\omega^n - e^{i\theta}\right|, \forall n$. With the assumption $\left|1 - \lambda_1^N\right| \sim Ch^{\delta'}$, we get $\left|\omega^n - e^{i\theta}\right| \geq Ch^{\delta'+1}$. Particularly, when $n = 0$, we have $|1 - \lambda_1| \geq Ch^{\delta'+1}$, hence $\delta/2 \leq \delta' + 1$.

Similar to (67) in Case 1, we can get

$$\begin{aligned}\begin{bmatrix}\alpha_{m,k-1}\\ \alpha_{m,k}\end{bmatrix} &= \sum_{j=0}^{N-1} r_j \begin{bmatrix}\Xi_{j+m}\\ \Theta_{j+m}\end{bmatrix}, \quad m = 1, \ldots, N,\\ &=: \mathcal{U}_1 V_1 + \mathcal{U}_2 V_2,\end{aligned}$$

where

$$\mathcal{U}_1 = Q_1 \sum_{j=0}^{N-1} \eta_{j+m} d_1^j + (I_2 - Q_1) \sum_{j=0}^{N-1} \eta_{j+m} d_2^j, \qquad d_1^j = \tfrac{e^{ij\theta}}{1-e^{iN\theta}},$$

$$\mathcal{U}_2 = Q_1 \sum_{j=0}^{N-1} \theta_{j+m} d_1^j + (I_2 - Q_1) \sum_{j=0}^{N-1} \theta_{j+m} d_2^j, \quad d_2^j = -d_1^{N-j} = \tfrac{-e^{i(N-j)\theta}}{1-e^{iN\theta}}.$$

We introduce:

$$\mathcal{S}_1 = \frac{1}{1 - e^{iN\theta}} \sum_{j=0}^{N-1} e^{ij\theta} u^{(k+1)}\left(x_{j+m+\frac{1}{2}}\right),$$

$$\mathcal{S}_2 = \frac{-1}{1 - e^{iN\theta}} \sum_{j=0}^{N-1} e^{i(N-j)\theta} u^{(k+1)}\left(x_{j+m-\frac{1}{2}}\right).$$

Then by Lemma 3.7:

$$\left|\mathcal{U}_1 - \mu h^{k+1} Q_1 \mathcal{S}_1 - \mu h^{k+1}(I_2 - Q_1)\mathcal{S}_2\right| \leq Ch^{k+1}(1 + \|Q_1\|_\infty)|u|_{W^{k+2,\infty}(I)},$$

$$\left|\mathcal{U}_2 - \rho h^k Q_1 \mathcal{S}_1 - \rho h^k (I_2 - Q_1)\mathcal{S}_2\right| \leq Ch^k(1 + \|Q_1\|_\infty)|u|_{W^{k+2,\infty}(I)}.$$

Therefore,

$$|\mathcal{U}_\nu| \leq Ch^{k+2-\nu}\left((1 + \|Q_1\|_\infty)|u|_{W^{k+2,\infty}} + \|Q_1\|_\infty(1 + \max(|\mathcal{S}_1|, |\mathcal{S}_2|))\right), \quad \nu = 1, 2. \tag{82}$$

By using similar Fourier expansion: $u^{(k+1)}(x_{j+\frac{1}{2}}) = \sum_{n=-\infty}^{\infty} \hat{f}(n)\omega^{jn}$. Since now we assume $u \in W^{k+3,\infty}(I)$, $\left|\hat{f}(n)\right| \leq C\frac{1}{1+|n|^2}|u|_{W^{k+3,1}(I)}$.

$$\mathcal{S}_1 = \frac{1}{1 - e^{iN\theta}} \sum_{n=-\infty}^{\infty} \hat{f}(n) \sum_{j=0}^{N-1} e^{ij\theta}\omega^{(j+m)n} = \sum_{n=-\infty}^{\infty} \hat{f}(n)\mathcal{W}_1(n),$$

$$\mathcal{S}_2 = \frac{-1}{1 - e^{iN\theta}} \sum_{n=-\infty}^{\infty} \hat{f}(n) \sum_{j=0}^{N-1} e^{i(N-j)\theta}\omega^{(j+m-1)n} = \sum_{n=-\infty}^{\infty} \hat{f}(n)\mathcal{W}_2(n),$$

where from simple algebra

$$\mathcal{W}_1(n) = \frac{\omega^{mn}}{1 - e^{i\theta}\omega^n}, \quad \mathcal{W}_2(n) = \frac{\omega^{(m-1)n}}{1 - e^{-i\theta}\omega^n}.$$

From the discussion at the beginning of the proof, we have $|\mathcal{W}_2(n)| = |\lambda_1 - \omega^n|^{-1} \leq Ch^{-(\delta'+1)}$, and similarly $|\mathcal{W}_1(n)| \leq Ch^{-(\delta'+1)}$. Since $\mathcal{S}_1$ and $\mathcal{S}_2$ can be estimated in the same way, we only show details for $\mathcal{S}_2$ in what follows. Similar to the proof of Lemma 3.8, we split $\mathcal{S}_2$ into blocks of size N,

$$\mathcal{S}_2 = \sum_{l=-\infty}^{\infty} \mathcal{S}_2^l, \quad \text{where} \quad \mathcal{S}_2^l = \sum_{n=lN}^{(l+1)N-1} \hat{f}(n)\mathcal{W}_2(n).$$

With the assumption that $0 \leq \delta/2 \leq 1$, there $\exists n_0 \sim O(h^{\delta/2-1})$ s.t. $2\pi\frac{n_0}{N} \leq \theta < 2\pi\frac{n_0+1}{N}$. Let $n_1 = \lfloor n_0/2 \rfloor$, $n_2 = 2n_0 - n_1$, then for $n_1 \leq n \leq n_2$, $\left|\hat{f}(n)\right| \leq C\frac{1}{1+n_1^2}|u|_{W^{k+3,1}(I)}$. For other n, $|\mathcal{W}_2(n)| \leq |\mathcal{W}_2(n_1)| \leq \frac{1}{2|\sin(\pi n_1/N - \theta/2)|} \leq Ch^{-\delta/2}$. Thus,

$$\begin{aligned}
\left|\mathcal{S}_2^0\right| &\leq Ch^{-\delta/2}\left(\sum_{n=0}^{n_1-1} + \sum_{n=n_2+1}^{N-1}\left|\hat{f}(n)\right|\right) + Ch^{-(\delta'+1)}\sum_{n=n_1}^{n_2}\left|\hat{f}(n)\right| \\
&\leq C\left(h^{-\delta/2}\sum_{n=0}^{N-1}\frac{1}{1+|n|^2} + h^{-(\delta'+1)}(n_2 - n_1 + 1)\frac{1}{1+n_1^2}\right)|u|_{W^{k+3,1}(I)} \\
&\leq C\left(h^{-\delta/2}\sum_{n=0}^{N-1}\frac{1}{1+|n|^2} + h^{-(\delta'+1)}h^{\delta/2-1}h^{2-\delta}\right)|u|_{W^{k+3,1}(I)} \\
&\leq C\left(h^{-\delta/2}\sum_{n=0}^{N-1}\frac{1}{1+|n|^2} + h^{-\delta'-\delta/2}\right)|u|_{W^{k+3,1}(I)}.
\end{aligned}$$

Using similar approaches, for $l \neq 0$,

$$\left|\mathcal{S}_2^l\right| \leq C\left(h^{-\delta/2}\sum_{n=lN}^{(l+1)N-1}\frac{1}{1+|n|^2} + h^{-\delta'+\delta/2}\frac{1}{|n_1/N + l|^2}\right)|u|_{W^{k+3,1}(I)}.$$

Summing up, we reach the estimation

$$\begin{aligned}
|\mathcal{S}_2| &\leq C\left(h^{-\delta/2}\sum_{n=-\infty}^{\infty}\frac{1}{1+|n|^2} + h^{-\delta'-\delta/2} + h^{-\delta'+\delta/2}\sum_{l\in\mathbb{N},l\neq 0}\frac{1}{|l|^2}\right)|u|_{W^{k+3,1}(I)} \\
&\leq Ch^{-\delta'-\delta/2}|u|_{W^{k+3,1}(I)}. \qquad (83)
\end{aligned}$$

Similarly, we obtain

$$|\mathcal{S}_1| \leq Ch^{-\delta'-\delta/2}|u|_{W^{k+3,1}(I)}. \qquad (84)$$

Combine (83), (84) and (82), we get

$$|\mathcal{U}_\nu| \leq Ch^{k+2-\nu}(1 + h^{-(\delta'+\delta/2)}\|Q_1\|_\infty)\|u\|_{W^{k+3,\infty}(I)}, \quad \nu = 1, 2,$$

and (31) follows.

A.6 Detailed Discussions on the Choice of the *T* Matrix as in (54) or (55)

We discuss what parameters result in $|b_1 \pm b_2| = 0$, under the assumption that α_1 has no dependence on h, $\beta_1 = \tilde{\beta}_1 h^{A_1}$, $\beta_2 = \tilde{\beta}_2 h^{A_2}$, $\tilde{\beta}_1$, $\tilde{\beta}_2$ are nonzero constants that do not depend on h.

$$\begin{aligned} b_1 - b_2 &= \left(-\beta_1 + \frac{k(k-1)}{2h}\right)\left(1 - \beta_2 \frac{2k(k-1)}{h}\right) + \frac{k(k-1)}{h} 2\alpha_1^2 \\ &= \left(-\tilde{\beta}_1 h^{A_1} + \frac{k(k-1)}{2} h^{-1}\right)\left(1 - 2k(k-1)\tilde{\beta}_2 h^{A_2-1}\right) + k(k-1)2\alpha_1^2 h^{-1}, \\ b_1 + b_2 &= \left(-\beta_1 + \frac{k(k+1)}{2h}\right)\left(1 - \beta_2 \frac{2k(k+1)}{h}\right) + \frac{k(k+1)}{h} 2\alpha_1^2 \\ &= \left(-\tilde{\beta}_1 h^{A_1} + \frac{k(k+1)}{2} h^{-1}\right)\left(1 - 2k(k+1)\tilde{\beta}_2 h^{A_2-1}\right) + k(k+1)2\alpha_1^2 h^{-1}. \end{aligned}$$

If $b_1 - b_2 = 0, \forall h < h_0$, then

- $\alpha_1 \neq 0$, then $A_1 = -1$, $A_2 = 1$ and $\tilde{\beta}_1, \tilde{\beta}_2$ satisfies

$$\left(-\tilde{\beta}_1 + \frac{k(k-1)}{2}\right)(1 - 2k(k-1)\tilde{\beta}_2) + k(k-1)2\alpha_1^2 = 0. \tag{85}$$

Similarly, for $b_1 + b_2 = 0, \forall h < h_0$, then

- $\alpha_1 \neq 0$, $A_1 = -1$, $A_2 = 1$ and $\tilde{\beta}_1, \tilde{\beta}_2$ satisfies

$$\left(-\tilde{\beta}_1 + \frac{k(k+1)}{2}\right)(1 - 2k(k+1)\tilde{\beta}_2) + k(k+1)2\alpha_1^2 = 0. \tag{86}$$

A.7 Detailed Discussions on Case 2

Parameter choices for $|\Gamma| = |\Lambda|$ imply

$$\begin{aligned} \Gamma \pm \Lambda &= \beta_1 + \frac{k^2(k^2-1)}{h^2}\beta_2 + \frac{k(k \pm 1)}{h}(-2\alpha_1^2 - 2\beta_1\beta_2) + \frac{-k^2 \pm k}{2h} \\ &= \left(\beta_1 - \frac{k(k \mp 1)}{2h}\right)\left(1 - 2\beta_2 \frac{k(k \pm 1)}{h}\right) - \frac{k(k \pm 1)}{h} 2\alpha_1^2 = 0, \end{aligned}$$

which indicates

- if $\alpha_1 \neq 0$, then $b_1 \pm b_2$ can be greatly simplified as follows.
  - If $\Gamma + \Lambda = 0$, then k is odd from Lemma 3.3, and

$$\begin{aligned} b_1 + b_2 &= \frac{k}{h}\left(1 - \beta_2 \frac{2k(k+1)}{h}\right), \\ b_1 - b_2 &= -\frac{2}{k+1}\left(\beta_1 - \frac{k(k-1)}{2h}\right), \\ \Lambda &= -\frac{1}{k+1}\left(\beta_1 - \frac{k^2}{h} + \frac{k^2(k^2-1)}{h^2}\beta_2\right). \end{aligned}$$

– If $\Gamma - \Lambda = 0$, then k is even from Lemma 3.3, and

$$b_1 + b_2 = \frac{2}{k-1}\left(\beta_1 - \frac{k(k+1)}{2h}\right),$$
$$b_1 - b_2 = -\frac{k}{h}\left(1 - \beta_2 \frac{2k(k-1)}{h}\right),$$
$$\Lambda = -\frac{1}{k-1}\left(\beta_1 - \frac{k^2}{h} + \frac{k^2(k^2-1)}{h^2}\beta_2\right), \quad k > 1.$$

- If $\alpha_1 = 0$, then

$$\beta_1 = \frac{k(k \pm 1)}{2h}, \quad \text{or } \beta_2 = \frac{h}{2k(k \pm 1)}. \tag{87}$$

References

1. Ainsworth, M., Monk, P., Muniz, W.: Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. J. Sci. Comput. **27**(1), 5–40 (2006)
2. Ascher, U.M., Ruuth, S.J., Spiteri, R.J.: Implicit–explicit Runge–Kutta methods for time-dependent partial differential equations. Appl. Numer. Math. **25**(2–3), 151–167 (1997)
3. Bona, J., Chen, H., Karakashian, O., Xing, Y.: Conservative, discontinuous Galerkin-methods for the generalized Korteweg–de Vries equation. Math. Comput. **82**(283), 1401–1432 (2013)
4. Cessenat, O., Despres, B.: Application of an ultra weak variational formulation of elliptic pdes to the two-dimensional Helmholtz problem. SIAM J. Numer. Anal. **35**(1), 255–299 (1998)
5. Chang, Q., Jia, E., Sun, W.: Difference schemes for solving the generalized nonlinear Schrödinger equation. J. Comput. Phys. **148**(2), 397–415 (1999)
6. Cheng, Y., Chou, C.-S., Li, F., Xing, Y.: L^2 stable discontinuous Galerkin methods for one-dimensional two-way wave equations. Math. Comput. **86**(303), 121–155 (2017)
7. Cheng, Y., Shu, C.-W.: A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives. Math. Comput. **77**(262), 699–730 (2008)
8. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems, Studies in Mathematics and Its Applications, vol. 4. North-Holland Publishing Co., Amsterdam (1978)
9. Cockburn, B., Hou, S., Shu, C.-W.: The Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. Math. Comput. **54**(190), 545–581 (1990)
10. Cockburn, B., Lin, S.-Y., Shu, C.-W.: TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. J. Comput. Phys. **84**(1), 90–113 (1989)
11. Cockburn, B., Shu, C.-W.: TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. Math. Comput. **52**(186), 411–435 (1989)
12. Cockburn, B., Shu, C.-W.: The local discontinuous Galerkin method for time-dependent convection–diffusion systems. SIAM J. Numer. Anal. **35**(6), 2440–2463 (1998)
13. Cockburn, B., Shu, C.-W.: The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. J. Comput. Phys. **141**(2), 199–224 (1998)
14. Dag, I.: A quadratic b-spline finite element method for solving nonlinear Schrödinger equation. Comput. Methods Appl. Mech. Eng. **174**(1–2), 247–258 (1999)
15. Douglas J., Dupont T.: Interior penalty procedures for elliptic and parabolic Galerkin methods. In: Glowinski, R., Lions, J.L. (eds.) Computing Methods in Applied Sciences. Lecture Notes in Physics, vol 58. Springer, Berlin, Heidelberg (1976)
16. Grote, M.J., Schneebeli, A., Schötzau, D.: Discontinuous Galerkin finite element method for the wave equation. SIAM J. Numer. Anal. **44**(6), 2408–2431 (2006)
17. Karakashian, O., Makridakis, C.: A space–time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method. Math. Comput. **67**(222), 479–499 (1998)
18. Karakashian, O., Makridakis, C.: A space–time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method. SIAM J. Numer. Anal. **36**(6), 1779–1807 (1999)

19. Liang, X., Khaliq, A.Q.M., Xing, Y.: Fourth order exponential time differencing method with local discontinuous Galerkin approximation for coupled nonlinear Schrödinger equations. Commun. Comput. Phys. **17**(2), 510–541 (2015)
20. Liu, H.: Optimal error estimates of the direct discontinuous Galerkin method for convection–diffusion equations. Math. Comput. **84**(295), 2263–2295 (2015)
21. Lu, W., Huang, Y., Liu, H.: Mass preserving discontinuous Galerkin methods for Schrödinger equations. J. Comput. Phys. **282**, 210–226 (2015)
22. Meng, X., Shu, C.-W., Wu, B.: Optimal error estimates for discontinuous Galerkin methods based on upwind-biased fluxes for linear hyperbolic equations. Math. Comput. **85**(299), 1225–1261 (2016)
23. Pathria, D., Morris, J.L.: Pseudo-spectral solution of nonlinear Schrödinger equations. J. Comput. Phys. **87**(1), 108–125 (1990)
24. Reed, W.H., Hill, T.: Triangular mesh methods for the neutron transport equation. Technical report, Los Alamos Scientific Lab., N. Mex. (USA) (1973)
25. Riviere, B., Wheeler, M.F.: Discontinuous finite element methods for acoustic and elastic wave problems. Contemp. Math. **329**, 271–282 (2003)
26. Rivière, B., Wheeler, M.F., Girault, V.: Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. Comput. Geosci. **3**(3), 337–360 (1999)
27. Rivière, B., Wheeler, M.F., Girault, V.: A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. SIAM J. Numer. Anal. **39**(3), 902–931 (2001)
28. Sheng, Q., Khaliq, A., Al-Said, E.: Solving the generalized nonlinear Schrödinger equation via quartic spline approximation. J. Comput. Phys. **166**(2), 400–417 (2001)
29. Shu, C.-W.: Discontinuous Galerkin methods for time-dependent convection dominated problems: basics, recent developments and comparison with other methods. In: Barrenechea, G., Brezzi, F., Cangiani, A., Georgoulis, E. (eds.) Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Lecture Notes in Computational Science and Engineering, vol 114. Springer, Cham (2016)
30. Taha, T.R., Ablowitz, M.I.: Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical, nonlinear Schrödinger equation. J. Comput. Phys. **55**(2), 203–230 (1984)
31. Wheeler, M.F.: An elliptic collocation-finite element method with interior penalties. SIAM J. Numer. Anal. **15**(1), 152–161 (1978)
32. Xu, Y., Shu, C.-W.: Local discontinuous Galerkin methods for nonlinear Schrödinger equations. J. Comput. Phys. **205**(1), 72–97 (2005)
33. Xu, Y., Shu, C.-W.: Optimal error estimates of the semidiscrete local discontinuous Galerkin methods for high order wave equations. SIAM J. Numer. Anal. **50**(1), 79–104 (2012)