INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 19, Number 6, Pages 864–886

AN ARBITRARY-ORDER ULTRA-WEAK DISCONTINUOUS GALERKIN METHOD FOR TWO-DIMENSIONAL SEMILINEAR SECOND-ORDER ELLIPTIC PROBLEMS ON CARTESIAN GRIDS

MAHBOUB BACCOUCH

This paper is dedicated to my mother, Rebah Jetlaoui, who passed away from COVID in July 31, 2021 while I was completing this work.

Abstract. In this paper, we present and analyze a new ultra-weak discontinuous Galerkin (UWDG) finite element method for two-dimensional semilinear second-order elliptic problems on Cartesian grids. Unlike the traditional local discontinuous Galerkin (LDG) method, the proposed UWDG method can be applied without introducing any auxiliary variables or rewriting the original equation into a system of equations. The UWDG scheme is presented in details, including the definition of the numerical fluxes, which are necessary to obtain optimal error estimates. The proposed scheme can be made arbitrarily high-order accurate in two-dimensional space. The error estimates of the presented scheme are analyzed. The order of convergence is proved to be $p+1$ in the L^2-norm, when tensor product polynomials of degree at most p and grid size h are used. Several numerical examples are provided to confirm the theoretical results.

Key words. Ultra-weak discontinuous Galerkin method; elliptic problems; convergence; *a priori* error estimation.

1. Introduction

In this paper, we develop a new ultra-weak discontinuous Galerkin (UWDG) finite element method for the semilinear second-order elliptic problems of the form

$$-\Delta u + f(\mathbf{x}, u) = \mathbf{0}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d, \quad d = 1,\ 2,\ 3. \tag{1a}$$

We shall assume that the nonlinear function $f(\mathbf{x}, u) : \ \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ is smooth with respect to its arguments $\mathbf{x}$ and u. To be more precise, we assume that f and its partial derivatives are continuous for $\mathbf{x} \in \overline{\Omega}$ and $u \in \mathbb{R}$ and satisfies the uniform bound

$$|f(\mathbf{x}, u)| \le M, \quad \forall\, \mathbf{x} \in \Omega, \quad \forall\, u \in \mathbb{R}, \tag{1b}$$

as well as the Lipschitz condition

$$|f_u(\mathbf{x}, u) - f_u(\mathbf{y}, v)| \le L\left(|\mathbf{x} - \mathbf{y}| + |u - v|\right), \quad \forall\, \mathbf{x},\ \mathbf{y} \in \Omega, \quad \forall\, u,\ v \in \mathbb{R}. \tag{1c}$$

For simplicity, we focus on two dimensions ($d = 2$) and write $\mathbf{x}$ as (x, y). In our analysis, we consider a rectangular domain denoted by $\Omega = \{\mathbf{x} = (x, y) : \ a < x < b,\ c < y < d\}$. We remark that our results remain true, with minor changes in the proofs, when the region Ω is a rectangular bounded domain of $\mathbb{R}^3$. In this paper,

Received by the editors May 4, 2022 and, in revised form, June 23, 2022; accepted June 28, 2022.

2000 *Mathematics Subject Classification.* 65N12, 65N15, 65N30.

we will consider either periodic boundary conditions

$$\begin{aligned} &u(a,y) = u(b,y),\ u(x,c) = u(x,d),\\ &u_x(a,y) = u_x(b,y), \quad u_y(x,c) = u_y(x,d),\ \mathbf{x} \in \partial\Omega, \end{aligned} \tag{1d}$$

or mixed Dirichlet-Neumann boundary conditions

$$u = g_D, \quad \mathbf{x} \in \partial\Omega_D, \quad \mathbf{n} \cdot \nabla u = \mathbf{n} \cdot \mathbf{g}_N, \quad \mathbf{x} \in \partial\Omega_N, \tag{1e}$$

or purely Dirichlet boundary conditions

$$u = g_D, \quad \mathbf{x} \in \partial\Omega. \tag{1f}$$

Here, $\mathbf{n}$ is the outward unit normal to the boundary, $\partial\Omega$, of Ω. For the mixed boundary conditions (1e), we always assume that the boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ is decomposed into two disjoint sets $\partial\Omega_D$ and $\partial\Omega_N$ where Dirichlet and Neumann boundary conditions are imposed, respectively. We further assume that the measure of $\partial\Omega_D$ is nonzero. In our analysis, we assume that the given functions f, g_D, and $\mathbf{g}_N$ are smooth functions on their domains such that the problem (1) has one and only one solution $u \in H^2(\Omega)$. We refer the reader to [23, 25, 29] and references therein for the existence and uniqueness of solutions to elliptic problems.

The origin of the discontinuous Galerkin (DG) finite element method (FEM) can be traced back to [32, 34] where it has been introduced for discretizing the neutron transport equation. Since then various types of DG schemes have been successfully used to discretize differential equations containing higher order spatial derivatives. DG methods for elliptic problems have been introduced in the late 90's. They are by now well-understood and rigorously analyzed in the context of linear elliptic problems (cf. [5] for the Poisson problem). The most successful DG schemes include symmetric interior penalty DG (SIPG) methods, non-symmetric interior penalty DG (NIPG) methods, local DG (LDG) methods, direct DG (DDG) methods, and ultra-weak DG (UWDG) methods. The class of SIPG methods (introduced in [4, 35]) and the class of NIPG methods (considered in [14]) are important methods for higher order differential equations. Some of the general attractive features of these methods are the local and high order of approximation, the flexibility due to local mesh refinement and the ability to handle unstructured meshes and discontinuous coefficients. The SIPG and NIPG methods use penalties to enforce weakly both continuity of the solution and the boundary conditions. The LDG method was first introduced to solve general convection-diffusion problems by Cockburn and Shu [21]. Nowadays, the LDG method has been successfully used in solving many linear and nonlinear problems. The key idea of the LDG method is to first rewrite the equation with higher order derivatives into a first order system, then apply the standard DG method on the system by properly choosing the so-called numerical fluxes. The DDG method was first introduced by Liu and Yan [33]. It involves the interior penalty methodology since the scheme is based on the direct weak formulation. Unlike the LDG method, the DDG method is based on the direct weak formulation and the construct of the suitable numerical flux on the cell edges. This method is called DDG since it does not introduce any auxiliary variables in contrast to the LDG.

The class of UWDG methods are proposed in [18]. These methods are based on repeated integration by parts so that all spatial derivatives are shifted from the exact solution to the test function in the weak formulation. Unlike the LDG method, the UWDG method can be applied without introducing any auxiliary variables or rewriting the original equation into a larger system. In [18], Cheng and Shu developed several UWDG methods to solve the third-order generalized KdV equation, the second-order convection-diffusion equation, the fourth-order biharmonic equation, and some fifth-order equations. They used the UWDG method in space and then they used the total variation diminishing (TVD) high-order Runge-Kutta method to discretize the resulting systems of ODEs in time. For each case, they proved the stability of the semi-discrete schemes by a careful choice of interface numerical fluxes. Their error estimates are sub-optimal. However, their numerical examples show that the scheme attains the optimal $(p+1)$-th order of accuracy. In the current work we design an UWDG method for the elliptic equation (1) and we investigate its convergence properties. To the best of our knowledge, this is the first paper to analyze an UWDG method for the elliptic equation (1).

In [2], Adjerid and Temimi proposed and analyzed a new UWDG finite element method to solve initial-value problems (IVPs) for ordinary differential equations (ODEs). They proved that the UWDG solution exhibits an optimal convergence rate in the L^2-norm. The order of convergence is proved to be of order $p+1$, when piecewise polynomials of degree at most p are used. They further showed that the p-degree UWDG solution of mth order ODEs and its first $m-1$ derivatives are superconvergent with order $2p+2-m$ at the end of each step. Also, they established that the p-degree discontinuous solution is superconvergent with order $p+2$ at the roots of $(p+1-m)$-degree Jacobi polynomial on each step. Finally, as an application of the superconvergence results, they constructed asymptotically exact *a posteriori* error estimates. Later, Adjerid and Temimi [3] presented a new DG method for solving the second-order wave equation (1) using the standard continuous finite element method in space and the UWDG method in time. They proved several optimal *a priori* error estimates in space-time norms for this new method and showed that it can be more efficient than existing methods. They also expressed the leading term of the local discretization error in terms of Lobatto polynomials in space and Jacobi polynomials in time which leads to superconvergence points on each space-time cell. Finally, they constructed efficient and asymptotically exact *a posteriori* estimates for space-time discretization errors.

In [13], Baccouch and Temimi studied the convergence and superconvergence properties of the UWDG method for a linear two-point boundary-value problem (BVP). They proved that the UWDG solution and its derivative exhibit optimal $\mathcal{O}(h^{p+1})$ and $\mathcal{O}(h^p)$ convergence rates in the L^2-norm, respectively, when p-degree piecewise polynomials with $p \geq 1$ are used. They further proved that the p-degree UWDG solution and its derivative are $\mathcal{O}(h^{2p})$ superconvergent at the downwind and upwind points, respectively. They observed optimal rates of convergence and superconvergence even in the presence of boundary layers when Shishkin meshes are used. More recently, Chen *et al.* [17] developed an UWDG method to solve the one-dimensional nonlinear Schrödinger equation. They provided stability conditions and error estimates for the scheme with a general class of numerical fluxes.

In this current work, we develop and analyze a new high order UWDG for the model (1). Convergence of the proposed scheme is rigorously analyzed. In particular, if h denotes the mesh size, then the UWDG solution u_h is shown to converge to the true solution u at $\mathcal{O}(h^{p+1})$ rate in the L^2-norm, when tensor product polynomials of degree at most p and grid size h are used. We would like to mention that the proposed scheme has several advantages over existing numerical methods such as those proposed in [6, 7, 15, 24, 26, 27, 31] as it leads to a smaller fully discrete problem for the solution while maintaining optimal convergence rates for both the solution and its first-order derivative. Although our error analysis is presented for the two-dimensional Poisson's equation, it can be readily extended to the three-dimensional Poisson's equation.

The rest of the paper is organized as follows. In Section 2, we present the UWDG method for the elliptic problem (1). In Section 3, we present an *a priori* error analysis for the two-dimensional elliptic problem. In Section 4, we present numerical results for the two-dimensional elliptic problems to confirm the theoretical results. Finally, we give concluding remarks in Section 5.

2. The UWDG method and Preliminaries

This section is devoted to the definition of the UWDG method. We also provide some notation and projection results needed for our *a priori* error estimates in Section 3.

2.1. The UWDG scheme. Here, we define the finite element spaces and proceed to construct the UWDG scheme. Let $\mathcal{T}_h$ be Cartesian mesh of the domain $\Omega = [a,b]\times[c,d]$. We assume that the mesh consists of $N = n\times m$ rectangular elements $K = I_i\times J_j$, where $I_i = [x_{i-1}, x_i],\ i = 1, 2, \dots, n$ and $J_j = [y_{j-1}, y_j],\ j = 1, 2, \dots, m$ where

$$a = x_0 < x_1 < \cdots < x_n = b, \quad c = y_0 < y_1 < \cdots < y_m = d.$$

For each rectangle $K \in \mathcal{T}_h$, we denote the mesh sizes as $h_i = x_i - x_{i-1}$ and $k_j = y_j - y_{j-1}$. The maximal mesh size is denoted by $h = \max\limits_{1\le i\le n,\ 1\le j\le m}(h_i, k_j)$. In this paper, we assume the mesh $\mathcal{T}_h$ is a shape regular triangulation of Ω, characterized by a small parameter h, namely that there exists a constant $c_0 > 0$ such that $c_K \ge c_0$, $\forall\ K \in \mathcal{T}_h$. Here c_K is the so-called chunkiness parameter defined by $c_K = h_K/d_K$, where $h_K = \max\limits_{1\le i\le n,\ 1\le j\le m}(h_i, k_j)$ is the local mesh size defined as the length of the longest edge of the element K and d_K is the diameter of the inscribed circle.

The UWDG weak formulation is obtained by multiplying equation (1a) by sufficiently smooth test function v, integrating over an arbitrary element $K \in \mathcal{T}_h$, and using Green's formula

$$(2)\quad -\iint_K u\Delta v\, dxdy + \int_{\partial K} \mathbf{n}\cdot\nabla v\, u\, ds - \int_{\partial K} \mathbf{n}\cdot\nabla u\, v\, ds + \iint_K f(x,y,u)\, v\, dxdy = 0,$$

where $\mathbf{n}$ is the outward normal unit vector to the boundary ∂K of the element $K \in \mathcal{T}_h$.

Let $\mathbb{P}^p(I_i)$ and $\mathbb{P}^p(J_j)$ be the spaces of polynomials of degree at most p on the intervals I_i and J_j, respectively. We define the piecewise polynomial finite element space V_h^p as the space of tensor product of $\mathbb{P}^p(I_i)$ and $\mathbb{P}^p(J_j)$, that is

$$V_h^p = \{v \in L^2(\Omega):\ \Omega \to \mathbb{R}|\ \ v|_K \in \mathcal{Q}^p(K),\ \forall\ K \in \mathcal{T}_h\},$$

where $\mathcal{Q}^p(K) = \mathbb{P}^p(I_i)\otimes\mathbb{P}^p(J_j)$. We note that V_h^p will be used for both our trial and test spaces.

The UWDG method is formulated as: find $u_h \in V_h^p$ such that $\forall\ v \in V_h^p$

(3a)

$$-\iint_K u_h\Delta v\, dxdy + \int_{\partial K} \mathbf{n}\cdot\nabla v\, \widehat{u}_h\, ds - \int_{\partial K} \mathbf{n}\cdot\widehat{\nabla u_h}\, v\, ds + \iint_K f(x,y,u_h)\, v\, dxdy = 0,$$

for all $K \in \mathcal{T}_h$. The "hat" quantities $\widehat{u}_h$ and $\widehat{\nabla u_h}$ are the so-called numerical fluxes. They take either the value from one side of the interface (namely inside or outside of the element K) or some linear combination of the values from both sides of the interface. The numerical fluxes need to be designed suitably to ensure consistency, stability, and convergence.

To define the numerical fluxes $\widehat{u}_h$ and $\widehat{\nabla u_h}$ on the boundary Γ, we introduce some definitions and notations. For $y \in J_j$, we let $v^+(x_i, y)$ and $v^-(x_i, y)$ be the values of the function v at the point (x_i, y) from the right element $I_{i+1} \times J_j$ and from the left element $I_i \times J_j$, respectively. Similarly, for $x \in I_i$, we use $v^+(x, y_j)$ and $v^-(x, y_j)$ to denote the values of v at the point (x, y_j) from the top element $I_i \times J_{j+1}$ and from the bottom element $I_i \times J_j$, respectively, *i.e.*, for $i = 0, 1, \ldots, n$ and $j = 0, 1, \ldots, m$

$$v^{\pm}(x_i, y) = v(x_i^{\pm}, y), \quad y \in J_j, \quad v^{\pm}(x, y_j) = v(x, y_j^{\pm}), \quad x \in I_i.$$

Let K^+ and K^- be two adjacent rectangle elements of the partition $\mathcal{T}_h$. Consider an arbitrary point (x, y) of the edge $\Gamma = K^+ \cap K^-$ sharing K^+ and K^-, and use $\mathbf{n}^{\pm}$ to denote the corresponding outward unit normal vectors at that point. Let $(v^{\pm}, \mathbf{w}^{\pm})$ be the traces of $(v, \mathbf{w})$ on Γ from the interior of the element $K^{\pm}$. The mean values $\{\cdot\}$ and jumps $[\![\cdot]\!]$ of a scalar-valued function $v \in V_h^p$ and a vector-valued function $\mathbf{w} \in (V_h^p)^2$ at the point $(x, y) \in \Gamma$ are defined as

$$\{v\} = \frac{1}{2}(v^+ + v^-), \quad \{\mathbf{w}\} = \frac{1}{2}(\mathbf{w}^+ + \mathbf{w}^-),$$
$$[\![v]\!] = v^+\mathbf{n}^+ + v^-\mathbf{n}^-, \quad [\![\mathbf{w}]\!] = \mathbf{n}^+ \cdot \mathbf{w}^+ + \mathbf{n}^- \cdot \mathbf{w}^-.$$

Now, we are in the position to introduce the numerical fluxes [11].

- Numerical fluxes associated with the periodic boundary conditions (1d): We use the following alternating numerical fluxes

$$\widehat{u}_h = u_h^- \quad \text{and} \quad \mathbf{n} \cdot \widehat{\nabla u_h} = \mathbf{n} \cdot (\nabla u_h)^+. \tag{3b}$$

  We remark that this choice is not particularly restrictive. For example, the other choice $\hat{u}_h = u_h^+$ and $\mathbf{n} \cdot \widehat{\nabla u_h} = \mathbf{n} \cdot (\nabla u_h)^-$ can be used.
- Numerical fluxes associated with the mixed boundary conditions (1e): Let $\mathbf{v}$ be a fixed vector that is not parallel to any normals of element interfaces. It is used to define artificial inflow and outflow boundaries of the domain Ω. The vector $\mathbf{v}$ is employed to provide a single rule for selecting the numerical fluxes $\widehat{u}_h$ and $\widehat{\nabla u_h}$. For simplicity, we choose $\mathbf{v} = [1, 1]^t$. Then, we define the following artificial inflow boundary $\partial\Omega^-$ and the artificial outflow boundary $\partial\Omega^+$ as

$$\partial\Omega^- = \{(x, y) \in \partial\Omega \mid \mathbf{n} \cdot \mathbf{v} \le 0\} = \partial\Omega_1^- \cup \partial\Omega_2^-,$$
$$\partial\Omega^+ = \{(x, y) \in \partial\Omega \mid \mathbf{n} \cdot \mathbf{v} > 0\} = \partial\Omega_1^+ \cup \partial\Omega_2^+,$$

  where $\partial\Omega_1^-$, $\partial\Omega_2^-$, $\partial\Omega_1^+$, and $\partial\Omega_2^+$ are, respectively, the left, bottom, right, and top edges of the physical domain Ω. We also define the inflow boundary Γ^- and the outflow boundary Γ^+ of each rectangle element $K \in \mathcal{T}_h$ as

$$\Gamma^- = \{(x, y) \in \Gamma \mid \mathbf{n} \cdot \mathbf{v} \le 0\} = \Gamma_1^- \cup \Gamma_2^-, \quad \Gamma^+ = \{(x, y) \in \Gamma \mid \mathbf{n} \cdot \mathbf{v} > 0\} = \Gamma_1^+ \cup \Gamma_2^+,$$

  where Γ_1^-, Γ_2^-, Γ_1^+, and Γ_2^+ are, respectively, the left, bottom,right, and top edges of the rectangle K.
  Now, we are ready to define the numerical fluxes associated with the mixed Dirichlet-Neumann boundary conditions (1e): If Γ is an interior edge then we take

$$\widehat{u}_h = u_h^- \quad \text{and} \quad \mathbf{n} \cdot \widehat{\nabla u_h} = \mathbf{n} \cdot (\nabla u_h)^+, \tag{3ca}$$

and if the edge Γ lies on $\partial\Omega$, we take

(3cb)
$$\hat{u}_h = \begin{cases} \Pi_h^- g_D, \ (x,y) \in \partial\Omega_D, \\ u_h^-, \ (x,y) \in \partial\Omega_N \cap \partial\Omega^+, \\ u_h^+, \ (x,y) \in \partial\Omega_N \cap \partial\Omega^-, \end{cases} \quad \mathbf{n} \cdot \widehat{\nabla u_h} = \begin{cases} \Pi_h^-(\mathbf{n} \cdot \mathbf{g}_N), \ (x,y) \in \partial\Omega_N, \\ \mathbf{n} \cdot (\nabla u_h)^+, \ (x,y) \in \partial\Omega_D \cap \partial\Omega^-, \\ \mathbf{n} \cdot (\nabla u_h)^-, \ (x,y) \in \partial\Omega_D \cap \partial\Omega^+, \end{cases}$$

where Π_h^- is a special projection, which will be defined later.

- Numerical fluxes associated with the purely Dirichlet boundary conditions (1f): If Γ is an interior edge then we take

(3da)
$$\widehat{u}_h = u_h^- \quad \text{and} \quad \mathbf{n} \cdot \widehat{\nabla u_h} = \mathbf{n} \cdot (\nabla u_h)^+,$$

and if the edge Γ lies on $\partial\Omega$, we take

(3db)
$$\hat{u}_h = \Pi_h^- g_D, \ (x,y) \in \partial\Omega, \quad \mathbf{n} \cdot \widehat{\nabla u_h} = \begin{cases} \mathbf{n} \cdot (\nabla u_h)^+, \ (x,y) \in \partial\Omega \cap \partial\Omega^-, \\ \mathbf{n} \cdot (\nabla u_h)^-, \ (x,y) \in \partial\Omega \cap \partial\Omega^+, \end{cases}$$

where Π_h^- is a special projection which will be defined later.

After we select the numerical fluxes $\widehat{u}_h$ and $\widehat{\nabla u_h}$, the discrete scheme (3) is equivalent to an algebraic system of nonlinear equations for the unknown coefficients appearing in u_h. The resulting system can be solved by direct or iterative methods for nonlinear system of equations.

Throughout the paper, we use the following notation:

- Γ_h: the set of all element interfaces of $\mathcal{T}_h$,
- Γ_B: the set of all boundary edges of the partition $\mathcal{T}_h$ on $\partial\Omega$,
- $\Gamma_I = \Gamma_h \backslash \Gamma_B$: the set of all interior interfaces of $\mathcal{T}_h$,
- Γ_B^-: the set of edges on the inflow boundary $\partial\Omega^-$,
- Γ_B^+: the set of edges on the outflow boundary $\partial\Omega^+$,
- $\Gamma_D^\pm = \partial\Omega_D \cap \partial\Omega^\pm$ and$\Gamma_N^\pm = \partial\Omega_N \cap \partial\Omega^\pm$,
- Γ_D and Γ_N, respectively, denote the sets of all edges in $\partial\Omega_D$ and $\partial\Omega_N$.

2.2. Norms. In this subsection, we define several norms that will be used throughout the paper. Denote $\|v\|_{0,K} = \left(\iint_K v^2(x,y)dxdy\right)^{1/2}$ to be the standard L^2-norm of v on the rectangle $K \in \mathcal{T}_h$. For any natural number ℓ and for $K \in \mathcal{T}_h$, we consider the H^ℓ-norm of the Sobolev space $H^\ell(K)$, defined by

$$\|v\|_{\ell,K} = \left(\sum_{0 \le \alpha+\beta \le \ell} \left\| \frac{\partial^{\alpha+\beta} v}{\partial x^\alpha \partial y^\beta} \right\|_{0,K}^2 \right)^{1/2}.$$

Let Γ_K be the edges of the element K, and we define

$$\|v\|_{\Gamma_K} = \left(\int_{\partial K} v^2(x(s), y(s))ds \right)^{1/2}.$$

We also define the broken Sobolev norm by

$$\|v\|_\ell = \left(\sum_{K \in \mathcal{T}_h} \|v\|_{\ell,K}^2 \right)^{1/2}.$$

Moreover, we define the H^ℓ-norm for a real-valued function v on the whole computational domain Ω as

$$\|v\|_{\Gamma_h} = \left(\sum_{K \in \mathcal{T}_h} \|v\|_{\Gamma_K}^2 \right)^{1/2}.$$

For a vector-valued function $\mathbf{w} = [w_1, w_2]^t \in \mathbf{H}^\ell(K) = \left(H^\ell(K)\right)^2$, we define the $\mathbf{H}^\ell$-norm on the rectangle K as

$$\|\mathbf{w}\|_{\ell,K} = \left(\|w_1\|_{\ell,K}^2 + \|w_2\|_{\ell,K}^2\right)^{1/2}.$$

The $\mathbf{H}^\ell$-norm of $\mathbf{w}$ on the whole domain Ω is

$$\|\mathbf{w}\|_\ell = \left(\sum_{K\in\mathcal{T}_h} \|\mathbf{w}\|_{\ell,K}^2\right)^{1/2}.$$

Finally, we define the semi-norm on the element K and the semi-norm on the computational domain Ω as

$$|v|_{\ell,K} = \left(\sum_{\alpha+\beta=\ell} \left\|\frac{\partial^{\alpha+\beta} v}{\partial x^\alpha \partial y^\beta}\right\|_{0,K}^2\right)^{1/2}, \quad |v|_\ell = \left(\sum_{K\in\mathcal{T}_h} |v|_{\ell,K}^2\right)^{1/2}.$$

2.3. Inverse/trace estimates. In this subsection, we recall the basic properties of the finite element space V_h^p. We summarize the classical inverse and trace properties in the following lemma [20].

Lemma 2.1. *Let $K = I_i \times J_j$ be an element in $\mathcal{T}_h$ and denote its boundary by Γ. Suppose that $v \in V_h^p$. Then there exists a constant C independent of the mesh size h and v such that*

$$\|\nabla v\|_{0,K} \le Ch^{-1} \|v\|_{0,K}, \tag{5a}$$

$$h\|v\|_{\infty,K} + h^{1/2}\|v\|_{0,\Gamma} \le C\|v\|_{0,K}, \tag{5b}$$

where $\|v\|_{0,\Gamma} = \left(\int_\Gamma v^2(x(s), y(s))\, ds\right)^{1/2}$ and $\|v\|_{\infty,K} = \max_{(x,y)\in K} |v(x,y)|$.

2.4. Projections. In this paper, we consider special projections in one and two dimensions. We use $\mathbb{P}^p(I_i)$ to denote the space of polynomials of degree not exceeding p on $I_i = [x_{i-1}, x_i]$. For $p \ge 1$, we define the special projection P_x^- into V_h^p such that, for any u and for all $i = 0, 1, \ldots, n$, the projection $P_x^- u$ satisfies:

$$\int_{I_i} (u - P_x^- u)v\, dx = 0, \ \forall\, v \in \mathbb{P}^{p-2}(I_i),$$
$$(u - P_x^- u)_x(x_{i-1}^+) = 0, \quad (u - P_x^- u)(x_i^-) = 0. \tag{6}$$

Similarly, for $p \ge 1$, we define another projection P_x^+ into V_h^p such that, for any u, the projection $P_x^+ u$ satisfies: for all $i = 0, 1, \ldots, n$

$$\int_{I_i} (u - P_x^+ u)v\, dx = 0, \ \forall\, v \in \mathbb{P}^{p-2}(I_i),$$
$$(u - P_x^+ u)_x(x_i^-) = 0, \quad (u - P_x^+ u)(x_{i-1}^+) = 0. \tag{7}$$

For the above one-dimensional projection, the following *a priori* error estimates hold

$$\left\|u - P_x^\pm u\right\| + h\left\|u - P_x^\pm u\right\|_\infty + h^{1/2}\left\|u - P_x^\pm u\right\|_{\Gamma_h} \le Ch^{p+1}\|u\|_{p+1}. \tag{8}$$

where Γ_h denotes the set of boundary points of all elements I_i, and C is a positive constant dependent on p but not on h.

Since Cartesian meshes are used in this paper, we apply the tensor product of the projections in the one-dimensional case. On the rectangle element $K = I_i \times J_j$, we define special projections Π_h^- for a real-valued function $u = u(x, y)$ into V_h^p as tensor product of the projections in one dimension

$$\Pi_h^{\pm} u = P_x^{\pm} \otimes P_y^{\pm} u, \tag{9}$$

with the subscripts x and y indicating the use of the one-dimensional projections $P_x^{\pm}$ with respect to the corresponding variable. To be more specific, the projection $\Pi_h^- u \in V_h^p$ satisfies the following $(p+1)^2$ conditions

$$\iint_K (u - \Pi_h^- u) v \, dxdy = 0, \quad \forall\, v \in \mathcal{Q}^{p-2}(K), \tag{10a}$$

$$\int_{J_j} (u - \Pi_h^- u)(x_i^-, y) v(y) \, dy = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(J_j), \tag{10b}$$

$$\int_{I_i} (u - \Pi_h^- u)(x, y_j^-) v(x) \, dx = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(I_i), \tag{10c}$$

$$\int_{J_j} (u - \Pi_h^- u)_x(x_{i-1}^+, y) v(y) \, dy = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(J_j), \tag{10d}$$

$$\int_{I_i} (u - \Pi_h^- u)_y(x, y_{j-1}^+) v(x) \, dx = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(I_i), \tag{10e}$$

$$(u - \Pi_h^- u)(x_i^-, y_j^-) = 0, \tag{10f}$$

$$(u - \Pi_h^- u)_x(x_{i-1}^+, y_j^-) = 0, \tag{10g}$$

$$(u - \Pi_h^- u)_y(x_i^-, y_{j-1}^+) = 0, \tag{10h}$$

$$(u - \Pi_h^- u)_{xy}(x_{i-1}^+, y_{j-1}^+) = 0. \tag{10i}$$

Similar to the one-dimensional case, we have the following error estimates.

Lemma 2.2. *The two-dimensional projection $\Pi_h^- u$ satisfying (10) exists and is unique. Furthermore, for $u \in H^{p+1}(\Omega)$, there exists a constant C independent of h such that*

$$\left\| u - \Pi_h^- u \right\| + h \left\| u - \Pi_h^- u \right\|_1 \le C h^{p+1} \left\| u \right\|_{p+1}. \tag{11}$$

Proof. Since (10) is a linear problem in finite dimension, existence is equivalent to uniqueness. Thus, it is enough to show uniqueness. We assume that there are two solutions $\Pi_h^- u$ and $\hat{\Pi}_h^- u$. The difference $Q_h^- u = \Pi_h^- u - \hat{\Pi}_h^- u \in \mathcal{Q}^p(K)$ satisfies

$$\iint_K Q_h^- u \; v \, dxdy = 0, \quad \forall\, v \in \mathcal{Q}^{p-2}(K), \tag{12a}$$

$$\int_{J_j} Q_h^- u(x_i^-, y) v(y) \, dy = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(J_j), \tag{12b}$$

$$\int_{I_i} Q_h^- u(x, y_j^-) v(x) \, dx = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(I_i), \tag{12c}$$

$$\int_{J_j} (Q_h^- u)_x(x_{i-1}^+, y) v(y) \, dy = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(J_j), \tag{12d}$$

$$\int_{I_i} (Q_h^- u)_y(x, y_{j-1}^+) v(x) \, dx = 0, \quad \forall\, v \in \mathbb{P}^{p-2}(I_i), \tag{12e}$$

(12f)
$$Q_h^- u(x_i^-, y_j^-) = (Q_h^- u)_x(x_{i-1}^+, y_j^-) = (Q_h^- u)_y(x_i^-, y_{j-1}^+) = (Q_h^- u)_{xy}(x_{i-1}^+, y_{j-1}^+) = 0.$$

Let $Q_h^- u = \sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}L_k(x)L_l(y)$. Using (12a) with $v = L_r(x)L_s(y),\ r,s = 0,1,\dots,p-2$, we get

$$\begin{aligned}0 &= \sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}\iint_K L_k(x)L_l(y)L_r(x)L_s(y)\,dxdy\\ &= \sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}\int_{I_i} L_k(x)L_r(x)\,dx\int_{J_j} L_l(y)L_s(y)\,dy = c_{r,s}\frac{h_i}{2r+1}\frac{k_j}{2s+1}.\end{aligned}$$

Therefore

$$c_{r,s} = 0, \quad r,s = 0,1,\dots,p-2. \tag{13}$$

Next, we use (12b) with $v = L_s(y),\ s = 0,1,\dots,p-2$ and we apply (13) to get

$$\begin{aligned}0 = \int_{J_j}\sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}L_k(x_i)L_l(y)L_s(y)\,dy &= \sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}\int_{J_j} L_l(y)L_s(y)\,dy\\ &= \frac{k_j}{2s+1}\sum_{k=0}^{p} c_{k,s} = \frac{k_j}{2s+1}\left(c_{p-1,s}+c_{p,s}\right).\end{aligned}$$

Thus, we get

$$c_{p-1,s} + c_{p,s} = 0, \quad s = 0,1,\dots,p-2. \tag{14}$$

Similarly, we use (12d) with $v = L_s(y),\ s = 0,1,\dots,p-2$ and $L_k'(x_i) = (-1)^{k+1}\frac{k(k+1)}{h_i}$ to get

$$\begin{aligned}0 &= \int_{J_j}\sum_{k=0}^{p}\sum_{l=0}^{p} c_{k,l}L_k'(x_{i-1})L_l(y)L_s(y)\,dy\\ &= \sum_{k=0}^{p}\sum_{l=0}^{p}(-1)^{k+1}\frac{k(k+1)}{h_i}c_{k,l}\int_{J_j} L_l(y)L_s(y)\,dy\\ &= \frac{k_j}{2s+1}\sum_{k=0}^{p}(-1)^{k+1}\frac{k(k+1)}{h_i}c_{k,s}\\ &= \frac{k_j}{h_i(2s+1)}\Big((-1)^p(p-1)pc_{p-1,s} + (-1)^{p+1}p(p+1)c_{p,s} + \sum_{k=0}^{p-2}(-1)^{k+1}k(k+1)c_{k,s}\Big)\end{aligned}$$

Therefore

$$(-1)^p(p-1)pc_{p-1,s} + (-1)^{p+1}p(p+1)c_{p-1,s} = 0, \quad s = 0,1,\dots,p-2. \tag{15}$$

Solving the system (14) and (15), we get

$$c_{p-1,s} = c_{p,s} = 0, \quad s = 0,1,\dots,p-2.$$

Similarly, we use (12c) and (12e) to get

$$c_{r,p-1} = c_{r,p} = 0, \quad r = 0,1,\dots,p-2.$$

Thus, we we have

$$\begin{aligned}Q_h^- u =& c_{p-1,p-1}L_{p-1}(x)L_{p-1}(y) + c_{p-1,p}L_{p-1}(x)L_l(p)\\ &+ c_{p,p-1}L_p(x)L_l(p-1) + c_{p,p}L_p(x)L_l(p).\end{aligned}$$

Finally, we use the four conditions in (12f) to get

(16)
$$c_{p-1,p-1} + c_{p-1,p} + c_{p,p-1} + c_{p,p} = 0,$$
(17)
$$(-1)^p \left((p-1)pc_{p-1,p-1} + (p-1)pc_{p-1,p} - p(p+1)c_{p,p-1} - p(p+1)c_{p,p}\right) = 0,$$
(18)
$$(-1)^p \left((p-1)pc_{p-1,p-1} - p(p+1)c_{p-1,p} + (p-1)pc_{p,p-1} - 1p(p+1)c_{p,p}\right) = 0,$$
(19)
$$(p-1)^2p^2c_{p-1,p-1} - (p^2-1)p^2c_{p-1,p} - (p^2-1)p^2c_{p,p-1} + p^2(p+1)^2c_{p,p} = 0.$$

Solving the above 4×4 linear system, we get $c_{p-1,p-1} = c_{p-1,p} = c_{p,p-1} = c_{p,p} = 0$, since the determinant of the coefficient matrix is $-16p^8$. Consequently, we get $Q_h^- u = 0$.
Finally, the standard approximation theory implies the *a priori* error estimate (11). $\square$

3. Error analysis

In this section, we prove the *a priori* error estimates of the UWDG scheme for the model BVP (1). Throughout this section, we use c, $C > 0$ (with or without subscripts) to denote generic constants, that may change from line to line, but are not depending on the crucial quantities such as the mesh size h.
To derive the error estimates, let us first denote the error by

$$e_u = u - u_h.$$

We denote the error between the UWDG solution and the projection of the exact solution by

$$\xi_u = u_h - \Pi_h^- u \in V_h^p.$$

We denote the projection error by

$$\eta_u = u - \Pi_h^- u.$$

Then, the actual error e_u can be written as

$$e_u = \eta_u - \xi_u. \tag{20}$$

In our analysis, we shall assume that the function f in (1a) is sufficiently differentiable function. Let $D = \{(x,y,u)|\ (x,y) \in \bar{\Omega},\ u \in \mathbb{R}\} \subset \mathbb{R}^3$. More precisely, we assume that the f satisfies the following conditions:

Assumption A1. *The function $f(x,y,u)$ and its partial derivative f_u are continuous on D and f_u is bounded on D i.e.,*

$$|f_u(x,y,u)| \le M, \quad \forall\ (x,y,u) \in D. \tag{21}$$

Assumption A2. *The function f_u is Lipschitz function. That is, for all $(x_1,y_1,u_1) \in D$ and $(x_2,y_2,u_2) \in D$, there exists a positive constant L such that*

$$|f_u(x_1,y_1,u_1) - f_u(x_2,y_2,u_2)| \le L\left(|x_1-x_2| + |y_1-y_2| + |u_1-u_2|\right). \tag{22}$$

3.1. Preliminary results. In our error analysis, we need the following bilinear form

$$B_K(u,v) = \iint_K u\Delta v \; dxdy - \int_{\partial K} u^- \; \mathbf{n}\cdot\nabla v \; ds + \int_{\partial K} \mathbf{n}\cdot(\nabla u)^+ \; vds. \tag{23}$$

In the next lemma, we state and prove essential results which will be needed to prove our main optimal error estimate.

Lemma 3.1. *Let B_K be defined by (23). Then we have*

$$B_K(\eta_u, v) = 0, \quad \forall\; u \in \mathcal{P}^{p+2}(K), \quad \forall\; v \in \mathcal{Q}^p(K), \tag{24}$$

where $\mathcal{P}^p(K)$ is the space of polynomials of degree at most p on $K \in \mathcal{T}_h$. In addition, for $u \in H^{p+3}(K)$, we have the following estimate

$$\left|\sum_{K\in\mathcal{T}_h} B_K(\eta_u, v)\right| \le Ch^{p+1}\,|u|_{p+3}\,\|v\|, \quad \forall\; v \in V_h^p. \tag{25}$$

Proof. Consider the rectangle $K = I_i \times J_j \in \mathcal{T}_h$, where $I_i = [x_{i-1}, x_i],\; i = 1,2,\ldots,n$ and $J_j = [y_{j-1}, y_j],\; j = 1,2,\ldots,m$. We have

$$\begin{aligned} B_K(\eta_u, v) = &\iint_K (u-\Pi^-u)(v_{xx}+v_{yy})dxdy \\ &+ \int_{J_j} (u-\Pi^-u)(x_{i-1}^-, y)v_x(x_{i-1}^+, y)dy - \int_{J_j} (u-\Pi^-u)(x_i^-, y)v_x(x_i^-, y)dy \\ &+ \int_{I_i} (u-\Pi^-u)(x, y_{j-1}^-)v_y(x, y_{j-1}^+)dx - \int_{I_i} (u-\Pi^-u)(x, y_j^-)v_y(x, y_j^-)dx \\ &- \int_{J_j} (u-\Pi^-u)_x(x_{i-1}^+, y)v(x_{i-1}^+, y)dy + \int_{J_j} (u-\Pi^-u)_x(x_i^+, y)v(x_i^-, y)dy \\ &- \int_{I_i} (u-\Pi^-u)_y(x, y_{j-1}^+)v(x, y_{j-1}^+)dx + \int_{I_i} (u-\Pi^-u)_y(x, y_j^+)v(x, y_j^-)dx. \end{aligned}$$

We remark that $u - \Pi^-u = 0$ for all $u \in \mathcal{Q}^p(K)$. Thus, (24) holds true for all $u \in \mathcal{Q}^p(K)$. Thus, we only need to consider the cases $u(x,y) = x^{p+1}$, $x^{p+1}y$, x^{p+2}, y^{p+1}, xy^{p+1}, y^{p+2}.
Let us start with $u(x,y) = x^{p+1}$ then $(u-\Pi^-u)_y = 0$ since u in independent of y. Also, by definition of the projection Π^-, we have $\iint_K (u-\Pi^-u)v_{xx}dxdy = 0$, $(u-\Pi^-u)(x_i^-, y) = (u-\Pi^-u)_x(x_{i-1}^+, y) = 0$. Thus, $B_K(\eta_u, v)$ reduces to

$$\begin{aligned} B_K(\eta_u, v) = &\iint_K (u-\Pi^-u)v_{yy}dxdy \\ &+ \int_{I_i} (u-\Pi^-u)(x, y_{j-1}^-)v_y(x, y_{j-1}^+)dx - \int_{I_i} (u-\Pi^-u)(x, y_j^-)v_y(x, y_j^-)dx. \end{aligned}$$

Since $u - \Pi^-u$ is independent of y, we write

$$\begin{aligned} &\iint_K (u-\Pi^-u)v_{yy}dxdy \\ &= \int_{I_i} \left[(u-\Pi^-u)(x, y_j^-)v_y(x, y_j^-) - (u-\Pi^-u)(x, y_{j-1}^+)v_y(x, y_{j-1}^+)\right] dx. \end{aligned}$$

Thus, $B_K(\eta_u, v) = 0$ since $(u-\Pi^-u)(x, y_{j-1}^+) = (u-\Pi^-u)(x, y_{j-1}^-)$. The proof of the case $u(x,y) = x^{p+2}$ is similar to that of the case $u(x,y) = x^{p+1}$. Details are omitted to save space.

Next, we consider the case $u(x,y) = x^{p+1}y$. In this case $u - \Pi^- u = y(x^{p+1} - P_x^-(x^{p+1}))$. Therefore

$$\iint_K (u-\Pi^- u)v_{xx}dxdy = \int_{J_j} y\left(\int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))v_{xx}dx\right)dy = 0,$$

$$\int_{J_j}(u-\Pi^- u)(x_{i-1}^-,y)v_x(x_{i-1}^+,y)dy = \int_{J_j}(u-\Pi^- u)(x_i^-,y)v_x(x_i^-,y)dy = 0,$$

$$\int_{J_j}(u-\Pi^- u)_x(x_{i-1}^+,y)v(x_{i-1}^+,y)dy = \int_{J_j}(u-\Pi^- u)_x(x_i^+,y)v(x_i^-,y)dy = 0.$$

Thus, $B_K(\eta_u, v)$ simplifies to

$$\begin{aligned} B_K(\eta_u,v) = &\iint_K (u-\Pi^- u)v_{yy}dxdy \\ &+\int_{I_i}(u-\Pi^- u)(x,y_{j-1}^-)v_y(x,y_{j-1}^+)dx - \int_{I_i}(u-\Pi^- u)(x,y_j^-)v_y(x,y_j^-)dx \\ &-\int_{I_i}(u-\Pi^- u)_y(x,y_{j-1}^+)v(x,y_{j-1}^+)dx + \int_{I_i}(u-\Pi^- u)_y(x,y_j^+)v(x,y_j^-)dx. \end{aligned}$$

Using $u - \Pi^- u = y(x^{p+1} - P_x^-(x^{p+1}))$, we obtain

$$\begin{aligned} B_K(\eta_u,v) = &\iint_K (u-\Pi^- u)v_{yy}dxdy + \int_{I_i} y_{j-1}(x^{p+1}-P_x^-(x^{p+1}))v_y(x,y_{j-1}^+)dx \\ &-\int_{I_i} y_j(x^{p+1}-P_x^-(x^{p+1}))v_y(x,y_j^-)dx \\ &-\int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))v(x,y_{j-1}^+)dx + \int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))v(x,y_j^-)dx. \end{aligned} \tag{26}$$

Next, we integrate by parts to write

$$\begin{aligned} &\iint_K (u-\Pi^- u)v_{yy}dxdy = \int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))\left(\int_{J_j} yv_{yy}dy\right)dx \\ &= \int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))\left(y_jv_y(x,y_j^-) - y_{j-1}v_y(x,y_{j-1}^+) - \int_{J_j} v_ydy\right)dx \\ &= \int_{I_i} y_j(x^{p+1}-P_x^-(x^{p+1}))v_y(x,y_j^-)dx - \int_{I_i} y_{j-1}(x^{p+1}-P_x^-(x^{p+1}))v_y(x,y_{j-1}^+)dx \\ &-\int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))v(x,y_j^-)dx + \int_{I_i}(x^{p+1}-P_x^-(x^{p+1}))v(x,y_{j-1}^+)dx. \end{aligned} \tag{27}$$

Combining (26) and (27), we get $B_K(\eta_u, v) = 0$.

We would like to remark that the proof of the cases $u(x,y) = y^{p+1}$, $u(x,y) = xy^{p+1}$, and $u(x,y) = y^{p+2}$ are very similar to the cases $u(x,y) = x^{p+1}$, $u(x,y) = x^{p+1}y$, and $u(x,y) = x^{p+2}$. We omit the details to save space. This completes the proof of (24).

Finally, we will prove (25). We first assume that K is the reference square element $\hat{K} = [-1,1]^2$. For $\hat{u} \in H^{p+3}(\hat{K})$ and for fixed $\hat{v} \in \mathcal{Q}^p(\hat{K})$, the linear functional $\hat{u} \to B_{\hat{K}}(\eta_{\hat{u}}, \hat{v})$ is continuous on $H^{p+3}(\hat{K})$ with norm bounded by $C\|\hat{v}\|_{0,\hat{K}}$ and, by (24), it vanishes for $\hat{u} \in \mathcal{P}^{p+2}(\hat{K})$. By the Bramble-Hilbert lemma (see *e.g.*, [19, Lemma 6]), we obtain for $\hat{u} \in H^{p+3}(\hat{K})$ and $\hat{v} \in \mathcal{Q}^p(\hat{K})$

$$\left|B_{\hat{K}}(\eta_{\hat{u}}, \hat{v})\right| \le C_1 |\hat{u}|_{p+3,\hat{K}} \|\hat{v}\|_{0,\hat{K}}.$$

Going back to the physical element K by using the transformation

$$x(\xi) = \frac{x_{i-1}+x_i}{2} + \frac{h_i}{2}\xi, \quad y(\eta) = \frac{y_{j-1}+y_j}{2} + \frac{k_j}{2}\eta, \tag{28}$$

we get for $u \in H^{p+3}(K)$ and $v \in \mathcal{Q}^p(K)$

$$|B(\eta_u, v)| \le C_1 \left|B_{\hat{K}}(\eta_{\hat{u}}, \hat{v})\right| \le C_2 \left|\hat{u}\right|_{p+3,\hat{K}} \left\|\hat{v}\right\|_{0,\hat{K}} \le C h^{p+1} \left|u\right|_{p+3,K} \left\|v\right\|_{0,K}. \tag{29}$$

Summing over all the elements, we establish the superconvergence result (25). □

Next, we derive the error equations that will be used for the error analysis.

3.2. Error equations. For simplicity we only consider the case of periodic boundary conditions or mixed Dirichlet-Neumann boundary conditions:

$$u = g_D, \quad (x,y) \in \partial\Omega_D = \partial\Omega^-, \quad \mathbf{n}\cdot\nabla u = \mathbf{n}\cdot\mathbf{g}_N, \quad (x,y) \in \partial\Omega_N = \partial\Omega^+. \tag{30}$$

We remark that this assumption is not essential. We note that if other boundary conditions are chosen, the numerical fluxes can be easily designed; see *e.g.*, [1, 8, 9, 10, 12]. When the mixed Dirichlet-Neumann boundary conditions (30) are used, the numerical fluxes (3c) simplify to

$$\widehat{u}_h = u_h^- \quad \text{and} \quad \mathbf{n}\cdot\widehat{\nabla u_h} = \mathbf{n}\cdot(\nabla u_h)^+, \tag{31a}$$

and if the edge Γ lies on $\partial\Omega$, we take

$$\hat{u}_h = \begin{cases} u_h^-, \ (x,y) \in \Gamma_I \cup \Gamma_B^+, \\ \Pi_h^- g_D, \ (x,y) \in \Gamma_B^-, \end{cases} \quad \mathbf{n}\cdot\widehat{\nabla u_h} = \begin{cases} \Pi_h^-(\mathbf{n}\cdot\mathbf{g}_N), \ (x,y) \in \Gamma_B^+, \\ \mathbf{n}\cdot(\nabla u_h)^+, \ (x,y) \in \Gamma_I \cup \Gamma_B^-, \end{cases} \tag{31b}$$

Subtracting the discrete formulation (3) from the continuous formulation (2) with $v \in V_h^p$, we obtain the following error equation on the element K: $\forall\ v \ \in V_h^p$,

$$\begin{aligned} -\iint_K e_u \Delta v \ dxdy + \int_{\partial K} e_u^- \ \mathbf{n}\cdot\nabla v \ ds - \int_{\partial K} \mathbf{n}\cdot(\nabla e_u)^+ \ v ds \\ + \iint_K (f(x,y,u) - f(x,y,u_h))\, v \ dxdy = 0. \end{aligned} \tag{32}$$

Next, we apply the integral mean-value theorem to write

$$f(x,y,u) - f(x,y,u_h) = R e_u, \quad \text{where } R = R(x,y) = \int_0^1 f_u\left(x,y,u - t e_u\right) dt. \tag{33}$$

By the assumption A1, we obtain

$$|R| \le M, \quad \forall\ (x,y) \in \bar{\Omega}. \tag{34}$$

Substituting (33) into (32), we get

$$\iint_K e_u(-\Delta v + R\ v) \ dxdy + \int_{\partial K} e_u^- \ \mathbf{n}\cdot\nabla v \ ds - \int_{\partial K} \mathbf{n}\cdot(\nabla e_u)^+ \ v ds = 0. \tag{35}$$

Splitting the errors as in (20), we obtain

$$\begin{aligned} \iint_K \xi_u(\Delta v - R v) \ dxdy - \int_{\partial K} \xi_u^- \ \mathbf{n}\cdot\nabla v \ ds + \int_{\partial K} \mathbf{n}\cdot(\nabla\xi_u)^+ \ v ds = \iint_K \eta_u \Delta v \ dxdy \\ - \int_{\partial K} \eta_u^- \ \mathbf{n}\cdot\nabla v \ ds + \int_{\partial K} \mathbf{n}\cdot(\nabla\eta_u)^+ \ v ds - \iint_K R\eta_u \ v \ dxdy. \end{aligned} \tag{36}$$

For simplicity, we rewrite (36) as

$$\begin{aligned}&\iint_K \xi_u(\Delta v - Rv)\; dxdy - \int_{\partial K} \xi_u^- \;\mathbf{n}\cdot\nabla v\; ds + \int_{\partial K} \mathbf{n}\cdot(\nabla\xi_u)^+\; vds\\ &= B_K(\eta_u, v) - \iint_K R\eta_u\; v\; dxdy,\end{aligned} \tag{37}$$

where the bilinear operator $B_K(\cdot,\cdot)$ is defined in (23). Summing over all elements $K \in \mathcal{T}_h$ yields

$$\begin{aligned}&\iint_\Omega \xi_u(\Delta v - Rv)\; dxdy - \sum_{\Gamma^-\in\Gamma_B^-}\int_{\Gamma^-} \xi_u^-\;\mathbf{n}\cdot(\nabla v)^+\; ds\\ &- \sum_{\Gamma^+\in\Gamma_B^+}\int_{\Gamma^+} \xi_u^-\;\mathbf{n}\cdot(\nabla v)^-\; ds - \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \xi_u^-\;\mathbf{n}\cdot[\![\nabla v]\!]\; ds\\ &+ \sum_{\Gamma^-\in\Gamma_B^-}\int_{\Gamma^-} \mathbf{n}\cdot(\nabla\xi_u)^+\; v^+\; ds + \sum_{\Gamma^+\in\Gamma_B^+}\int_{\Gamma^+} \mathbf{n}\cdot(\nabla\xi_u)^+\; v^- ds\\ &+ \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \mathbf{n}\cdot(\nabla\xi_u)^+\; [\![v]\!] ds = \sum_{K\in\mathcal{T}_h} B_K(\eta_u, v) - \iint_\Omega R\eta_u\; v\; dxdy.\end{aligned} \tag{38}$$

If the periodic boundary conditions are used then (38) reduces to

$$\begin{aligned}&\iint_\Omega \xi_u(\Delta v - Rv)\; dxdy - \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \xi_u^-\;\mathbf{n}\cdot[\![\nabla v]\!]\; ds + \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \mathbf{n}\cdot(\nabla\xi_u)^+\; [\![v]\!] ds\\ &= \sum_{K\in\mathcal{T}_h} B_K(\eta_u, v) - \iint_\Omega R\eta_u\; v\; dxdy.\end{aligned} \tag{39}$$

If the mixed Dirichlet-Neumann boundary conditions (30) are used then (38) reduces to

$$\begin{aligned}&\iint_\Omega \xi_u(\Delta v - Rv)\; dxdy - \sum_{\Gamma^+\in\Gamma_B^+}\int_{\Gamma^+} \xi_u^-\;\mathbf{n}\cdot(\nabla v)^-\; ds - \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \xi_u^-\;\mathbf{n}\cdot[\![\nabla v]\!]\; ds\\ &+ \sum_{\Gamma^-\in\Gamma_B^-}\int_{\Gamma^-} \mathbf{n}\cdot(\nabla\xi_u)^+\; v^+\; ds + \sum_{\Gamma\in\Gamma_I}\int_{\Gamma} \mathbf{n}\cdot(\nabla\xi_u)^+\; [\![v]\!] ds = \sum_{K\in\mathcal{T}_h} B_K(\eta_u, v)\\ &- \iint_\Omega R\eta_u\; v\; dxdy,\end{aligned} \tag{40}$$

since $\xi_u^- = 0$ on Γ_B^- and $\mathbf{n}\cdot(\nabla\xi_u)^+ = 0$ on Γ_B^+.

3.3. Regularity estimates. Here, we present some regularity estimates, which will be needed to prove the main error estimates of our UWDG scheme. When the periodic boundary conditions are used, we consider the following linear elliptic problem: find $V \in H^2(\Omega)$ such that

$$-\Delta V + RV = \xi_u, \quad \text{for } (x,y)\in\Omega, \tag{41a}$$

$$u(a,y) = u(b,y),\; u(x,c) = u(x,d),\; u_x(a,y) = u_x(b,y),\; u_y(x,c) = u_y(x,d),\; \mathbf{x}\in\partial\Omega. \tag{41b}$$

When the mixed Dirichlet-Neumann boundary conditions are used, we consider the following linear elliptic problem: find $V \in H^2(\Omega)$ such that

$$-\Delta V + RV = \xi_u, \quad \text{for } (x,y) \in \Omega, \tag{42a}$$

$$V = 0, \quad \text{on } \Gamma_B^-, \quad \mathbf{n} \cdot \nabla V = 0, \quad \text{on } \Gamma_B^+. \tag{42b}$$

The following lemma will be used to estimate $\|\xi_u\|$ and $\|e_u\|$.

Lemma 3.2. *Suppose that $V \in H^2(\Omega)$ satisfies either (41) or (42). Then, we have the following regularity estimate*

$$\|V\|_2 \le C \|\xi_u\|. \tag{43}$$

Proof. This elliptic regularity estimate is classical. We refer the reader to *e.g.*, [16, 28, 36]. □

3.4. *A priori* error estimates. In the next theorem, we state the *a priori* error estimates for e_u in the L^2-norm.

Theorem 3.1. *Let $u \in H^{p+3}(\Omega)$ be the solution of (1). Suppose that $p \ge 1$ and u_h be the UWDG solution defined in (3). Then, for sufficiently small h, there exists a constant C such that*

$$\|\xi_u\| \le Ch^{p+1}. \tag{44}$$

Consequently, we have the following error estimates

$$\|e_u\| \le Ch^{p+1}, \tag{45}$$

$$\|e_u\|_\infty \le Ch^{p}. \tag{46}$$

Proof. We first estimate $\|\xi_u\|$ in the L^2-norm by using a duality argument. When the periodic boundary conditions are used, we assume that V satisfies (41) and when the mixed Dirichlet-Neumann boundary conditions are used, we assume that V satisfies (42). Then (39) and (40) with $v = V$ both reduce to

$$\|\xi_u\|^2 = -\sum_{K\in\mathcal{T}_h} B_K(\eta_u, V) + \iint_\Omega R\eta_u \, V \, dxdy. \tag{47}$$

Adding and subtracting $\Pi_h^- V$, we get

$$\|\xi_u\|^2 = T_1 + T_2 + T_3, \tag{48}$$

where

$$T_1 = -\sum_{K\in\mathcal{T}_h} B_K(\eta_u, \eta_V), \quad T_2 = -\sum_{K\in\mathcal{T}_h} B_K(\eta_u, \Pi_h^- V), \quad T_3 = \iint_\Omega R\eta_u \, V \, dxdy.$$

Next, we will estimate T_1, T_2, and T_3 one by one.

Estimate of T_1: We first assume that K is the reference square element $\hat{K} = [-1,1]^2$. Define the bilinear form $\hat{B}_{\hat{K}}(\hat{u}, \hat{V}) = B_{\hat{K}}(\eta_{\hat{u}}, \eta_{\hat{V}})$. Clearly, for $p \ge 1$, the bilinear form $\hat{B}_{\hat{K}} : H^{p+1}(\hat{K}) \times H^2(\hat{K}) \to \mathbb{R}$ satisfies

(i) $\hat{B}_{\hat{K}}(\hat{u}, \hat{V}) = 0$ for all $\hat{u} \in H^{p+1}(\hat{K})$ and $\hat{V} \in \mathcal{P}^p(\hat{K})$,
(ii) $\hat{B}_{\hat{K}}(\hat{u}, \hat{V}) = 0$ for all $\hat{u} \in \mathcal{P}^p(\hat{K})$ and $\hat{V} \in H^2(\hat{K})$.

By the Bramble-Hilbert lemma, we obtain for $\hat{u} \in H^{p+1}(\hat{K})$ and $\hat{V} \in H^2(\hat{K})$

$$\left|\hat{B}_{\hat{K}}(\hat{u}, \hat{V})\right| \le C_1 \left|\hat{u}\right|_{p+1,\hat{K}} \left|\hat{V}\right|_{2,\hat{K}}.$$

Going back to the physical element K by using the transformation (28), we get for $u \in H^{p+1}(K)$ and $V \in H^2(K)$

$$\begin{aligned} |B(\eta_u, \eta_V)| \leq & C_1 \left| B_{\hat{K}}(\eta_{\hat{u}}, \eta_{\hat{V}}) \right| = C_1 \left| \hat{B}_{\hat{K}}(\hat{u}, \hat{V}) \right| \\ \leq & C_2 \left| \hat{u} \right|_{p+1,\hat{K}} \left| \hat{V} \right|_{2,\hat{K}} \leq C_3 h^{p+1} \left| u \right|_{p+1,K} \left| V \right|_{2,K}. \end{aligned}$$

Summing over all the elements, we obtain

$$\left| \sum_{K \in \mathcal{T}_h} B_K(\eta_u, \eta_V) \right| \leq C h^{p+1} \left| V \right|_2.$$

Using the regularity estimate (43) yields

$$|T_1| \leq C_1 h^{p+1} \|\xi_u\|. \tag{49}$$

<u>**Estimate of T_2:**</u> Using (25) with $v = \Pi_h^- V$, we obtain

$$T_2 \leq C_1 h^{p+1} \left\| \Pi_h^- V \right\|.$$

Next, we show that $\left\| \Pi_h^- V \right\| \leq C_2 \|V\|$ by writing

$$\begin{aligned} \left\| \Pi_h^- V \right\| = \left\| \Pi_h^- V - V + V \right\| & \leq \left\| \Pi_h^- V - V \right\| + \|V\| \\ & \leq C_2 h^2 \|V\|_2 + \|V\| \\ & \leq C_3 \|V\|_2. \end{aligned}$$

Thus, we have

$$T_2 \leq C_1 C_3 h^{p+1} \|V\|_2.$$

Using the regularity estimate (43) yields

$$T_2 \leq C_2 h^{p+1} \|\xi_u\|. \tag{50}$$

<u>**Estimate of T_3:**</u> Using (34) and applying the Cauchy-Schwarz inequality, we get

$$T_3 \leq M \|\eta_u\| \|V\|.$$

Using the regularity estimate (43) and the standard interpolation error estimate (11) yields

$$T_3 \leq C_3 h^{p+1} \|\xi_u\|. \tag{51}$$

Now, combining (48) with (49), (50) and (51), we get

$$\|\xi_u\|^2 \leq C h^{p+1} \|\xi_u\|.$$

Dividing by $\|\xi_u\|$, we get

$$\|\xi_u\| \leq C h^{p+1},$$

which completes the proof of (44).

Splitting the error as in (20), then applying the standard interpolation error estimate (11) and the estimate (44), we get

$$\|e_u\| = \|\eta_u - \xi_u\| \leq \|\eta_u\| + \|\xi_u\| \leq C h^{p+1},$$

which completes the proof of (45). We note that the estimate (46) follows from the inverse property, the triangle inequality, the interpolation property (11), and the estimate (44)

$$\|e_u\|_\infty = \|\eta_u - \xi_u\|_\infty \leq \|\eta_u\|_\infty + \|\xi_u\|_\infty \leq C_1 h^p + C_2 h^{-1} \|\xi_u\| \leq C h^p,$$

which completes the proof of the theorem. □

4. Computational examples

In this section, we present several numerical results to illustrate the theoretical results outlined in this paper. We apply the UWDG scheme (3), described above, to solve numerically the elliptic problem (1). The results are given in the form of tables, which reveal the convergence of the proposed UWDG method. In our numerical experiments, we demonstrate that our theoretical results are valid for periodic, purely Dirichlet, and mixed Dirichlet-Neumann boundary conditions. As estimation of the convergence rates, the experimental order of convergence r is computed using

$$r = -\frac{\ln(\|e\|^{(n_1)} \,/\, \|e\|^{(n_2)})}{\ln(n_1/n_2)},$$

where $\|e\|^{(n)}$ denotes the error $\|e\| = \|u - u_h\|$ using $N = n^2$ elements.

Example 4.1. In this example, we consider the following semilinear Poisson-Boltzmann (PB) equation, which appears in many applications, including semiconductor modeling [22] and charged particles in solutions [30]

$$-u_{xx} - u_{yy} = e^{-u} + g(x,y), \quad (x,y) \in [0,1]^2, \tag{52a}$$

where the function $g(x,y) = 8\pi^2 \sin(2\pi(x+y)) - e^{-\sin(2\pi(x+y))}$ and boundary conditions are extracted from the exact solution

$$u(x,y) = \sin(2\pi(x+y)).$$

We perform several numerical tests on this example to study the effect of boundary conditions and investigate the convergence of the proposed UWDG solutions. We will consider the mixed Dirichlet-Neumann, periodic, and purely Dirichlet boundary conditions. First, we consider (52a) subject to the purely Dirichlet boundary conditions

$$\begin{aligned} &u(0,y) = \sin(2\pi y),\ u(x,0) = \sin(2\pi x), \\ &u(1,y) = \sin(2\pi y), \quad u(x,1) = \sin(2\pi x), \quad (x,y) \in \partial\Omega. \end{aligned} \tag{52b}$$

We solve (52a) subject to the dirichlet boundary conditions (52b) using the UWDG scheme presented in Section 2 on a uniform Cartesian mesh having $N = 25$, 100, 225, 400, and 625 elements obtained by dividing the computational domain $[0,1]^2$ into n^2 squares with $n = 5$, 10, 15, 20, and 25. We compute the UWDG solution u_h in the finite element spaces Q^p with $p = 1,2,3,4$. Both the L^2 errors and orders of accuracy are shown in Table 1. We observe that our UWDG scheme gives the optimal $(p+1)-$th order of the accuracy for this nonlinear problem. The results are in agreement with the theoretical estimates.

We repeat the previous example with all parameters kept unchanged except that we use the mixed Dirichlet-Neumann boundary conditions

$$\begin{aligned} &u(0,y) = \sin(2\pi y),\ u(x,0) = \sin(2\pi x), \\ &u_x(1,y) = \cos(2\pi y), \quad u_y(x,1) = \cos(2\pi x), \quad (x,y) \in \partial\Omega. \end{aligned} \tag{52c}$$

The L^2 errors $\|e_u\|$ are listed in Table 4.1. Clearly, these results suggest optimal convergence rates. Again, our results are in full agreement with the theoretical results.

Finally, we solve the same problem but using the periodic boundary conditions

$$\begin{aligned} &u(0,y) = u(1,y),\ u(x,0) = u(x,1), \\ &u_x(0,y) = u_x(1,y), \quad u_y(x,0) = u_y(x,1), \quad (x,y) \in \partial\Omega. \end{aligned} \tag{52d}$$

TABLE 1. Convergence rates for $\|e_u\|$ for the BVP (52a) subject to (52b) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p = 1$-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	1.0082e-1	NA	9.7831e-3	NA	7.3755e-4	NA	4.6402e-5	NA
100	2.4299e-2	2.0528	1.2083e-3	3.0173	4.6522e-5	3.9868	1.4510e-6	4.9991
225	1.0724e-2	2.0173	3.5784e-4	3.0012	9.2121e-6	3.9940	1.9133e-7	4.9968
400	6.0175e-3	2.0085	1.5104e-4	2.9982	2.9176e-6	3.9966	4.5434e-8	4.9977
625	3.8475e-3	2.0043	7.7366e-5	2.9981	1.1957e-6	3.9976	1.4893e-8	4.9984

TABLE 2. Convergence rates for $\|e_u\|$ for the BVP (52a) subject to (52c) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p = 1$-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	9.1671e-2	NA	9.2366e-3	NA	7.2601e-4	NA	4.5540e-5	NA
100	2.3778e-2	1.9468	1.1960e-3	2.9491	4.6397e-5	3.9679	1.4473e-6	4.9757
225	1.0631e-2	1.9853	3.5672e-4	2.9837	9.2036e-6	3.9896	1.9119e-7	4.9923
400	5.9916e-3	1.9932	1.5084e-4	2.9920	2.9164e-6	3.9948	4.5420e-8	4.9962
625	3.8381e-3	1.9959	7.7315e-5	2.9951	1.1954e-6	3.9968	1.4891e-8	4.9977

TABLE 3. Convergence rates for $\|e_u\|$ for the BVP (52a) subject to (52d) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p = 1$-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	9.2238e-2	NA	9.2305e-3	NA	7.2573e-4	NA	4.5533e-5	NA
100	2.3809e-2	1.9539	1.1959e-3	2.9483	4.6395e-5	3.9674	1.4472e-6	4.9756
225	1.0636e-2	1.9874	3.5671e-4	2.9836	9.2035e-6	3.9895	1.9119e-7	4.9921
400	5.9933e-3	1.9939	1.5084e-4	2.9919	2.9164e-6	3.9948	4.5420e-8	4.9962
625	3.8388e-3	1.9964	7.7314e-5	2.9951	1.1954e-6	3.9968	1.4891e-8	4.9977

We present the L^2 errors in Table 4.1. These results indicate that $\|e_u\| = \mathcal{O}(h^{p+1})$. These results are consistent with the theoretical results.

Example 4.2. In this example, we consider the following semilinear elliptic problem

$$-u_{xx} - u_{yy} + u^3 = g(x,y), \quad (x,y) \in [0,1]^2, \tag{53a}$$

where the function $g(x,y)$ and the boundary conditions are chosen so that the exact solution is

$$u(x,y) = \sin(2\pi x)\sin(2\pi y).$$

First, we consider (53a) subject to the purely Dirichlet boundary conditions

$$u(0,y) = u(x,0) = u(1,y) = u(x,1) = 0, \quad (x,y) \in \partial\Omega. \tag{53b}$$

We solve (53a) subject to (53b) using the proposed UWDG scheme presented in Section 2 on a uniform Cartesian mesh having $N = 25$, 100, 225, 400, and 625 elements obtained by dividing the computational domain $[0,1]^2$ into n^2 squares with $n = 5$, 10, 15, 20, and 25. We compute the UWDG solution u_h in the finite

TABLE 4. Convergence rates for $\|e_u\|$ for the BVP (53a) subject to (53b) on uniform meshes having $N = n^2$ square elements with $n = 5, 10, 15, 20$, and 25 using $p = 1$-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	8.3350e-2	NA	6.0738e-3	NA	5.6930e-4	NA	3.0515e-5	NA
100	1.8514e-2	2.1706	8.2751e-4	2.8758	3.3519e-5	4.0861	1.0126e-6	4.9134
225	7.8031e-3	2.1309	2.5040e-4	2.9481	6.5533e-6	4.0254	1.3474e-7	4.9744
400	4.3111e-3	2.0624	1.0632e-4	2.9776	2.0684e-6	4.0086	3.2070e-8	4.9896
625	2.7394e-3	2.0321	5.4577e-5	2.9884	8.4659e-7	4.0033	1.0521e-8	4.9948

TABLE 5. Convergence rates for $\|e_u\|$ for the BVP (53a) subject to (53c) on uniform meshes having $N = n^2$ square elements with $n = 5, 10, 15, 20$, and 25 using $p = 1$-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	6.5408e-2	NA	6.5320e-3	NA	5.1313e-4	NA	3.2199e-5	NA
100	1.6851e-2	1.9566	8.4573e-4	2.9493	3.2805e-5	3.9673	1.0234e-6	4.9756
225	7.5242e-3	1.9885	2.5225e-4	2.9837	6.5078e-6	3.9894	1.3519e-7	4.9923
400	4.2390e-3	1.9946	1.0666e-4	2.9921	2.0622e-6	3.9948	3.2117e-8	4.9961
625	2.7149e-3	1.9968	5.4670e-5	2.9951	8.4526e-7	3.9969	1.0529e-8	4.9979

element spaces Q^p with $p = 1, 2, 3, 4$. We report the L^2 errors $\|e_u\|$ with their orders of accuracy in Table 4.2. We observe that our UWDG scheme gives the optimal $(p+1)-$th order of the accuracy for this nonlinear problem. The results are in agreement with the theoretical estimates.

We repeat the previous example with all parameters kept unchanged except that we use the mixed Dirichlet-Neumann boundary conditions

(53c)
$$u(0, y) = u(x, 0) = 0, \quad u_x(1, y) = \cos(y), \quad u_y(x, 1) = \cos(x), \quad (x, y) \in \partial\Omega.$$

The L^2 errors $\|e_u\|$ are listed in Table 4.2. Clearly, these results suggest optimal convergence rates. Again, our results are in full agreement with the theoretical results.

Finally, we solve the same problem but using the periodic boundary conditions

$$\begin{aligned} &u(0, y) = u(1, y),\ u(x, 0) = u(x, 1), \\ &u_x(0, y) = u_x(1, y), \quad u_y(x, 0) = u_y(x, 1), \quad (x, y) \in \partial\Omega. \end{aligned} \tag{53d}$$

We present the L^2 errors in Table 4.2. These results indicate that $\|e_u\| = \mathcal{O}(h^{p+1})$. These results are consistent with the theoretical results.

Example 4.3. We consider the following nonlinear reaction-diffusion problem with mixed boundary conditions

$$-u_{xx} - u_{yy} = -\frac{u^3}{u^2+1} - u + g(x, y), \quad (x, y) \in [0, 1]^2, \tag{54a}$$

$$u(0, y) = u(x, 0) = 0, \quad x \in [0, 1], \quad y \in [0, 1], \tag{54b}$$

$$u_x(1, y) = (1-e)(1-y)(e^{y^2} - 1), \quad y \in [0, 1], \tag{54c}$$

$$u_y(x, 1) = (1-e)(1-x)(e^{x^2} - 1), \quad x \in [0, 1]. \tag{54d}$$

TABLE 6. Convergence rates for $\|e_u\|$ for the BVP (53a) subject to (53d) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p =$ 1-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	5.6923e-1	NA	5.8930e-2	NA	4.5929e-3	NA	2.8749e-4	NA
100	1.4879e-1	1.9357	7.5423e-3	2.9659	2.9203e-4	3.9752	9.1045e-6	4.9808
225	6.6666e-2	1.9800	2.2450e-3	2.9887	5.7873e-5	3.9920	1.2019e-6	4.9939
400	3.7605e-2	1.9902	9.4862e-4	2.9945	1.8332e-5	3.9961	2.8547e-7	4.9969
625	2.4098e-2	1.9943	4.8606e-4	2.9966	7.5129e-6	3.9975	9.3580e-8	4.9982

TABLE 7. Convergence rates for $\|e_u\|$ for the BVP (54) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p =$ 1-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	1.2489e-3	NA	8.1907e-5	NA	4.2346e-6	NA	2.1006e-7	NA
100	2.9839e-4	2.0654	1.2501e-5	2.7119	3.2665e-7	3.6964	7.8462e-9	4.7427
225	1.4633e-4	1.7573	3.9643e-6	2.8325	6.8450e-8	3.8543	1.0836e-9	4.8826
400	8.6806e-5	1.8152	1.7215e-6	2.8995	2.2189e-8	3.9158	2.6217e-10	4.9327
625	5.7216e-5	1.8680	8.9459e-7	2.9335	9.2011e-9	3.9449	8.6751e-11	4.9562

TABLE 8. Convergence rates for $\|e_u\|$ for the BVP (54a) subject to (54e) on uniform meshes having $N = n^2$ square elements with $n = 5$, 10, 15, 20, and 25 using $p =$ 1-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
25	8.0274e-3	NA	5.5325e-4	NA	3.1077e-5	NA	1.4830e-6	NA
100	2.1409e-3	1.9067	7.8183e-5	2.8230	2.1761e-6	3.8360	5.1574e-8	4.8457
225	9.9632e-4	1.8865	2.4053e-5	2.9073	4.4318e-7	3.9247	6.9763e-9	4.9338
400	5.7383e-4	1.9179	1.0309e-5	2.9451	1.4198e-7	3.9568	1.6733e-9	4.9628
625	3.7209e-4	1.9413	5.3211e-6	2.9637	5.8519e-8	3.9720	5.5124e-10	4.9761

The function $g(x, y)$ is selected such that the exact solution is given by

$$u(x, y) = (1 - x)(e^{x^2} - 1)(1 - y)(e^{y^2} - 1).$$

We solve (54). For different values of $p \in \{1, 2, 3, 4\}$ and for a series of triangulations consisting of uniform Cartesian meshes having $N = n^2$ squares with $n = 5$, 10, 15, 20, and 25, the L^2 errors and the experimental orders of convergence are computed and presented in Table 4.3. In each case, we observe a convergence ratio of about $r = p+1$ for $p = 1, 2, 3, 4$, as predicted by Theorem 3.1. Finally, we solve the same problem, but using the dirichlet boundary conditions

$$u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0, \quad (x, y) \in \partial\Omega. \tag{54e}$$

We present the L^2 errors in Table 4.3. These results indicate that $\|e_u\| = \mathcal{O}(h^{p+1})$. These results are consistent with the theoretical results.

Example 4.4. In this final example, we consider the following semilinear elliptic problem

$$-u_{xx} - u_{yy} + \sin(u) = 2\sin(x + y) + \sin(\sin(x + y)), \quad (x, y) \in [0, 2\pi]^2, \tag{55a}$$

TABLE 9. Convergence rates for $\|e_u\|$ for the BVP (55) on rectangular meshes having $N = \frac{3n}{2} \times n$ square elements with $n = 4$, 8, 12, 16, and 20 using $p =$ 1-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order	$\|e_u\|$	order
24	5.7572e-1	NA	5.8145e-2	NA	4.5654e-3	NA	2.8624e-4	NA
96	1.4932e-1	1.9470	7.5171e-3	2.9514	2.9157e-4	3.9688	9.0941e-6	4.9762
216	6.6775e-2	1.9848	2.2416e-3	2.9842	5.7831e-5	3.9899	1.2013e-6	4.9923
384	3.7640e-2	1.9927	9.4783e-4	2.9921	1.8325e-5	3.9949	2.8539e-7	4.9962
600	2.4113e-2	1.9956	4.8580e-4	2.9953	7.5109e-6	3.9970	9.3563e-8	4.9978

TABLE 10. Convergence rates for $\|e_u\|_\infty$ for the BVP (55) on rectangular meshes having $N = \frac{3n}{2} \times n$ square elements with $n = 4$, 8, 12, 16, and 20 using $p =$ 1-4.

N	$p=1$		$p=2$		$p=3$		$p=4$	
	$\|e_u\|_\infty$	order	$\|e_u\|_\infty$	order	$\|e_u\|_\infty$	order	$\|e_u\|_\infty$	order
24	4.0606	NA	7.9735e-1	NA	8.3009e-2	NA	6.4848e-3	NA
96	2.0868	0.96040	2.0971e-1	1.9268	1.0645e-2	2.9631	4.1270e-4	3.9739
216	1.3991	0.98603	9.4106e-2	1.9763	3.1706e-3	2.9871	8.1806e-5	3.9914
384	1.0510	0.99446	5.3117e-2	1.9880	1.3402e-3	2.9932	2.5917e-5	3.9956
600	8.4131e-1	0.99728	3.4050e-2	1.9927	6.8684e-4	2.9957	1.0622e-5	3.9973

subject to the mixed Dirichlet-Neumann boundary conditions

$$u(0,y) = \sin(y), \quad u(x,0) = \sin(x),$$
$$u_x(2\pi, y) = \cos(y), \quad u_y(x, 2\pi) = \cos(x), \quad (x,y) \in \partial\Omega. \tag{55b}$$

The exact solution is given by

$$u(x,y) = \sin(x+y).$$

We solve (55a) using the proposed UWDG scheme presented in Section 2. We assume that the mesh consists of $N = \frac{3n}{2} \times n$ rectangular elements with $N = 24$, 96, 216, 384, and 600, where $n = 4, 8, 12, 16, 20$. We compute the UWDG solution u_h in the finite element spaces Q^p with $p = 1, 2, 3, 4$. We report the L^2 errors $\|e_u\|$ and $\|e_u\|_\infty$ with their orders of accuracy in Tables 4.4 and 4.4. We observe that our UWDG scheme gives $\|e_u\| = \mathcal{O}(h^{p+1})$ and $\|e_u\|_\infty = \mathcal{O}(h^p)$. These results are in full agreement with the theoretical estimates given in Theorem 3.1.

5. Concluding remarks

In this paper we presented and analyzed a new ultra-weal discontinuous Galerkin (UWDG) method for the two-dimensional semilinear second-order elliptic problems on Cartesian grids. We performed an *a priori* error analysis in the L^2-norm. The proposed scheme can be made arbitrarily high-order accurate in two-dimensional space. The UWDG solution u_h is shown to converge to the exact solution u at $\mathcal{O}(h^{p+1})$ rate in the L^2-norm, when tensor product polynomials of degree at most p and grid size h are used. Although our error analysis is presented for the two-dimensional problems, it can be readily extended to the three-dimensional problems. The error analysis for the two- and three- dimensional problems on triangular meshes will be investigated in a separate paper. In future work, we will study the superconvergence and error estimation of the UWDG method for nonlinear elliptic problems in multidimensional cases on rectangular meshes. We are also planning to develop *a posteriori* error estimators for the UWDG method applied to two-dimensional parabolic and hyperbolic problems on rectangular and triangular meshes.

Acknowledgement

: This research was supported by the NASA Nebraska Space Grant (Federal Grant/Award Number 80NSSC20M0112). The author would like to thank the anonymous reviewers for the valuable comments and suggestions which improved the quality of the paper.

Conflict of interest: The author declares that he has no conflicts of interest.

Ethical statement: The author declares that the material is his original work, which has not been previously published and is not currently being considered elsewhere.

Data Availability Statement: Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] S. Adjerid, M. Baccouch, A superconvergent local discontinuous Galerkin method for elliptic problems, Journal of Scientific Computing 52 (2012) 113–152.

[2] S. Adjerid, H. Temimi, A discontinuous Galerkin method for higher-order ordinary differential equations, Computer Methods in Applied Mechanics and Engineering 197 (2007) 202–218.

[3] S. Adjerid, H. Temimi, A discontinuous Galerkin method for the wave equation, Computer Methods in Applied Mechanics and Engineering 200 (2011) 837–849.

[4] D. N. Arnold, An interior penalty finite element method with discontinuous elements, SIAM Journal on Numerical Analysis 39 (1982) 742–760.

[5] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM Journal on Numerical Analysis 39 (2002) 1749–1779.

[6] M. Baccouch, A local discontinuous Galerkin method for the second-order wave equation, Computer Methods in Applied Mechanics and Engineering 209–212 (2012) 129–143.

[7] M. Baccouch, Superconvergence of the local discontinuous Galerkin method applied to the one-dimensional second-order wave equation, Numerical methods for partial differential equations 30 (2014) 862–901.

[8] M. Baccouch, A superconvergent local discontinuous Galerkin method for the second-order wave equation on Cartesian grids, Computers and Mathematics with Applications 68 (2014) 1250–1278.

[9] M. Baccouch, Superconvergence of the local discontinuous Galerkin method for the sine-Gordon equation on Cartesian grids, Applied Numerical Mathematics 113 (2017) 124 – 155.

[10] M. Baccouch, Optimal error estimates of the local discontinuous Galerkin method for the two-dimensional sine-Gordon equation on Cartesian grids, International Journal of Numerical Analysis and Modeling 16 (2019) 436–462.

[11] M. Baccouch, Convergence and superconvergence of a local discontinuous galerkin method for semilinear second-order elliptic problems on cartesian grids, Communications on Applied Mathematics and Computation 87:76 (2021) DOI: 10.1007/s42967–021–00123–8.

[12] M. Baccouch, S. Adjerid, *A posteriori* local discontinuous Galerkin error estimation for two-dimensional convection-diffusion problems, Journal of Scientific Computing 62 (2014) 399–430.

[13] M. Baccouch, H. Temimi, Analysis of optimal error estimates and superconvergence of the discontinuous Galerkin method for convection-diffusion problems in one space dimension, International Journal of Numerical Analysis and Modeling 13 (2016) 403–434.

[14] C. E. Baumann, J. T. Oden, A discontinuous *hp* finite element method for convection-diffusion problems, Computer Methods in Applied Mechanics and Engineering 175 (1999) 311–341.

[15] W. Cao, D. Li, Z. Zhang, Optimal superconvergence of energy conserving local discontinuous Galerkin methods for wave equations, Communications in Computational Physics 21 (1) (2017) 211–236.

[16] P. Castillo, B. Cockburn, I. Perugia, D. Schötzau, An *a priori* error analysis of the local discontinuous Galerkin method for elliptic problems, SIAM Journal on Numerical Analysis 38 (2000) 1676–1706.

[17] A. Chen, F. Li, Y. Cheng, An ultra-weak discontinuous Galerkin method for schrödinger equation in one dimension, Journal of Scientific Computing 78 (2) (2019) 772–815.

[18] Y. Cheng, C.-W. Shu, A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives, Mathematics of computation 77 (2008) 699–730.

[19] P. Ciarlet, P. Raviart, General Lagrange and Hermite interpolation in R^n with applications to finite element methods, Arch. Ration. Mech. Anal. 46 (1972) 177–199.

[20] P. G. Ciarlet, The finite element method for elliptic problems, SIAM, 2002.

[21] B. Cockburn, C. W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM Journal on Numerical Analysis 35 (1998) 2440–2463.

[22] P. Degond, H. Liu, D. Savelief, M.-H. Vignal, Numerical approximation of the Euler-Poisson-Boltzmann model in the quasineutral limit, Journal of Scientific Computing 51 (1) (2012) 59–86.

[23] H. Dret, Nonlinear elliptic partial differential equations : an introduction, Springer, Cham, 2018.

[24] D. French, T. Peterson, A continuous space-time finite element method for the wave equation, Mathematic of Computation 65 (1996) 491–506.

[25] D. Gilbarg, Elliptic partial differential equations of second order, Springer, Berlin, 2001.

[26] M. Grote, A. Schneebeli, D. Schötzau, Discontinuous Galerkin finite element method for the wave equation, SIAM Journal on Numerical Analysis 6 (2006) 2408–2431.

[27] M. Grote, D. Schötzau, Optimal error estimates for the fully discrete interior penalty DG method for the wave equation, Journal of Scientific Computing 40 (2009) 257–272.

[28] T. Gudi, N. Nataraj, A. Pani, An *hp*-local discontinuous Galerkin method for some quasilinear elliptic boundary value problems of nonmonotone type, Mathematics of Computation 77 (2008) 731–756.

[29] Q. Han, Elliptic partial differential equations, Courant Institute of Mathematical sciences American Mathematical Society, New York, N.Y. Providence, R.I, 2011.

[30] B. Hille, Ionic channels of excitable membranes, Sinauer Associates, 1984.

[31] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems, Computer Methods in Applied Mechanics and Engineering 107 (1) (1993) 117 – 129.

[32] P. LeSaint, P. A. Raviart, On a finite element method for solving the neutron transport equation, in: C. de Boor (ed.), Mathematical Aspects of Finite Elements in Partial Differential Equations, Academic Press, New York, 1974.

[33] H. Liu, J. Yan, The direct discontinuous Galerkin (DDG) methods for diffusion problems, SIAM Journal on Numerical Analysis 47 (2009) 675–698.

[34] W. H. Reed, T. R. Hill, Triangular mesh methods for the neutron transport equation, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos (1991).

[35] M. F. Wheeler, An elliptic collocation-finite element method with interior penalties, SIAM Journal on Numerical Analysis 15 (1978) 152–161.

[36] S. Yadav, A. K. Pani, Superconvergence of a class of expanded discontinuous Galerkin methods for fully nonlinear elliptic problems in divergence form, Journal of Computational and Applied Mathematics 333 (2018) 215 – 234.

Department of Mathematics, University of Nebraska at Omaha, Omaha, NE 68182, USA
E-mail: mbaccouch@unomaha.edu